Volume 5 of 9 of Submission

FINAL REPORT

Study Title
Equivalency of Microbial and Maize Expressed Cry1F Protein;
Characterization of Test Substances for Biochemical and Toxicological Studies

Data Requirement
Not Applicable

Author
[redacted]

Facility Coordinating Study
Biochemistry Department
Mycogen Corporation
5501 Oberlin Drive
San Diego, CA 92121
Lab Identifier: MYCO98-001

Study Sponsor
[redacted] Product Registration Manager
Mycogen Corporation
5501 Oberlin Drive
San Diego, CA 92121

Testing Facilities
Biochemical Equivalency
Biochemistry Department
Mycogen Corporation
5501 Oberlin Drive
San Diego, CA 92121
Lab Identifier: MYCO98-001

Biochemical – ELISA
Pioneer Hi Bred Intl.
7300 NW 62nd Ave.
Johnston, IA 50134
Lab Identifier: MYCO98-001

Biochemical - N-terminal:
UCSD Biology Dept.
Protein Sequencing Facility
9500 Gilman Dr., Mail Code 0322
La Jolla, CA 92093
Lab Identifier: Myco98-001

Biological #1:
Ricerca, Inc.
Department of Analytical and Biological Sciences
7528 Auburn Rd.
Painesville, OH 44077
Lab Identifier: 7442

Biological #2:
Pioneer Hi Bred Intl.
7300 NW 62nd Ave.
Johnston, IA 50134
Lab Identifier: 980914.ECB

Project ID
MYCO98-001

Study Completed on
October 19, 1998

STATEMENT OF NO DATA CONFIDENTIALITY CLAIMS

No claim of confidentiality is made for any information contained in this study on the basis of its falling within the scope of FIFRA § 10(d) (1) (A), (B), or (C).

These data are the property of Mycogen Corporation, and as such, are considered to be confidential for all purposes other than compliance with FIFRA §10. Submission of these data in compliance with FIFRA does not constitute a waiver of any right to confidentiality that may exist under any other statute or in any other country.

Company: Mycogen Corporation

Company Agent: [redacted]

[redacted]
Product Registration Manager

Date: October 19, 1998

COMPLIANCE STATEMENT

All aspects of this study were in accordance with the Environmental Protection Agency Good Laboratory Practice Standards (40 CFR 160, October 16, 1989) with the following exception:

The N-terminal work was performed at a university laboratory not adhering to GLP standards as a matter of course. For this work, Mycogen QA performed an audit of the facility, and specific arrangements were made to obtain a new column, to perform calibrations frequently and to document these in notebooks.

[redacted] 10/19/98

Study Director Date
Mycogen

[redacted] 10/19/98

Sponsor Date

DOCUMENT DISTRIBUTION

Original	Sponsor
Copy 1	Sponsor
Copy 2	QA Archives
Copy 3	Study Director

APPROVALS

Study Title: Equivalency of Microbial and Maize Expressed Cry1F Protein; Characterization of Test Substances for Biochemical and Toxicological Studies

Coordinating Laboratory: Biochemistry Department
Mycogen
5501 Oberlin Drive
San Diego, CA 92121

[redacted] 10/19/98
Study Director Date

[redacted] 10/19/98
Study Sponsor Date

QUALITY ASSURANCE STATEMENT

STUDY TITLE: Equivalency of Microbial and Maize Expressed Cry1F Protein; Characterization of Test Substances for Biochemical and Toxicological Studies

STUDY NUMBER: MYCO98-001
STUDY DIRECTOR: [redacted]

The Quality Assurance Unit has conducted the following inspections during this study:

Inspection Date	Phase	Reported to Study Director & Management	Auditor
8/25/98	Protocol Review/Approval	10/19/98	Bob Shutter
9/4/98	Biological Testing-*Heliothosis virescens*	10/19/98	Donald Mayer (Ricerca, Inc.)
9/9/98	SDS Page/Western Blot	10/19/98	Bob Shutter
9/11/98	Glycosylation by Electrophorisis	10/19/98	Bob Shutter
9/18/98	Western Transfer	10/19/98	Robert Sibley
9/25/98	Invitro Digestibility	10/19/98	Robert Sibley
9/28/98	N-Terminal Analysis	10/19/98	Robert Sibley
9/30/98	Colormetric Test, Total Protein Content	10/19/98	Robert Sibley
10/19/98	Raw Data vs. Draft Report	10/19/98	Bob Shutter
10/19/98	Final Report	10/19/98	Bob Shutter

This report has been reviewed by the Quality Assurance Unit in accordance with the EPA GLP Standards set forth in 40 CFR §160.35(b)(6). This report describes the methods and procedures used in this study, and the reported results accurately reflect the raw data.

[redacted] 10/19/98
Quality Assurance Supervisor Date

TABLE OF CONTENTS

List of Abbreviations

BCW	Black cutworm, Agrotis ipsilon
BSA	Bovine serum albumin
Bt	Bacillus thuringiensis
CAPS	3-[cyclohelylamine]-1-propanesulfonic acid
CEW	Corn earworm, Helicoverpa zea
Cry1F	designation of subject toxin based on older nomenclature
Cry1Fa2	designation of subject toxin based on revised nomenclature
ECB	European corn borer, Ostrinia nubilalias
EDTA	Ethylene diamine tetraacetic acid
ELISA	Enzyme Linked Immunosorbent Assay
FAW	Fall armyworm, Spodoptera frugiperda
IgG	Immunoglobulin G
LC_{50}	concentration of toxin causing 50% mortality as determined by statistical evaluation of dose-mortality curves in insect bioassay
MR872	strain designation for Pseudomonas fluorescens containing full-length gene for CryIF protein
MWCO	Molecular weight cut off
NBT/BCIP	Nitroblue tetrazolium chloride / 5-bromo-4-chloro-3'-indoylphosphate p-toluidine salt
PBS	Phosphate buffered saline
ppm	Parts per million
PS81I	Bacillus thuringiensis subspecies aizawai strain PS81I
PVDF	Polyvinylidene difluoride
SDS PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
SGF	Simulated gastric fluid
TBS	Tris buffered saline
TBW	Tobacco budworm, Heliothis virescens
Tris	tris(hydroxymethyl)aminomethane

SUMMARY

CryIF toxin extracted from corn or from a microbial source was found to react similarly in a variety of biochemical and biological tests. This equivalency supports the use of microbially derived toxin in certain toxicology tests, which is necessary due to the lack of abundance of plant derived toxin. These data also demonstrate that the plant-produced toxin has characteristics consistent with the protein expected from the inserted transgene with respect to molecular weight, immunoreactivity, lack of post-translational modification (glycosylation), N-terminal amino acid sequence and spectrum of bioactivity.

Toxin extracted from transgenic corn demonstrated immunoreactive species corresponding to proteins of ~68 and ~65 kDa. The ~68 kDa species corresponds to the expected molecular weight of the gene product from the transgene inserted into corn. The ~65 kDa species was predominant, and corresponds to the expected product of limited N-terminal proteolysis, potentially arising during the purification procedure from plant material. The ~65 kDa species extracted from corn migrated in SDS-PAGE with a mobility corresponding to that of the trypsinized, microbially derived CryIF and of authentic CryIF standard (also trypsinized). There were no other immunoreactive species apparent in Western blots of CryIF derived from corn. The material yielded a quantification of 0.158% toxin by ELISA.

Toxin extracted from a microbial source and truncated by trypsin revealed a predominant species corresponding to a protein of ~65 kDa by SDS-PAGE with Coomassie Brilliant Blue staining. This is the anticipated product of trypsinolysis of the cryIF gene product. Contaminating bands in low abundance were observed by Coomassie staining at mobilities corresponding to ~100, 45, 20 and 15 kDa. Western blotting demonstrated a predominat immunoreactive species at a mobility corresponding to a protein of ~65 kDa, plus other immunoreactive species at mobilities corresponding to ~25 and ~20 kDa. This material yielded a quantification of 11.4% toxin by SDS-PAGE. ELISA was not attempted due to the low molecular weight immunoreactive species.

Though the N-terminus of plant-derived CryIF was blocked significantly, a five amino acid sequence corresponding to the expected N-terminus of proteolytically cleaved CryIF was obtained. This observed sequence was ^{28}STGRL, where the superscript refers to the amino acid residue in the subject protein. This is the expected N-terminus following cleavage by a trypsin-like activity (^{20}VEILNEER$\Downarrow^{28}$STGRLPLD). Two sequences and a fragment were obtained on the microbial derived CryIF; ^{28}STGRLPL, ^{43}FLLSEFV, and ^{33}PLDI. Likewise, the fragment beginning at ^{43}F follows a trypsin site (^{39}SLTR$\Downarrow^{43}$FLLSEFVP).

These two materials performed indistinguishably in bioassays on susceptible species when normalized by reported toxin content.

	Top Load Bioassays at Ricerca			Diet Incorporation at Pioneer
	LC_{50} ng CryIF per cm^2			LC_{50} ppm CryIF
Source of toxin	European Corn Borer	Tobacco Budworm	Fall Armyworm	European Corn Borer
Corn CryIF	0.58	1.74	2.21	0.11
Microbial CryIF	0.58	1.88	2.49	0.14

The rank order (LC_{50}) of susceptibility for species tested for both toxin preparations was:

most susceptible: ECB > TBW = FAW >>CEW >BCW :least susceptible

The control plant extract had no detectable CryIF toxin or crossreactivity by ELISA and produced negligible effects in bioassays. Neither corn derived nor microbially derived CryIF showed evidence of glycosylation. Microbially derived CryIF was very susceptible to digestion in simulated gastric conditions in the presence of pepsin. Molar ratios of CryIF to pepsin of 1:100 yielded complete digestion in 1-5 minutes.

INTRODUCTION

An insecticidal delta-endotoxin originally identified in Bacillus thuringiensis subspecies aizawai strain PS81I is the subject of this report. The designation of this protein toxin is Cry1F. It is of commercial interest because CryIF has activity on the European corn borer (ECB, Ostrinia nubilalis), one of the major pests of maize in North America and western Europe. Data collected here are for the purpose of characterizing the biochemical and toxicological properties of this insecticidal endotoxin in support of a regulatory submission.

Study Title

Equivalency of Microbial and Maize Expressed Cry1F Protein; Characterization of Test Substances for Biochemical and Toxicological Studies

Study Identifier

MYCO98-001

Study Sponsor

[redacted], Product Registration Manager
Mycogen
5501 Oberlin Drive
San Diego, CA 92121

Testing Facilities

Purification/Biochemical Assays:	Mycogen Corporation 5501 Oberlin Drive San Diego, CA 92121
Biochemical - ELISA:	Pioneer Hi Bred Intl. 7300 NW 62nd Ave. Johnston, IA 50134
Biochemical - N-Terminal:	UCSD Biology Dept. Protein Sequencing Facility 9500 Gilman Dr., Mail Code 0322 La Jolla, CA 92093
Biological #1:	Ricerca, Inc. Department of Analytical and Biological Sciences 7528 Auburn Rd. Painesville, OH 44077
Biological #2:	Pioneer Hi Bred Intl. 7300 NW 62nd Ave. Johnston, IA 50134

Study Director and Principle Investigators involved in the study

Study Director: [redacted]
Mycogen
[redacted]

Biological Principal Investigators: [redacted]
Ricerca
[redacted]
[redacted]
Pioneer Hi-Bred
[redacted]

Biochemical Principal Investigators:
ELISA: [redacted]
Pioneer HI-Bred
[redacted]

N-Terminal sequencing: [redacted]
UCSD
[redacted]

Biochemical tests: [redacted]
Mycogen
[redacted]

Biochemical tests: [redacted]
Mycogen
[redacted]

Significant Study Dates

Protocol Approved 8/25/98

Experimental Sections	Initiated	Completed	Report
Biochemical Assays	8/28/98	10/12/98	10/19/98
ELISAs	8/27/98	9/1/98	10/7/98
N-Terminal sequencing	9/28/98	10/01/98	10/17/98
Biological testing, Pioneer	9/14/98	9/21/98	10/13/98
Biological testing, Ricerca	9/4/98	10/8/98	10/12/98

Final Report Drafted 10/12/98
Final Report Completed 10/19/98

OBJECTIVES

The objective of this study was to determine and compare selected biochemical and biological properties of maize-expressed Cry1F protein and bacterially expressed Cry1F protein. It is necessary to use bacterially expressed CryIF for certain toxicology studies, due to limited abundance of the plant-expressed toxin, therefore demonstration of the equivalence of the toxin from these two sources is necessary. The biochemical and biological data generated during this study will also be used to provide characterization information for biochemical and toxicological studies performed to support the registration of transgenic corn expressing this protein.

Justification of Test Systems

The test systems for the biochemical portion of this study are widely utilized and accepted in the field of protein characterization. The procedures are consistent with those known to have been used by other applicants for the submission of data to support registration of Bt toxins in other transgenic plants. The insect species used for the biological portion of this study vary in susceptibility to the target protein, thus allowing for observation of significant shifts in bioactivity due to the expression system.

Overview of Procedures

Experimental Design

The characterization of toxins from the bacterial source and the plant source were conducted using two types of analyses; biochemical and biological.

The biochemical tests performed on both materials were:

- Molecular weight determination by SDS PAGE
- Total protein content by Lowry
- Immunoreactivity by Western analysis and ELISA
- Glycosolation by electrophoresis and staining
- N-terminal sequencing by Edman analysis

The biological activity of both materials was determined on sensitive and insensitive insects:

Sensitive	Insensitive
Ostrinia nubilalis (European corn borer)	Agrotis ipsilon (Black cutworm)
Heliothis virescens (Tobacco budworm)	Helicoverpa zea (Corn earworm)
Spodoptera frugiperda (Fall armyworm)	

Finally, in vitro digestibility of the microbially expressed toxin in simulated gastric fluid was determined.

Test Methods - Biochemical

The following methods were used during the biochemical portion of the study.

- Sodium dodecyl sulfate polyacrylamide electrophoresis (SDS PAGE)
- Colorimetric test for total protein content (Lowry)
- Antibody reactivity in solution (ELISA)
- Colorimetric test for total protein content (Bradford) incidental to the ELISA
- Western blotting of toxin following SDS PAGE
- Susceptibility of Cry1F protein to proteolytic degradation in simulated mammalian gastric fluid (SGF) containing pepsin
- N-terminal sequence analysis by Edman degradation coupled with automated gas-phase sequencing of toxins which had been separated by SDS PAGE and blotted by Western transfer

- Glycosylation by electrophoresis and staining with the Immuno-Blot Kit for Glycoprotein Detection (Bio-Rad)

Test Methods – Biological at Ricerca

Experimental Design: Dilutions of test materials were formulated and dispensed onto the surface of artificial diet contained in 96 well microtiter plates. Neonate insect larvae were exposed to the treated diet and mortality was assessed. The concentration causing 50% mortality (LC_{50}) was determined by statistical analysis of the resulting data. The following species were tested in this format:

- Ostrinia nubilalis (European corn borer)
- Heliothis virescens (Tobacco budworm)
- Spodoptera frugiperda (Fall armyworm)
- Agrotis ipsilon (Black cutworm)
- Helicoverpa zea (Corn earworm)

Test Methods – Biological at Pioneer

Experimental Design: Bioassays were conducted against European corn borer larvae (ECB) to determine the LC_{50} of the test materials. Neonate larvae are placed on insect diet into which had been incorporated a series of rates of the toxin protein. The larvae are allowed to grow and develop on the diet, after which percent mortality was determined for each treatment and the LC_{50} then calculated from these data.

Statistical Evaluation of Data

Biochemical Data

Means and standard deviations were generated as appropriate. Correlation coefficients for standard curves were determined. No other statistical tests were performed on the biochemical data.

ELISA Data

Each sample was analyzed at least twice in duplicate. Interpolated values were derived only from standard curves which passed the quality control (QC) criteria described in SOP PHIAP-SP-0006/02 and from samples wherein the duplicate sample wells/analysis had a % coefficient of variation (CV) of <15%.

Bioassays – Ricerca

Percent mortality was calculated at each dose level. Abbott's formula was used to correct for control mortality.

Abbott's formula:

$$\%\ \text{Corrected Mortality} = \frac{100 * (\%\ \text{mortality}_{\text{treatment}} - \%\ \text{mortality}_{\text{control}})}{100 - \%\ \text{mortality}_{\text{control}}}$$

Probit analysis using POLO-PC (LeOra Publishing, Berkeley, CA) was used to calculate LC_{50} values and slopes.

Bioassays – Pioneer

The LC_{50} was determined by plotting the protein rate (x axis) vs. % mortality (y axis) using linear regression. Using the slope formula and JMP software (SAS Institute), the LC_{50} is calculated as $y = m(x) + b$ where $y = 0.50$ for 50% mortality, m = slope of the best fit line calculated by the software, b = y intercept (also calculated by the software). The LC_{50} was obtained by solving for x.

TEST, CONTROL AND REFERENCE SUBSTANCES

The subject toxins were purified from transgenic corn and from a microbial production system. The details of production for the toxin extracted from corn are found in Appendix A. Briefly, corn plants were harvested, frozen and shipped to the laboratory. Frozen tissue was homogenized in an extraction buffer and clarified by filtration and centrifugation. The resulting soluble material containing the CryIF toxin was concentrated by ammonium sulfate precipitation. The ammonium sulfate pellet was collected by centrifugation, dialyzed to remove salt, and lyophilized. An isogenic corn line was used to prepare a control powder lacking CryIF. Plant material from that line was treated in an identical fashion to produce a control powder. The details of this production also are found in Appendix A.

A bacterial production system was used to produce larger quantities of CryIF toxin. A gene encoding the full-length CryIF toxin was placed in Psuedomonas fluroescens strain MR872 (amino acids 1 – 1174). This strain was grown in large-scale fermentation and the toxin extracted from concentrated cell pellets. The details of production of the bacterial CryIF powder are listed in Appendix C. Briefly, cell pellets were enzymatically lysed and full-length toxin was isolated as insoluble inclusions. These inclusions were washed extensively, then suspended in alkaline conditions and treated with trypsin to truncate the toxin. The soluble, truncated toxin was then diafiltered and finally concentrated by lyophilization. The bulk bacterial powder was produced by repetitively applying the same procedure to small batches of cell pellets (80 – 450 grams each). The resulting powders were combined into two batches and designated as the lots indicated below.

The trypsin-truncated microbially derived toxin represents amino acids 28 to 612 of the full CryIF toxin sequence. The toxin gene inserted in transgenic corn codes for amino acids 1- 605. Thus, the truncated toxin produced by trypsin treatment of the full length bacterial toxin adequately represents that which is expected to be found in plant material.

Test Substances

Plant CryIF powder

Lot Number:	1568-022B
Date Lot Created:	7/15/98
Description of production:	Appendix A
Quantified value:	0.158% CryIF toxin by ELISA
Quantification Certificate of Analysis:	Appendix B
Storage:	Room temperature

- Synonyms used in studies

Biochemical equivalence tests (Mycogen):	CryIF corn
Biological (Ricerca):	Plant CryIF, TCR 198-00092
Biological (Pioneer):	Bot. #2, CryIF plant tissue
ELISA (Pioneer):	Bottle 2, CryIF plant extract
N-terminal (UCSD):	FC

Bacterial CryIF powder

Lot Number:	1599-39
Date Lot Created:	8/19/98
Description of production:	Appendix C
Quantified value:	11.4% CryIF toxin by SDS PAGE
Quantification Certificate of Analysis:	Appendix D
Storage:	4°C

- Synonyms used in studies

Biochemical equivalence tests (Mycogen):	PHB powder, microbial CryIF
Biological (Ricerca):	Bacterial CryIF, TCR 198-00090
N-terminal (UCSD):	PHB

Lot Number:	1599-45
Date Lot Created:	8/19/98
Description of production:	Appendix C
Quantified value:	11.4% CryIF toxin by SDS PAGE
Analysis Report:	Appendix K
Storage:	4°C

- Synonyms used in studies

Biochemical equivalence tests (Mycogen):	Powder 1, microbial CryIF
N-terminal (UCSD):	P1

Control Substances

Plant extract from non-expressing line (powder)

Lot Number:	1568-022A
Date Lot Created:	7/15/98
Description of production:	Appendix A
Quantified value:	Below limit of detection of ELISA, Less than 0.00004% CryIF toxin
Quantification Certificate of Analysis:	Appendix E
Storage:	Room temperature

- Synonyms used in studies

Biochemical equivalence tests (Mycogen):	control corn
Biological (Ricerca):	Control plant tissue, TCR 198-00091
Biological (Pioneer):	Bot. #1, neg. control
ELISA (Pioneer):	Bottle 1, CryIF plant extract

Bovine Serum Albumin (BSA)

Product Number:	23208
Lot Number:	98011372
Description of production:	Obtained from Pierce, Rockford, IL
Quantified value:	7 individual standards from 125 micrograms per ml to 2000 micrograms per ml
Storage:	4°C
Stability:	Expire Jan., 2000
Use in study:	SDS PAGE

Product Number:	23209
Lot Number:	98051269A
Description of production:	Obtained from Pierce, Rockford, IL
Quantified value:	2.0 mg per ml
Storage:	Room temperature
Stability:	Expire 8/28/99
Use in study:	Lowry

Non Bacillus thuringiensis purified CryIF truncated protein

Lot Number:	8/13/98
Date Lot Created:	082597
Description of production:	Obtained final formulation from Pioneer, initial production described in Mycogen Notebook 1496, page 52
Quantified value:	1.4 mg per ml
Storage:	-20°C
Use in study:	SDS PAGE, ELISA

Reference Substances

Dipel® 2X

Lot Number:	Ricerca Identifier JK-0
Date Lot Created:	unknown
Description of production:	Locally purchased by Ricerca
Quantified value:	unknown
Storage:	Room temperature
Use in study:	Bioassay control

Stability of the microbial and plant derived toxins is under evaluation in a long-term characterization study. Dry powders of Bt toxins are generally considered stable in the art.

CHANGES TO APPROVED PROTOCOL MYCO98-001

Deviations

Relevant documents in Appendix F.

Mycogen test facility

1. 9/18/98 Section of protocol affected, 9.6.
   Deviation: Biorad PVDF membrane was substituted for Immobilon Psq membrane, and 0.1% Coomassie Brilliant Blue stain used instead of 1% Ponceau S.
   Reason for deviations: The Immobilon membranes were past the manufacturers expiration date so an equivalent membrane was substituted. Coomassie Brilliant Blue is considered more sensitive than Ponceau S so the substitution was made to increase detection ability on the blots for N-terminal sequencing.
   Documentation: Deviation form dated 9/21/98
   Impact on study: None.
2. 9/9/98 Section of protocol affected, 7.4.1.
   Deviation: Storage temperature for bulk containers of microbial powder lots 1599-39 and 1599-45 deviated from the label recommendation for a period up to three days.
   Reason for deviations: Power to the entire research building was interrupted for several hours due to a construction project. The project involved rewiring electrical equipment, including the cold room (4°C), over a 3-day holiday weekend. A critical connection to one apparatus in the cold room was not properly rewired resulting in the loss of cooling. The rewiring occurred the evening of 9/5/98 and the temperature deviation was discovered the morning of 9/8/98. The cold room was at ambient (23°C). The wiring anomaly was corrected and temperature returned to normal 9 AM on 9/8/98.
   Documentation: Letter from Principle Investigator to Study Director dated 9/9/98
   Impact on study: Minimal to none. (1) All samples used in biological or toxicological studies had already been shipped. (2) The decision to store the lyophilized powder at 4°C was not based on a need for the sample to be chilled for short-term stability. Mycogen routinely stores powders in the cold for long-term storage, and either uses them from cold storage (cycles from storage to room temperature, then back) or prepares 'working quantities' which are kept at room temperature (DSOP 6004, Storage of Standard Powders). (3) Since the bulk powders were not exposed to excessive temperature, no deleterious effect is anticipated.

Ricerca test facility

3. 9/30/98 Section of protocol affected, 10.3.2.
   Deviation: With corn earworm and black cutworm the criteria established (at least 2 doses producing mortality >50%) could not be achieved with the plant CryIF (1568-022B).
   Reason for deviations: The insensitivity of the insect to the toxin coupled with the toxin concentration in the plant material. Could not deliver sufficient toxin in the dosage.
   Documentation: Correspondence file dated 9/30/98 from Study Director to Principal Investigator at Ricerca
   Impact on study: None. These species were expected to be insensitive. Reasonable dose responses were obtained with the bacterial powder, and the Dipel control validated the test.

Pioneer facility

4. 8/27/98, 8/31/98 and 9/1/98. Section of protocol affected. 9.3.
   Deviation: Extraction procedure used for test powders deviated from Pioneer SOP # PHIAR-SM-0001/01.
   Reason for deviation: Study director requested that powders be handled via quantitative dilution in

volumetric glass ware.
Documentation: Letter to Principal Investigator dated 8/25/98 and Pioneer "Observations and Remarks" sheets for each assay.
Impact on study: Positive, improves quantitative results of analysis.

5. 8/25/98. Section of protocol affected. 9.3
Deviation: Microbial CryIF (Lot 1599-39) was not subjected to ELISA quantification as was the Plant derived CryIF (Lot 1568-022B).
Reason for deviation: During pilot tests immediately preceding implementation of the Study it was observed that there were potential cross reactions with the antibody in the ELISA, thus yielding suspect quantification.
Documentation: Correspondence file. Microbial test material (Lot 1599-39) was not submitted by Sponsor for ELISA. See Results and Discussion section for more information.
Impact on study: Minimal, SDS PAGE analysis yielded quantification for this material.
6. 9/14/98 Section of protocol affected. 10.
Deviation: Pioneer insect bioassay unit added to Study.
Reason for deviation: Ricerca management notified Sponsor that their facility would not be available in the future for biological testing of Bt materials. Mycogen was employing Ricerca for long-term characterization of stability of microbial test material, and the loss of Ricerca therefore jeopardized the study. Inclusion of a new test facility at the inception of the work allowed for generation of early time point data. Pioneer facility has capability for ECB dose-response studies, which are adequate for the objective.
Documentation: Adequate study designs were available under coincident Study Protocol 98-08-RA-GLP-009 entitled "Characterization of Corn Tissues (pollen, grain, leaves) Expressed CryIF Protein, sponsored by Mycogen Corporation and dated 8/4/98. Sponsor arranged for additional bioassay space for the subject proteins of this Study.
Impact on study: Positive, provides independent assessment of toxin equivalency and continues opportunity for long-term characterization of the protein.

Amendments

Relevant documents in Appendiz G.

1. 9/2/98. Section of protocol affected. 10.4.
Change: protocol statement of use of 1% BSA changed to 0.1% BSA
Reason for change: The entry of 1% in the approved protocol was a typo, the correct entry should be 0.1%.
Documentation: Amendment sheet signed 9/2/98 by Study Director.
Impact on study: None. No tests were performed at the incorrect dilution of 1%.
2. 10/12/98. Section affected. 10.2.2
Change: protocol amended to indicate that all insects would be purchased from French Ag. Res., Lamberton, MN.
Reason for change: Anticipated in protocol, amendment clarifies exact source of insects.
Documentation: Amendment sheet signed 10/12/98 by Study Director.
Impact on study: None.
3. 10/12/98. Section affected. 10.1
Change: protocol amended to indicate that assessments would be made after six days for all insects tested.
Reason for change: Anticipated in protocol, amendment clarifies exact duration of test.
Documentation: Amendment sheet signed 10/12/98 by Study Director.
Impact on study: None.
4. 10/12/98. Section affected. 10.3.2
Change: protocol amended to indicate the exact doses in the Ricerca bioassays.

Reason for change: Anticipated in protocol, amendment clarifies exact dilution series of tests.
Documentation: Amendment sheet signed 10/12/98 by Study Director.
Impact on study: None.

5. 9/24/98. Section affected. Certificates of Analysis
Change: Certificates of Analysis of Lots 1568-022B (CryIF from Corn) and 1568-022A (control corn extract) that has been issued on 9/11/98 were amended due to a computation error.
Reason for change: A conversion factor was applied twice, resulting in a 10-fold error in the computed result.
Documentation: Amendment to Certificate of Analysis signed 9/24/98 by Study Director. Correspondence file.
Impact on study: None. The bioassays were not affected by this because the Ricerca tests were set up based on weight of powder, irrespective of estimated toxin content, and Pioneer tests, which were based on toxin estimates, were setup with the corrected value as discovered 9/14/98.

METHODS

Biochemical

SDS-PAGE

Sodium dodecyl sulfate polyacrylamide electrophoresis (SDS PAGE) was performed essentially as described by Laemmli (1970) using Bio-Rad Laboratories (Hercules, CA) reagents and molecular weight standards (Broad Range, no. 161-031) according to kit instructions. For quantitative SDS PAGE, prediluted BSA standards (Pierce 23208) were used according to manufacturers instructions. Sample preparations were mixed 1:2 with Sample Buffer and boiled 5min, then were subjected to SDS PAGE using 10% polyacrylamide gels (Bio-Rad Laboratories, Hercules, CA). Electrophoresis was conducted at 200V 35 – 60 min according to instructions. The proteins were visualized by staining in a solution of Coomassie Brilliant Blue R-250 (Bio-Rad Laboratories, Hercules, CA). Quantitative densitometry was performed on a Molecular Dynamics 325 scanning laser densitometer using ImageQuant version 3.3 software (Molecular Dynamics, Sunnyvale, CA). Integration was performed in the 'volume' mode, and quantification was performed by linear regression analysis of amount of protein loaded (x axis) versus integrated volume (y axis) in Excel 5 on a PC. Photodocumentation was performed on a UVP ImageStore 7500 CCD system (Ultra-Violet Products, Upland, CA).

Colorimetric test for total protein (Lowry)

Protein content was determined using a Modified Lowry Protein Assay (Pierce Chem. Co., Rockford, IL). The kit differs from the standard Lowry assay in that the alkaline cupric sulfate - tartrate reagent has been modified to contain potassium iodide. The modification produces a more stable stock solution. Various dilutions of the sample extracts were diluted in water and mixed with the modified Lowry alkaline cupric sulfate - tartrate reagent. Following a 10 min incubation at room temperature, 0.1ml of 1.0N Phenol Reagent was added and incubation resumed at room temperature for 30 min. The samples were then read at 750nm in a UV-Vis Scanning Spectrophotometer (Shimadzu UV-2101PC). The protein concentrations for the unknown samples were calculated from a standard curve (absorbance vs. standard concentration). The standard curve was fit by a quadratic equation. Bovine serum albumin (no. 23209, Pierce Chem. Co., Rockford, IL) served as the standard.

Antibody reactivity in solution (ELISA)

Antibody reactivity in solution was performed by ELISA to estimate the concentration of toxin in preparations using authentic Cry1F as a standard. Samples were evaluated by SOP #PHIAR-SM-0001/01 and PHIAR-SP-0001/01. The former SOP describes a sandwich ELISA using polyclonal

rabbit anti-CryIF for both the capture and detection phase. The detecting antibody is a biotinylated derivative of the same polyclonal rabbit anti-CryIF reagent. The latter SOP is for a microtiter version of the Biorad Protein Assay based on the Bradford colorimetric method. The only modifications to these SOPs was the original dilution phase which used volumetric glassware in place of standard microfuge tubes (described above in Deviations).

Western Blotting

Proteins separated by SDS PAGE were blotted and probed with rabbit anti-CryIF antibody essentially as described Towbin (see references in Investigator's Report by George Schwab). SDS PAGE was performed as described above with the addition of Kaleidoscope prestained molecular weight standards (no. 161-0324, Bio-Rad Laboratories, Hercules, CA) to facilitate molecular weight estimation following blotting. Proteins from the SDS gel were transferred to Protran nitrocellulose (Schleicher & Schuell, Keene, NH) in the presence of Towbin Transfer Buffer, then the unoccupied binding sites on the membrane were blocked with a 1% (w/v) solution of Rad-Free blocking powder (Schleicher & Schuell, Keene, NH) in PBS. The membranes were probed with a 1:1000 dilution of rabbit polyclonal anti-Cry1F antibody (lot number 091796-P, Pioneer HiBred, Johnston, IA), by incubating 1hr at room temperature in the blocking solution described above. Immunoblots were developed by incubation with an alkaline phosphatase-conjugated goat anti-rabbit (no. A3687, Sigma Chemical Co., St. Louis, MO.) at a 1:5000 dilution in the blocking solution described above. Following extensive rinsing in PBS, the Western blot was colorimetrically developed by a 10min incubation in 10ml nitro-blue tetrazolium chloride/5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt (1-Step NBT/BCIP, Pierce Chem. Co., Rockford, IL).

Susceptibility of CryIF to proteolytic degradation in simulated gastric fluid with pepsin.

Crystallized pepsin A [EC 3.4.23.1] from porcine gut mucosa was obtained from Sigma Chemical Co., St. Louis, MO (no. P6887). The assay was performed in a 1.5ml eppendorf polypropylene microfuge tube containing 90ul 0.15M HCl to which was added 10ul of a 5mg per ml solution of trypsinolyzed CryIF from microbial lot 1599-045. For the time-course of digestion, 10ul of a 50ug per ml solution of pepsin was added and allowed to incubate 0, 1, 2, 4, 10, 15, 30 and 60min. For the concentration-dependence of digestion experiments, 10ul of a pepsin solution serially diluted from a 5.0mg per ml stock pepsin solution (0.5mg, 50ug, 5.0ug, 0.5ug, 50ng and 5.0ng per ml) were added and incubated 5min at room temperature. These amounts correspond to a 1:1, 1:10, 1:100, 1:1000, 1:10000: 1:100000 and 1: 1000000 ratio of pepsin:CryIF (w/w) or 1:1.9, 1:18.8, 1:188, 1:1883, 1:18830, 1:188300 and 1:1883000 molar molar ratio of pepsin:CryIF2. Since pepsin is unstable above pH 6.0, the reactions were terminated by the addition of 2M Tris base to a final concentration of 167mM, imparting a final pH of 7.0 to the sample. Samples were analyzed by SDS PAGE as described above.

N-Terminal analysis

SDS PAGE and Western blotting were as described above with the following exceptions. Transblot PVDF (no. 162-0180, Bio-Rad Laboratories, Hercules, CA) was used in place of nitrocellulose and CAPS Transfer Buffer was used in place of Towbin Transfer Buffer. Material from plant (Lot 1568-022B) was concentrated by affinity chromatography using anti-CryIF antibody. Following transfer of the proteins from the SDS gel, the membrane was rinsed in distilled water and the proteins of interest were localized by staining with 0.1% (w/v) Coomassie Brilliant Blue R-250 in 50% (v/v) methanol for 5 min at room temperature followed by partial destaining in several changes of 50% (v/v) methanol containing 10% (v/v) acetic acid. The membrane was allowed to dry. The bands of interest were excised from the membrane and subjected to Edman degradation coupled with automated gas-phase sequencing at the UCSD Protein Sequencing Facility (La Jolla, CA).

Glycosylation

Potential glycosylation of CryIF in planta was examined by the Biorad Immun-Blot kit for Glycoprotein detection (170-6490) according to manufacturers instructions. Samples were first treated to SDS PAGE and transfer Protran nitrocellulose (Schleicher & Schuell, Keene, NH) as described in the preceding sections. Ovalbumin was used as a positive glycoprotein standard. The nitrocellulose was washed with PBS at room temperature, then immersed in 10ml of Reagent A and incubated in the dark at room temperature for 20min with gentle agitation. The membrane was washed three times with PBS at room temperature and immersed in 10ml Biotinylation Solution for 60min at room temperature with gentle agitation. The membrane was washed three times with TBS at room temperature for 10min each with gentle agitation, then incubated in 10ml Blocking Solution overnight at 4°C. The next day the membrane was washed three times with TBS at room temperature and immersed in 10ml Conjugate Solution and incubated at room temperature for 60min with gentle agitation. The membrane was washed three times with TBS at room temperature and carbohydrate was localized by immersing the membrane in the alkaline phosphatase detection reagent, 10ml nitro-blue tetrazolium chloride/5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt (1-Step NBT/BCIP, Pierce Chem. Co., Rockford, IL) for 10min at room temperature. The reaction was stopped by rinsing in water and drying the membrane.

Biological -- Ricerca

Test Systems at Ricerca

Route of Administration

The bioassays at Ricerca employed a diet surface overlay technique, sometimes called a "top-load" delivery of toxin. Molten diet was dispensed into 96-well microtiter dishes and allowed to solidify. Test materials were diluted in phosphate buffer containing 0.1% BSA and then pipetted on the surface of the diet and allowed to dry.

System Description

Eggs of the following insects were purchased from French Agricultural Research, Lamberton, MN. Ostrinia nubilalis (European corn borer), Heliothis virescens (Tobacco budworm), Spodoptera frugiperda (Fall armyworm), Agrotis ipsilon (Black cutworm), and Helicoverpa zea (Corn earworm). Upon arrival, eggs were place in an incubator and maintained until needed. Prior to use, eggs were sterilized in a bleach solution and transferred to the Insectary and maintained till hatch. Artificial diet was prepared and dispensed into microtiter dishes. Test materials, including Dipel® 2X as a positive control, were diluted and pipetted onto the surface of the agar. One dose was tested in 48 wells, and neonate larvae were infested one to a well. Thus, there was a target of 48 insects per dose per assay. Trays were covered with a mylar film and placed in an incubator for the duration of the assay. During scoring, larvae that were clearly dead or stunted were classified as dead.

Dosages at Ricerca

Dipel

Rate

ng Powder/cm2	TBW	FAW	ECB	CEW	BCW
0	X	x	x	x	X
0.1			x	x	
0.3	X	x	x	x	X
1	X	x	x	x	X
3	X	x	x	x	X
10	X	x	x	x	X
30	X	x	x	x	X
100	X	x			X

Bacterial
Cry1F

Rate

Ng Powder/cm2	TBW	FAW	ECB	CEW	BCW
0	X	x	x	x	X
0.1		x	x		
0.3	X	x	x		
1	X	x	x	x	X
3	X	x	x	x	X
10	X	x	x	x	X
30	X	x	x	x	X
100	X	x	x	x	X
300	X			x	X
500		x	x		X
1000	X			x	X
3000				x	

Plant Cry1F

Rate

Ng Powder/cm2	TBW	FAW	ECB	CEW	BCW
0	X	x	x	x	X
3	X	x	x		
10	X	x	x	x	X
30	X	x	x	x	X
100	X	x	x	x	X
300	X	x	x	x	X
1000	X	x	x	x	X
3000	X	x	x	x	X
5000	X	x	x	x	X
10000				x	X

Doses were of weight of powder delivered to the assay arena. Conversion of the weight of powder to the amount of toxin delivered was made by computation from the Certificate of Analysis. The plant control extract was used at the same weight as was the CryIF plant extract.

Duration

Mortality was assessed after six days.

Biological -- Pioneer

Test Systems at Pioneer

Route of Administration

The bioassays at Pioneer employed a diet incorporation technique for delivery of toxin. Molten diet is mixed with dilutions of the samples, dispensed into microiter dishes and allowed to solidify. Test materials were diluted in phosphate buffer prior to mixing with diet.

System Description

Eggs of Ostrinia nubilalis (European corn borer) were obtained from the Pioneer ECB laboratory and held for hatch. One dose was tested in 14 wells, and neonate larvae were infested one to a well. Thus, there was a target of 14 insects per dose per assay. Trays were covered with a perforated cover and placed in an incubator for the duration of the assay.

Dosages at Pioneer

The plant extracted toxin, the plant control and the microbial toxin were tested at rates of 0.03, 0.05, 0.07, 0.1, 0.3, and 0.5 ppm of toxin, as determined by the ELISA and SDS PAGE.

Duration

Mortality was assessed after seven days.

RESULTS

Biochemical

Molecular weight and immunoreactivity

SDS PAGE of plant derived CryIF (Lot 1568-022B) and the corresponding control extract from non-expressing corn (Lot 1568-022A) were nondescript following Coomassie Brilliant Blue staining (Appendix I, Report by George Schwab, Figure 2, lanes 4 and 6). This was expected given the anticipated abundance of CryIF toxin in plant extracts. Microbially derived CryIF (Lots 1599-39 and 1599-45) yielded a predominant species with the expected mobility corresponding to the trypsin-resistant core of CryIF (Appendix I, Figure 2, lanes 8 and 10). The apparent mobility corresponds to a molecular mass of ~66kDa. The predicted trypsin-resistant core for a microbially produced toxin is 65,912 Da (residues 28 to 612, Appendix I, Figure 1). Additional species with apparent molecular weights of approximately 100, 45, 20 and 15 kDa were apparent in both microbial lots.

Processing by the Western technique with rabbit anti-CryIF antibody revealed a doublet at a mobility approximating 65kDa in the plant-derived CryIF (Lot 1568-022B) (Appendix I, Figure 3, lane 4). Neither of these proteins was evident by Coomassie Blue staining. Examination of the amino acid sequence of CryIF, as illustrated in Appendix I, Figure 1, shows that the molecular weight difference between a protein with the natural N-terminus and one resulting from subsequent proteolytic processing at a trypsin site near the N-terminus, is 3,220 Da. This is the approximate difference observed in the estimated molecular weights of the two immunoreactive polypeptide species observed in lane 4 of Figure 3. Both preparations of bacterially-derived, trypsin-treated CryIF (lots 1599-39 and 1599-45) contain a number of species reactive with the anti-CryIF antibody (Appendix I, Figure 3 lanes 6 and 8). The most prominent of these have apparent molecular masses equivalent to 65, 25 and 20kDa. The 65kDa species is the principle Coomassie-stained species observed in lanes 8 and 10 of Figure 2 and correspond to the molecular weight predicted for trypsinolyzed CryIF (ca., 66 kDa).

CryIF content

Plant derived CryIF (Lot 1568-022B) yielded a toxin content of 0.158% by ELISA (Appendix J). The powder from non-expressing corn was below detectable limits of the ELISA (10 picograms per well), which corresponds to <0.00004% CryIF in the powder. Both microbial lots (1599-39 and 1599-45) were analyzed by SDS PAGE and yielded toxin contents of 11.4% each (Appendix K). SDS PAGE allowed specific integration around the protein band of the appropriate molecular weight for the toxin, thus avoiding any potential artifacts in quantification due to the lower molecular weight immunoreactive species present in the microbial powders.

Colorimetric Tests For Total Protein

The results of the Lowry and the Bradford were as follows (percent total protein in powder):

Sample	Lot	Lowry	Bradford
CryIF from corn	1568-022B	81.6%	84.2%
Control Corn	1568-022A	93.4%	87.4%
Microbial CryIF	1599-39	81.8%	nd
Microbial CryIF	1599-45	85.4%	nd

The Lowry was the procedure prescribed in the Study Protocol, but the Bradford is incidentally necessary for performance of the ELISA. Results were therefore compiled for comparison. For the Bradford, three determinations were averaged to produce the value from which the reported percent

protein was computed. 'nd' means not determined as these microbial samples were not analyzed by ELISA (Data in Appendices I and J).

N-terminal sequence analysis

The N-terminus of plant-derived CryIF was blocked significantly, but a five amino acid sequence corresponding to the expected N-terminus of proteolytically cleaved CryIF was obtained. This observed sequence was ^{28}STGRL, where the superscript refers to the amino acid residue in the subject protein. This is the expected N-terminus following cleavage by a trypsin-like activity (^{20}VEILNEER$\Downarrow^{28}$STGRLPLD). Two sequences and a fragment were obtained on the microbial derived CryIF; ^{28}STGRLPL, ^{43}FLLSEFV, and ^{33}PLDI. Likewise, the fragment beginning at ^{43}F follows a trypsin site (^{39}SLTR$\Downarrow^{43}$FLLSEFVP) (Appendix L and I).

Glycosylation

The immunoblot detection technique used here demonstrated that there was no apparent post-translational modification of CryIF in maize-derived extracts involving carbohydrates. Appendix I, Figure 4B, lanes 8 and 10, shows that the bacterially-derived CryIF materials are essentially bereft of reactivity. The maize-derived materials exhibit a number of discrete molecular weight species (Appendix I, Figure 4B, lanes 4 and6), however, none of these co-migrate with mobilities concordant with the immunoreactive species obtained in the western blot (Appendix I, Figure 3).

Susceptibility to digestion under simulated gastric conditions

The digestibility of bacterially derived CryIF protein toxin was determined by exposing suspensions of the protein at room temperature to a simulated gastric juice, containing either varying amounts of pepsin for a set incubation period or a set amount of pepsin incubated for varying periods of time. Ratios of CryIF:pepsin (w/w) tested ranged from 1:1 to 1:1000000. This corresponds to a molar ratio of CryIF:pepsin ranging from approximately 1:2 to 1:1,883,000. CryIF was completely proteolyzed to amino acids and small peptides within 5 min at molar ratios approximating 1:100 (CryIF:pepsin) as shown in Appendix I, Figure 6B. In a time-course experiment, a molar ratio of 1:100 (CryIF:pepsin) effected nearly complete proteolysis of the Bt toxin within one minute (Appendix I, Figure 6A).

Biological

The LC_{50} data from the top-load tests at Ricerca (Data in Appendix M) are tabulated below:

Source	LC_{50} ng CryIF per cm^2				
	European Corn Borer	Tobacco Budworm	Fall Armyworm	Corn Earworm	Black Cutworm
Plant (1568-022B)	0.58	1.74	2.21	>>15.8*	22.7*
Microbial (1599-39)	0.58	1.88	2.49	51.6	69.2

* The top dose tested failed to produce more than 50% mortality on the insect species. LC_{50}s were computed by the POLO-PC program, but represent extrapolations beyond the actual data.

The assay positive control, Dipel® 2X, yielded the following LC_{50}s (in ppm substance):

Source	LC_{50} ppm Dipel® 2X				
	European Corn Borer	Tobacco Budworm	Fall Armyworm	Corn Earworm	Black Cutworm
Dipel control	1.03	0.70	10.5	0.75	189

The behavior of the control plant material (1568-022A) in the bioassays is summarized in tabular form by presenting the average mortality observed at the indicated dose. The '0' treatment represents the negative control. The value in parentheses is the number of assays represented in the average.

	Average Percent Mortality Plant Control Powder (1568-022A) Average of All Assays (*N* in parenthesis)				
LC_{50} ng Control Plant Powder per cm^2	European Corn Borer	Tobacco Budworm	Fall Armyworm	Corn Earworm	Black Cutworm
0	5.2 (4)	15.2 (5)	8.2 (4)	4.0 (4)	8.3 (4)
500	6.3 (1)	-	-	-	-
5,000	7.4 (3)	12.4 (5)	12.1 (4)	22.2 (1)	29.2 (1)
10,000	-	-	-	14.4 (3)	14.7 (2)
20,000	-	-	-	-	44.7 (1)

* "-" dose not part of dilution test scheme for this insect

The LC_{50} of plant derived CryIF (Lot 1568-022B) in diet incorporation tests with European corn borer was found to be 0.11 ppm CryIF toxin (Appendix N). The LC_{50} of bacterial derived CryIF (Lot 1599-39) was found to be 0.14 ppm CryIF toxin. The non-expressing plant extract (Lot 1568-022A) produced percent mortality ratings of 0 to 14% at all doses but the top dose (versus 14% mortality observed in the buffer control). Mortality in the top dose was 36%.

CIRCUMSTANCES POTENTIALLY AFFECTING THE QUALITY OR INTEGRITY OF THE DATA

All deviations and amendments were disclosed above. The area of greatest potential concern would be in the use of two different quantification methods on the test materials (ELISA on the plant derived material and SDS PAGE on the microbially derived material). The main reason driving this decision was the appearance of strongly cross reacting bands in the lower molecular weight range of the microbially produced toxin, as observed in the Western blots. The appearance of such cross reaction was expected since the source of antigen from which the anti-CryIF antibodies was produced was the identical strain used to prepare the microbial lots (Pseudomonas fluorescens MR872). In addition, Luo and Adang (1994) described that 10 and 25 kDa contaminants were present in their preparations of CryIC. These fragments were determined to have arisen during trypsinolysis of the toxin and were from the N-terminus of CryIC. They adhered tightly to the truncated toxin and could not be removed by customary biochemical procedures. The nearly identical LC_{50}s of the preparations as determined on three susceptible insect species in two different dosing formats (top load and diet incorporation) strongly support the validity of the two different toxin quantification methods.

SUMMARY AND ANALYSIS

Data were obtained sufficient to compare the biological and biochemical properties of CryIF toxin from two sources. The first source was from corn genetically modified to express the toxin, and the second from a microbial expression system that produces recombinant proteins in high yield. Toxin was observed by SDS PAGE and Western to be present at mobilities corresponding to the expected molecular weight (68 kDa and 65 kDa N-terminally processed). The expected immunological reactivities were observed with the plant derived material both on solid phase (Western) and in liquid phase (ELISA). No unexpected cross reactivity was observed in control corn. There were low molecular weight immunoreactive species present in the microbial lots, thus only solid phase immunoreactivity was assessed.

The respective materials were assayed for toxin content and tested in bioassay on susceptible and non-susceptible species. The observed LC_{50} values on susceptible species were indistinguishable between toxin derived from plant and microbial sources. This strongly indicates the biological equivalence of these test materials and indicates that the analytical procedures used accurately assessed the toxin content of the preparations. The relative activity of the toxin preparations was likewise indistinguishable.

The N-terminal sequence analysis of the materials did indicate blockage in the plant material, but a sequence was able to be deduced by inspection of the chromatograms. The sequence corresponded to that of the protein after limited N-terminal processing by a protease with a trypsin-like specificity (cuts after Arg). This likely arose during the steps in preparing and concentrating the toxin from plant material. Likewise, the same N-terminus was found for microbially derived material.

There was no evidence for glycosylation of either the plant or microbially derived toxin. In addition, the toxin from microbial source was easily digested by pepsin under simulated gastric conditions. The digestion showed both time and concentration dependency.

STORAGE LOCATIONS

Raw data and Study Director files are located in the QA Archives, Mycogen Corporation, San Diego. Retain samples of the test materials are located in the QA Retain area at the same location.

REFERENCES

Luo, KE and M. J. Adang. 1994. Removal of Adsorbed Toxin Fragments that Modify *Bacillus thuringiensis* CryIC δ-Endotoxin Iodination and Binding by Sodium Dodecyl Sulfate Treatment and Renaturation. *Applied and Environmental Microbiology*. 60:2905 – 2910.

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Aug. 1994, p. 2905–2910 Vol. 60, No. 8
0099-2240/94/$04.00+0

Removal of Adsorbed Toxin Fragments That Modify *Bacillus thuringiensis* CryIC δ-Endotoxin Iodination and Binding by Sodium Dodecyl Sulfate Treatment and Renaturation

KE LUO AND MICHAEL J. ADANG*

Department of Entomology, University of Georgia, Athens, Georgia 30602

Received 21 January 1994/Accepted 16 May 1994

We report that 10- and 25-kDa toxin fragments adhere to CryIC prepared from *Bacillus thuringiensis* insecticidal crystals, block iodination, and alter membrane binding. There is no apparent affect on CryIC toxicity against *Spodoptera exigua*. Associated peptides remained bound to CryIC in the presence of 50 mM dithiothreitol or 6 M urea. A novel detergent-renaturation procedure was developed for the purification of *B. thuringiensis* CryIC toxin. Sodium dodecyl sulfate (SDS) treatment followed by gel filtration chromatography yielded a homogeneous 62-kDa CryIC toxin. After removal of SDS and renaturation, the purified CryIC toxin was fully insecticidal to *S. exigua* larvae. ^{125}I-labeled CryIC bound with high affinity to brush border membrane vesicles from *S. exigua* larvae.

Bacillus thuringiensis produces insecticidal protein crystals made of subunits called δ-endotoxins. Of the numerous δ-endotoxins, the lepidopteran-active CryIC protein is particularly important because of its toxicity to *Spodoptera* species (16, 17, 27, 35). These insects are common agricultural pests which are not susceptible to most *B. thuringiensis* δ-endotoxins (37).

The CryI δ-endotoxins are protoxins of 130 to 140 kDa which are activated by trypsin-like enzymes in the insect midgut to 60- to 65-kDa toxins (1, 11, 14). CryI toxins bind with high affinity to proteins located in the midgut brush border membrane of susceptible insects (15, 33). For example, CryIA(c) toxin binds to a 120-kDa protein in *Manduca sexta* (13, 19). This 120-kDa protein was recently identified as aminopeptidase N (18, 28). Binding proteins enhance the ability of the toxin to form ion channels in midgut cells. A mixture of *M. sexta* aminopeptidase N and a 65-kDa toxin binding protein in phospholipid vesicles catalyzed toxin-induced pore formation 1,000-fold (28).

CryIC toxin recognizes different sites than the CryIA toxins (33). A 40-kDa protein was identified as a candidate CryIC receptor in *Spodoptera litoralis* (23, 26). The unique specificity of CryIC is apparent in insects that have acquired resistance to CryIA toxins. *Plodia interpunctella* (Indian meal moth) and *Plutella xylostella* (diamondback moth) resistant to CryIA toxins still bind and are killed by CryIC toxin (12, 34). In *P. xylostella*, resistance is reversible and correlated with the return of CryIA toxin binding sites (30).

Purified δ-endotoxin is the critical reagent in many *B. thuringiensis* experiments, and yet little is known regarding the effects of contaminating proteins. Previous reports mention peptides that copurify with toxins. Pfannenstiel et al. reported that 22- and 38-kDa peptide fragments were associated with *B. thuringiensis* subsp. *israelensis* toxin and were not removed after high-performance liquid chromatography (25). Schwartz et al. (29) found minor peptides associated with CryIC toxin prepared from *B. thuringiensis* crystals but not with CryIC produced in recombinant *Escherichia coli*. Van Rie et al. (33) extracted CryIC from recombinant *E. coli* and found that a contaminating small peptide blocked ^{125}I labeling. In this study, we found that 25- and 10-kDa peptides are tightly associated with CryIC toxin isolated from *B. thuringiensis*. These peptides were preferentially labeled with ^{125}I and caused nonspecific attachment to brush border membranes.

The primary goal of this work was to develop a method of purifying CryIC toxin that yields a homogeneous 62-kDa protein. The sodium dodecyl sulfate (SDS) treatment removed toxin-associated peptides, and the renaturation treatment restored binding and insect toxicity. The toxin-associated peptides were identified as N-terminal fragments of CryIC toxin. We also show that a ^{125}I-labeled 10-kDa peptide binds nonspecifically to *S. exigua* membrane vesicles and that adhering peptides modify the interaction of CryIC toxin with *S. exigua* brush border membrane.

MATERIALS AND METHODS

Biological materials. A *B. thuringiensis* strain harboring a single δ-endotoxin gene (*cryIC*) was provided by Ecogen, Inc. (Langhorne, Pa.). This strain was constructed by transferring a *cryIC* gene (16) on a self-replicating plasmid (2) into a crystal-negative *B. thuringiensis* host. *S. exigua* was obtained from the USDA/ARS Southern Field Crop Insect Management Laboratory (Stoneville, Miss.).

Bacterial growth and toxin isolation. *B. thuringiensis* was grown at 30°C in 1 liter of L broth until sporulation and cell lysis. The crystal-spore-debris mixture was centrifuged at 7,500 × *g* for 30 min, and the pellet was washed two times with distilled H_2O. The pellet was treated with 30 ml of 50 mM KOH–1% 2-mercaptoethanol for 30 min to dissolve crystals and then centrifuged at 27,000 × *g* to remove insoluble debris. Protoxin was precipitated by lowering the pH to 5.0 with HCl. The precipitate was recovered after centrifugation at 27,000 × *g* and then dissolved in 5 ml of 50 mM 3-(cyclohexylamino)-1-propanesulfonic acid (CAPS; pH 10.0). L-1-Tosylamide-2-phenylethylchloromethylketone (TPCK)-treated trypsin (5 mg) was added, and the mixture was incubated at room temperature for 15 min. The resulting toxin was precipitated at pH 5.0 and recovered by centrifugation at 27,000 × *g* for 30 min. This toxin is called $CryIC_{crude}$ ($CryIC_c$). The $CryIC_c$ precipitate was stored at −20°C in distilled H_2O.

Toxin was also prepared from crystals purified by buoyant

* Corresponding author. Phone: (706) 542-2436. Fax: (706) 542-2279. Electronic mail address: Adang@bscr.cc.uga.edu.

TABLE 1. Insect toxicity of CryIC preparations

Toxin	Description	LC_{50}[a]
$CryIC_c$	Trypsin-activated toxin	186 (127–247)
$CryIC_p$	CryIC with 10- and 25-kDa peptides	192 (136–254)
$CryIC_s$	CryIC in 0.1% SDS	599 (426–1187)
$CryIC_r$	CryIC purified in SDS and then renatured	188 (109–261)
^{125}I-$CryIC_r$	Iodinated $CryIC_r$	200 (136–287)

[a] Expressed in nanograms per square centimeter; the 95% confidence intervals are in parentheses.

density gradient centrifugation using Renografin-76 (Squibb) (22). The spore-crystal pellet was sonicated in distilled H_2O and then washed two times in 1 M NaCl containing 0.1% Triton X-100. This suspension was layered on 64% Renografin and centrifuged at 43,000 × *g* for 1 h in a swinging-bucket rotor. The central region containing crystals was collected, washed, and further purified by using 60, 67% Renografin step gradients and the same centrifugation conditions. Crystals were washed two times in distilled H_2O, lyophilized, and stored dry at −20°C. Toxin was prepared from crystals by the procedure described above. Activation of toxin was done by using 1 mg of trypsin per 2 mg of protoxin in 1 ml of CAPS (pH 10.0).

Toxin purification methods. Toxin descriptions and designations are listed in Table 1. (i) $CryIC_{peptides}$ ($CryIC_p$) was prepared by dissolving $CryIC_c$ toxin in 50 mM CAPS (pH 10.0) and then separating it over a Sephacryl S-300 (Pharmacia) column (1.5 by 40 cm) equilibrated with the same buffer.

(ii) For anion-exchange chromatography, the $CryIC_c$ toxin was dissolved in 20 mM piperazine (pH 9.5) and applied to a Q-Sephadex (Pharmacia) column, and proteins were eluted with a 0 to 1 M NaCl gradient in 20 mM piperazine (pH 9.5).

The following treatments and purification steps were taken to remove adhering peptides from $CryIC_p$. $CryIC_p$ was precipitated at pH 5.0 and washed as described above for $CryIC_c$. (i) $CryIC_p$ was dissolved in 50 mM CAPS (pH 10.0) containing 50 mM dithiothreitol, incubated for 30 min, and then fractionated in the same buffer on a Superose 12 fast-performance liquid chromatography (FPLC; Pharmacia) column. (ii) $CryIC_p$ was dissolved in 6 M urea (pH 7.0), incubated at room temperature for 30 min, and then passed through the Superose column in the same buffer. Urea and dithiothreitol in the samples were removed by dialysis overnight against 50 mM CAPS (pH 10.0) at 4°C. (iii) $CryIC_p$ was dissolved in 50 mM CAPS (pH 10.0) containing 1% SDS, incubated at room temperature for 10 min, and then applied to the column equilibrated in 50 mM CAPS (pH 10.0) containing 0.1% SDS. Toxin eluted in 0.1% SDS is called $CryIC_{SDS}$ (abbreviated $CryIC_s$).

Purification and sequence analysis of 10- and 25-kDa peptides. During the purification of $CryIC_s$ on the Superose 12 column, 10- and 25-kDa peptides coeluted in a broad peak. The 10- or 25-kDa protein peak was pooled and then concentrated by lyophilization. The dried proteins were dissolved in distilled H_2O and rechromatographed on a Sephacryl S300 (Pharmacia) gel filtration column (1.5 by 40 cm) in 50 mM CAPS (pH 10.0) containing 0.1% SDS. The 10- and 25-kDa peptides which eluted in separate fractions were collected and then stored at −20°C for later analysis. For N-terminal amino acid sequence analysis, proteins were resolved by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (21) using 10 to 20% polyacrylamide gels and electroblotted onto a polyvinylidene difluoride membrane (Bio-Rad) in 0.5× Towbin buffer (12.5 mM Tris, 96 mM glycine [pH 8.3], 10% methanol) (31). Blotted membranes were rinsed with distilled H_2O, stained with Ponceau S (Sigma) as described in the manufacturer's (Bio-Rad) instructions, and sequenced on an Applied Biosystems model 477A protein sequenator at Tufts University (Boston, Mass.).

Removal of SDS and renaturation of toxin. Toxin eluted in 0.1% SDS was treated as described by Zahler et al. (38). Four volumes of cold acetone (−20°C) were added to the $CryIC_s$ toxin and then placed at −20°C for 20 min. The resulting precipitate was collected by centrifugation at 27,000 × *g* for 20 min at 4°C. The pellet was dissolved in 6 M guanidine hydrochloride and dialyzed against 50 mM CAPS (pH 10.0) containing 5% (vol/vol) glycerol, 0.1 M NaCl, and 1 mM EDTA at 4°C overnight to renature the toxin. SDS-treated and renatured toxin is called $CryIC_{renatured}$ ($CryIC_r$). The 10-kDa peptide was treated by the same renaturation method used in the binding studies.

Iodination. To ^{125}I-labeled toxin, two Iodobeads (Pierce) were added to 100 μl of phosphate-buffered saline (PBS; 10 mM Na_2HPO_4, 1.7 mM KH_2PO_4, 2.7 mM KCl, 136.9 mM NaCl [pH 7.4]) containing 0.5 mCi of $Na^{125}I$ and allowed to react for 5 min at room temperature. Ten micrograms of toxin (in 50 mM CAPS [pH 10.0]) was added to the reaction vial. The mixture was incubated for 15 min at room temperature. The reaction solution was applied to a Sephadex G-50 (Pharmacia) column to remove free iodine. The radioactivity was measured on a Beckman Gamma 4000 counter. Specific activities were 8.8 μCi/μg of input toxin for SDS-purified toxin and 6.6 μCi/μg for the dithiothreitol- and urea-treated toxins.

Preparation of BBMVs. Brush border membrane vesicles (BBMVs) from *S. exigua* were prepared by the method of Wolfersberger et al. (36) as modified by Ferre et al. (12). Fifth-instar larvae were chilled on ice for 15 min. The midguts were excised in ice-cold MET buffer {0.3 M mannitol, 5 mM EGTA [ethylene glycol-bis(β-aminoethyl ether)-*N,N,N',N'*-tetraacetic acid], 17 mM Tris-Cl [pH 7.5]} and then frozen on dry ice. Frozen tissue was ground in a Potter-Elvenhem homogenizer in ice-cold MET buffer containing 0.1 mM phenylmethylsulfonyl fluoride. An equal volume of cold 24 mM $MgCl_2$ was added, and the mixture was incubated on ice for 15 min. The mixture was centrifuged at 2,000 × *g* for 10 min at 4°C, and the pellet was discarded. The supernatant was clarified by centrifugation at 2,000 × *g* for 10 min and then centrifuged at 27,000 × *g* for 30 min. The pellet was resuspended in MET buffer and stored at −80°C until use.

Membrane vesicle binding assays. Immediately before use in binding assays, BBMVs were thawed on ice and then centrifuged at 15,000 × *g* for 5 min. The storage buffer was replaced with PBS containing 0.1% bovine serum albumin (BSA). Binding assays were performed as described by Garczynski et al. (13).

Gel electrophoresis. CryIC toxin preparations (either 20 μg for gels to be stained with Coomassie blue or 100,000 cpm for ^{125}I-labeled proteins) were analyzed by SDS-PAGE (20). Gels were either stained with Coomassie brilliant blue R-250 or exposed for autoradiography. Coomassie blued-stained gels were scanned with a laser densitometer (Molecular Dynamics). Autoradiography was carried out at −80°C with Kodak XAR-5 film and an intensifying screen.

Insect bioassays. Toxin preparations were diluted with PBS containing 0.1% BSA. Eight toxin concentrations were tested with 24 neonate *S. exigua* larvae per concentration. Samples (50 μl) were applied uniformly to the diet surface and then allowed to dry. For toxin in 0.1% SDS, an equivalent amount of SDS was added to buffer controls. One larva was placed onto the diet surface and reared at 26°C. Insect mortality was

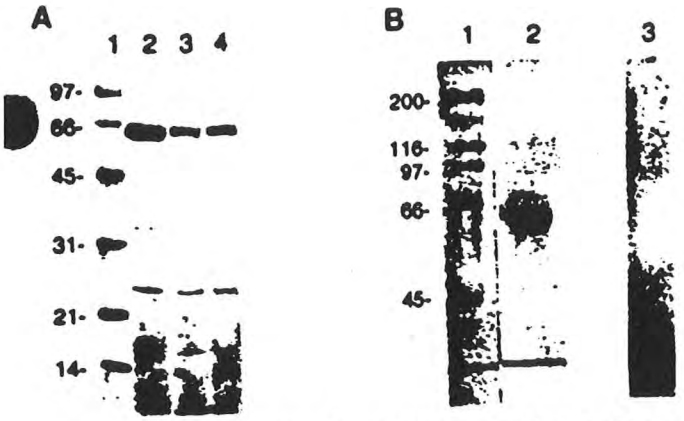


FIG. 1. SDS-PAGE of crude and partially purified CryIC. (A) CryIC prepared from a spore-crystal mixture. Lanes: 2, $CryIC_c$; 3, $CryIC_c$ passed through a Sephacryl S-300 gel filtration column in 50 mM CAPS (pH 10); 4, $CryIC_c$ applied to a Q-Sephadex ion-exchange resin and eluted with 0 to 1 M NaCl in 20 mM piperazine (pH 9.5). The 5 to 20% polyacrylamide gel was stained with Coomassie blue. (B) CryIC prepared from purified crystals. Lanes: 2, $CryIC_c$ after treatment in 50 mM dithiothreitol–50 mM CAPS (pH 10) and FPLC chromatography on Superose 12 gel filtration resin; 3, autoradiogram of the ^{125}I-labeled CryIC preparation shown in lane 2. The gel was 10% polyacrylamide. Molecular mass markers (in kilodaltons) are shown in lane 1 of both panels.

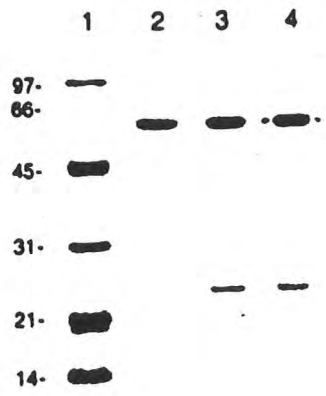


FIG. 2. Effect of various chemical treatments on the CryIC-peptide complex. Lanes: 1, molecular mass markers (in kilodaltons); 2, $CryIC_r$ (1% SDS treated and renatured); 3, $CryIC_p$ (50 mM dithiothreitol); 4, $CryIC_p$ (6 M urea). After the indicated treatment, proteins were fractionated by using FPLC and Superose 12 gel filtration. The 5 to 20% gel was stained with Coomassie blue.

scored after 7 days, and the results were analyzed by probit analysis (9).

Determination of protein concentration. The protein contents of BBMVs and CryIC toxin were measured as described by Bradford (4) with BSA as a standard. The concentration of CryIC in the presence of SDS or urea was determined with the bicinchoninic acid protein assay reagent (Pierce) with BSA as a standard. Toxin concentrations were identical when determined by both methods.

RESULTS

Partial purification of CryIC toxin. Protoxin was extracted from a spore-crystal mixture and then cleaved by trypsin to the expected 62-kDa protein and small peptides (Fig. 1A, lane 2). Crude $CryIC_c$ toxin was passed through a gel filtration or anion-exchange column. Instead of pure toxin, the 62-kDa toxin fractions had contaminating proteins of 25 and 10 kDa plus several other peptides (Fig. 1A, lanes 3 and 4). Because we could not iodinate the CryIC preparations shown in Fig. 1 (data presented below), we investigated methods to eliminate copurifying peptides. We were initially concerned that the adhering peptides were caused by our use of a spore-crystal mixture instead of purified crystals. Crystals were purified on Renografin density gradients, and CryIC toxin was prepared. Figure 1B (lane 2) shows the toxin derived from crystals after separation on a Superose 12 gel filtration column in 50 mM dithiothreitol. This column procedure was recommended previously for *E. coli*-derived CryIC as a means to eliminate peptides that block iodination (33). The small peptides migrated near the dye front on the SDS–10% polyacrylamide gel (Fig. 1B, lane 2), but their effects on iodination are clear (Fig. 1B, lane 3). When toxin was iodinated, the label was not attached to 62-kDa toxin but to the small peptide near the dye front. Use of a spore-crystal mixture was not a factor in toxin purity. These results showed that peptides adhered to CryIC toxin and blocked toxin iodination.

SDS treatment, identification of peptides, and CryIC renaturation. To identify a treatment that dissociated toxin from contaminating peptides, $CryIC_c$ was treated with 50 mM dithiothreitol, 6 M urea, or 1% SDS and then fractionated by gel filtration chromatography. Figure 2 shows a stained SDS-polyacrylamide gel of the toxin fractions. Neither dithiothreitol nor urea eliminated the 25- and 10-kDa peptides (Fig. 2, lanes 3 and 4). SDS treatment followed by gel filtration in the presence of 0.1% SDS resulted in a single protein of 62 kDa (Fig. 2, lane 2).

After separation from CryIC toxin in 0.1% SDS, the 25- and 10-kDa peptides were further purified and used to obtain a partial amino acid sequence. Both peptides were found to be N-terminal fragments of CryIC toxin. The first nine amino acid residues of the 25-kDa peptide matched residues 28 to 36 of the CryIC protoxin (10), while the 10-kDa N terminus corresponded to residues 50 to 58 of CryIC protoxin. We did not determine the first residue of trypsin-activated CryIC toxin, but the N terminus predicted by comparison with other CryI toxins is residue 29.

A drawback of SDS is its strong denaturant effect that destroys the function of many proteins. $CryIC_r$ had only 31% of the normal $CryIC_c$ toxicity to *S. exigua* larvae in bioassays (Table 1). Pfannenstiel et al. (24) reported a similar reduction in toxicity for SDS-treated *B. thuringiensis* var. *israelensis* δ-endotoxins. SDS can be removed from protein by a number of procedures. We circumvented this problem by a method in which SDS is stripped from the protein with guanidine and then the guanidine is removed by dialysis (38). This toxin preparation is designated $CryIC_{renatured}$ ($CryIC_r$).

When $CryIC_r$ was labeled with ^{125}I, a single 62-kDa peptide was visualized by autoradiography (Fig. 3). In contrast, preparations of CryIC with adhering peptides (both $CryIC_c$ and $CryIC_p$) had the ^{125}I residue attached to the 10-kDa peptide. On the basis of the relative intensities of the stained 10-, 25-, and 62-kDa bands seen in Fig. 2, a small amount of adhering peptide was sufficient to completely block iodination of CryIC toxin.

Functional assay of renatured CryIC. Renatured toxin, i.e., $CryIC_r$, was compared with native CryIC ($CryIC_c$ and $CryIC_p$)

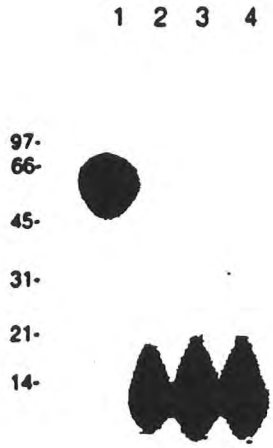


FIG. 3. Autoradiography of ^{125}I-labeled CryIC preparations. Lanes: 1, CryIC$_r$; 2, CryIC$_p$; 3; CryIC treated and purified in 6 M urea; 4, ^{125}I-labeled CryIC$_c$. Approximately 10^5 cpm was loaded per lane on an SDS–5 to 20% polyacrylamide gel. The autoradiogram was obtained after a 2-h exposure to film at −80°C. The positions of molecular mass markers (kilodaltons) are indicated on the left.

by the following tests: (i) bioassay of CryIC toxins against *S. exigua* larvae and (ii) ^{125}I-CryIC binding to *S. exigua* BBMV. The results presented below show that CryIC$_r$ was biologically active.

The results are presented in Table 1. If SDS was removed and toxin was successfully renatured, then the 50% lethal concentrations (LC_{50}) of CryIC$_c$, CryIC$_r$, and CryIC$_p$ should be equivalent. The data show that while the CryIC$_s$ LC_{50} value was threefold higher, renatured CryIC$_r$, CryIC$_c$, and CryIC$_p$ had nearly identical LC_{50} values. Also, iodination of CryIC$_r$ did not change its toxicity to *S. exigua*. These data demonstrate the biological activity of CryIC$_r$ and show that iodination did not decrease CryIC$_r$ toxicity significantly.

The next step in evaluating CryIC$_r$ was in binding assays with *S. exigua* BBMV. ^{125}I-CryIC$_r$ was added to increasing concentrations of *S. exigua* vesicles. Figure 4A shows maximal binding (23% of input ^{125}I-CryIC$_r$) at a vesicle concentration of 200 μg of protein per ml. The curve shape and extent of binding are characteristic of Cryl toxins that bind saturably and with high affinity to receptor sites. Figure 4A also shows the results of a binding experiment using CryIC$_p$ (labeled at the 10-kDa peptide) as the ligand. The rationale was that if the 10-kDa peptide remained associated with toxin, CryIC$_p$ binding would be similar to CryIC$_r$ binding. Figure 4A shows that only 5% of the input label bound to BBMV. We purified and ^{125}I labeled the 10-kDa peptide and showed that it bound to the same extent as CryIC$_p$ (Fig. 4A). This interaction between CryIC, small peptides, and membranes was interesting and explored further.

Figure 4B shows the results of an experiment in which ^{125}I-CryIC$_r$ and the amount of vesicles were kept constant and increasing amounts of unlabeled CryIC$_r$ were added. CryIC$_r$ bound with high affinity, and binding was reduced by increasing amounts of unlabeled CryIC$_r$. Data were analyzed by the computer program LIGAND. The dissociation constant (K_d) for CryIC$_r$ was 4.8 nM, and the binding site concentration (B_{max}) was 2.1 pmol/mg of vesicle protein.

We next performed a heterologous binding experiment between ^{125}I-CryIC$_r$ and unlabeled CryIC$_p$. CryIC$_p$ reduced ^{125}I-CryIC$_r$ binding, but the calculated affinity was significantly

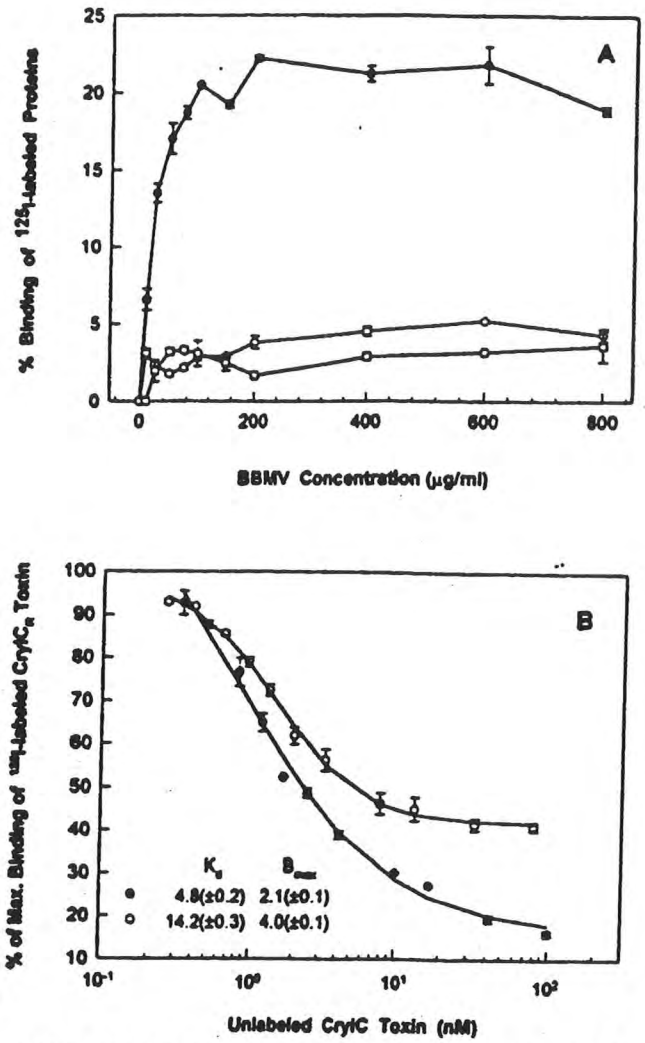


FIG. 4. CryIC binding to *S. exigua* vesicles. (A) Binding of ^{125}I-CryIC$_r$ and ^{125}I-labeled 10-kDa protein as a function of vesicle concentration. The indicated concentrations of vesicles were incubated with 0.1 nM ^{125}I-CryIC$_r$ (●), 0.1 nM CryIC$_p$ with attached ^{125}I-labeled 10-kDa peptide (○), or purified ^{125}I-labeled 10-kDa peptide (□). (B) Inhibition of ^{125}I-CryIC$_r$ binding by unlabeled CryIC preparations. Vesicles were incubated with 0.1 nM ^{125}I-CryIC$_r$ plus the indicated concentrations of unlabeled CryIC$_r$ (●) and CryIC$_p$ (○). The CryIC$_p$ preparation was 79% toxin as determined by scanning densitometry. Binding is expressed as a percentage of the maximum amount of toxin bound upon incubation with labeled toxin alone. The binding parameters (mean ± standard error) estimated by LIGAND analysis are presented in the inset. K_d is expressed in nanomolar concentrations, and B_{max} is expressed in picomoles per milligram of vesicle protein. For both panels, each point is the mean of duplicate samples. Standard error between samples is shown by bars.

lower than that with CryIC$_r$ (14.2 versus 4.8 nM). In contrast to the homologous preparation CryIC$_r$, CryIC$_p$ was not able to displace 40% of the bound toxin. The amount of CryIC$_r$ bound remained at 40% in the presence of 500 nM unlabeled CryIC$_p$. These data showed that adhering toxin fragments modified CryIC toxin interaction with vesicles. We also incubated CryIC$_p$ with associated ^{125}I-labeled 10-kDa peptide with vesicles and used unlabeled CryIC$_p$ as the competitor. As expected from the results in Fig. 4A, this complex showed a low level of

nonspecific binding, but the binding was not displaced by a 1,000-fold excess of unlabeled $CryIC_p$ toxin (data not shown).

In summary, $CryIC_r$ bound saturably and with high affinity to *S. exigua* BBMV. The K_d and B_{max} were close to those reported for high-affinity CryIC in *S. litoralis* (33). These binding data supported the bioassay results which showed $CryIC_r$ to be successfully renatured. A surprising observation was that the 10- and 25-kDa toxin peptides modified CryIC binding. Furthermore, labeled 10-kDa peptide bound nonspecifically to *S. exigua* vesicles.

DISCUSSION

This study describes a method to purify CryIC in SDS that removes adhering 10- and 25-kDa toxin fragments. These toxin fragments blocked iodination at tyrosine residues. How a 10-kDa toxin fragment with a single tyrosine residue (predicted from the *cryIC* DNA sequence) blocks labeling at all of the 12 to 13 tyrosine residues in the CryIC toxin is not clear. We found that increasing the $Na^{125}I$ concentration in the labeling reaction or using the chloramine-T method as described in reference 13 did not yield ^{125}I-CryIC (data not shown).

Contaminating peptides in CryIC preparations have been reported previously (29, 33). Van Rie et al. (33) concluded that peptides associated with CryIC (isolated from *E. coli*) impeded ^{125}I labeling. Those researchers (33) found that exposure to 50 mM dithiothreitol followed by gel filtration eliminated this difficulty. Neither dithiothreitol nor 6 M urea had any effect on our *B. thuringiensis*-derived toxin. The discrepancy between our results and those of Van Rie et al. (33) may be caused by several factors. One factor may be the different host organisms. For example, to isolate CryIC and investigate toxin-induced ion channels, Schwartz et al. (29) prepared CryIC from both *B. thuringiensis* crystals and from *E. coli* inclusions. Only CryIC prepared from *B. thuringiensis* crystals had contaminating peptides. Fermentation conditions may also be a factor in δ-endotoxin production. Culture media do not only have the obvious impact on yield but also seem to cause subtle changes in toxin chemistry. Recently, Bhattacharya et al. (3) showed that sugars are nonenzymatically attached to CryI proteins released into culture media. Those researchers (3) convincingly presented arguments as to how glycosylation can account for toxin aggregation and altered toxicity to insects.

SDS-treated and renatured $CryIC_r$ was active in insect and binding assays. The capacity of CryIC to refold to a biologically active state was remarkable but predictable. Choma and Kaplan (6) denatured CryIA(c) with guanidine hydrochloride and then restored toxicity to cultured cells by removal of guanidine hydrochloride. In our experiments, $CryIC_r$ bound with high affinity (K_d = 4.8 nM) to *S. exigua* vesicles. Van Rie et al. (33) measured ^{125}I-CryIC binding to *S. litoralis* BBMV. Analysis by the LIGAND program revealed high-affinity (K_d = 0.18 nM) and low-affinity (K_d = 13.9 nM) binding sites (33). Our data do not eliminate the possibility of multiple CryIC sites in *S. exigua*. The slight upward deflection of the $CryIC_r$ binding curve (Fig. 4B) at 10 nM could be interpreted as being due to a population of low-affinity binding sites. Assays using different CryI toxins are necessary to aid in choosing between one- and two-site models of CryIC binding to *S. exigua*.

The distinct binding characteristics obtained with unlabeled $CryIC_r$ and $CryIC_p$ were dramatic. Substituting $CryIC_p$ with adhering peptides for $CryIC_r$ decreased the affinity (K_d) from 4.8 to 14.2 nM. Also, a 1,000-fold excess of $CryIC_p$ did not displace 43% of the ^{125}I-$CryIC_r$. We obtained identical results with a second independent set of vesicles and toxin preparations (data not shown). Three plausible explanations for these results follow. (i) $CryIC_r$ binding fits a two-site model in which $CryIC_p$ competes for the low- but not the high-affinity site. This pattern is observed with CryI toxins that have some, but not all, midgut sites in common (32, 33). (ii) Adhering 10- and 25-kDa toxin fragments change the affinity of CryIC for the receptor. (iii) The nonspecific interaction of 10- and 25-kDa sub-toxin peptides with the membrane alters $CryIC_r$ binding.

Amino acid sequencing revealed that the 10- and 25-kDa peptides originated from the CryIC molecule. The peptides correspond in molecular size to known proteolytic fragments of CryIA toxins. CryIA toxins are cleaved under conditions of excess trypsin, or while partially denatured, to peptides of 30 to 34 kDa (5, 7, 8). Intensive proteolysis of *B. thuringiensis* subsp. *alesti* CryIA δ-endotoxin leads to the release of 10-kDa peptides from the toxin molecule (5). These peptides were hydrophobic and formed stable aggregates (5). We observed CryIC toxin to be particularly susceptible to overdigestion by trypsin (unpublished data). Our observation that the 10-kDa toxin fragment binds nonspecifically to BBMV is not surprising since it originates from the hydrophobic domain predicted to enter the membrane upon insertion of toxin (21).

The purification-renaturation method presented here will be useful in the preparation of CryIC toxin and may have further applications for other classes of δ-endotoxins. The method further demonstrates the structural stability of CryI toxins. While this study showed that adhering sub-toxin peptides modify CryIC labeling and membrane binding, adhering peptides probably affect other in vitro and possibly in vivo toxicity of CryIC δ-endotoxin.

ACKNOWLEDGMENTS

This work was supported through grant 9302031 from the USDA and by Ecogen, Inc.

We thank Yang-Jiang Lu and Steve Garczynski for critically reading the manuscript.

REFERENCES

1. **Adang, M. J.** 1991. *Bacillus thuringiensis* insecticidal crystal proteins: gene structure, action and utilization, p. 3–24. *In* K. Maramorosch (ed.), Biotechnology for biological control of pests and vectors. CRC Press, Inc., Boca Raton, Fla.
2. **Baum, J. A., D. M. Coyle, M. P. Gilbert, C. S. Jany, and C. Gawron-Burke.** 1990. Novel cloning vectors for *Bacillus thuringiensis*. Appl. Environ. Microbiol. 56:3420–3428.
3. **Bhattacharya, M., B. A. Plantz, J. D. Swanson-Kobler, and K. W. Nickerson.** 1993. Nonenzymatic glycosylation of lepidopteran-active *Bacillus thuringiensis* protein crystals. Appl. Environ. Microbiol. 59:2666–2672.
4. **Bradford, M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248–254.
5. **Chestukhina, G. G., S. A. Tyuria, L. I. Kostina, A. L. Osterman, I. A. Zalunin, O. A. Khodova, and V. M. Stepanov.** 1990. Subdomain organization of *Bacillus thuringiensis* entomocidal proteins' N-terminal domains. J. Protein Chem. 9:501–507.
6. **Choma, C. T., and H. Kaplan.** 1990. Folding and unfolding of the protoxin from *Bacillus thuringiensis*: evidence that the toxic moiety is present in an active conformation. Biochemistry 29:10971–10977.
7. **Choma, C. T., W. K. Surewicz, P. K. Carey, M. Pozcgay, T. Raynor, and H. Kaplan.** 1990. Unusual proteolysis of the protoxin and toxin from *Bacillus thuringiensis*: structural implication. Eur. J. Biochem. 189:523–527.
8. **Convents, D., C. Houssier, I. Lasters, and M. Lauwereys.** 1990. The *Bacillus thuringiensis* δ-endotoxin: evidence for a two domain structure of the minimal toxic fragment. J. Biol. Chem. 265:1369–1375.
9. **Daum, R.** 1970. Revision of two computer programs for probit analysis. Bull. Entomol. Soc. Am. 16:10–15.

10. **English, L., T. L. Readdy, and A. E. Bastian.** 1991. Delta endotoxin-induced leakage of $^{86}Rb^{+}$-K^{+} and H_2O from phospholid vesicles is catalyzed by reconstituted midgut membrane. Insect Biochem. 21:177–184.
11. **English, L., and S. L. Slatin.** 1992. Mode of action of delta-endotoxins from *Bacillus thuringiensis*: a comparison with other bacterial toxins. Insect Biochem. Mol. Biol. 22:1–7.
12. **Ferre, J., M. D. Real, J. Van Rie, S. Jansens, and M. Peferoen.** 1991. Resistance to the *Bacillus thuringiensis* bioinsecticide in a field population of *Plutella xylostella* is due to a change in a midgut membrane receptor. Proc. Natl. Acad. Sci. USA 88:5119–5123.
13. **Garczynski, S. F., J. W. Crim, and M. J. Adang.** 1991. Identification of putative brush border membrane-binding molecules specific to *Bacillus thuringiensis* δ-endotoxin by protein blot analysis. Appl. Environ. Microbiol. 57:2816–2820.
14. **Gill, S., E. A. Cowles, and P. V. Pietrantonio.** 1992. The mode of action of *Bacillus thuringiensis* endotoxins. Annu. Rev. Entomol. 37:615–636.
15. **Hofmann, C., H. Vanderbruggen, H. Hofte, J. Van Rie, S. Jansens, and H. Van Mellaert.** 1988. Specificity of *Bacillus thuringiensis* delta-endotoxins is correlated with the presence of high-affinity binding sites in the brush border membrane of target insect midguts. Proc. Natl. Acad. Sci. USA 85:7844–7848.
16. **Honee, G., T. van der Salm, and B. Visser.** 1988. Nucleotide sequence of crystal protein gene isolated from *B. thuringiensis* subspecies *entomocidus* 60.5 coding for a toxin highly active against *Spodoptera* species. Nucleic Acids Res. 16:6240.
17. **Kalman, S., K. L. Kiehne, J. L. Libs, and T. Yamamoto.** 1993. Cloning of a novel *cryIC*-type gene from a strain of *Bacillus thuringiensis* subsp. *galleriae*. Appl. Environ. Microbiol. 59:1131–1137.
18. **Knight, P. J. K., N. Crickmore, and D. J. Ellar.** 1994. The receptor for *Bacillus thuringiensis* CryIA(c) δ-endotoxin in the brush border membrane is aminopeptidase N. Mol. Microbiol. 11:429–436.
19. **Knowles, B. H., P. J. Knight, and D. J. Ellar.** 1991. N-acetyl galactosamine is part of the receptor in insect gut epithelia that recognizes an insecticidal protein from *Bacillus thuringiensis*. Proc. R. Soc. London Ser. B 3:31–35.
20. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680–685.
21. **Li, J., J. Carroll, and D. J. Ellar.** 1991. Crystal structure of insecticidal δ-endotoxin from *Bacillus thuringiensis* at 2.5Å resolution. Nature (London) 353:815–821.
22. **Mahillon, J., and J. Delcour.** 1984. A convenient procedure for the preparation of highly purified crystals of *Bacillus thuringiensis*. J. Microbiol. Methods 3:69–76.
23. **Oddou, P., H. Hartmann, and M. Geiser.** 1993. Immunologically unrelated *Heliothis* sp. and *Spodoptera* sp. midgut membrane-proteins bind *Bacillus thuringiensis* CryIA(b) δ-endotoxin. Eur. J. Biochem. 212:145–150.
24. **Pfannenstiel, M. A., G. A. Couche, G. Muthukumar, and K. W. Nickerson.** 1985. Stability of the larvicidal activity of *Bacillus thuringiensis* subsp. *israelensis*: amino acid modification and denaturants. Appl. Environ. Microbiol. 50:1196–1199.
25. **Pfannenstiel, M. A., W. C. Cray, G. Couche, and K. W. Nickerson.** 1990. Toxicity of protease-resistant domains from the delta-endotoxin of *Bacillus thuringiensis* subsp. *israelensis* in *Culex quinquefasciatus* and *Aedes aegypti* bioassays. Appl. Environ. Microbiol. 56:162–166.
26. **Sanchis, V., and D. J. Ellar.** 1993. Identification and partial purification of a *Bacillus thuringiensis* CryIC δ-endotoxin binding protein from *Spodoptera litoralis* gut membranes. FEBS Lett. 3:264–268.
27. **Sanchis, V., D. Lereclus, G. Menou, J. Chaufaux, S. Guo, and M.-M. Lecadet.** 1989. Nucleotide sequence and analysis of the N-terminal coding region of the *Spodoptera*-active δ-endotoxin gene of *Bacillus thuringiensis aizawai* 7.29. Mol. Microbiol. 3:229–238.
28. **Sangadala, S., F. Walters, L. H. English, and M. J. Adang.** 1994. A mixture of *Manduca sexta* aminopeptidase and alkaline phosphatase enhances *Bacillus thuringiensis* insecticidal CryIA(c) toxin binding and $^{86}Rb^{+}$-K^{+} leakage *in vitro*. J. Biol. Chem. 269:10088–10092.
29. **Schwartz, J.-L., L. Garneau, D. Savaria, L. Masson, R. Brousseau, and E. Rousseau.** 1993. Lepidopteran-specific crystal toxins from *Bacillus thuringiensis* form cation- and anion-selective channels in planar lipid bilayers. J. Membr. Biol. 132:53–62.
30. **Tabashnik, B. E., N. Finson, F. R. Groeters, W. J. Moar, M. W. Johnson, K. Luo, and M. J. Adang.** 1994. Reversal of resistance to *Bacillus thuringiensis* in *Plutella xylostella*. Proc. Natl. Acad. Sci. USA 91:4120–4124.
31. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA 76:4350–4354.
32. **Van Rie, J., S. Jansens, H. Hofte, D. Degheele, and H. Van Mellaert.** 1989. Specificity of *Bacillus thuringiensis* delta-endotoxins. Eur. J. Biochem. 186:239–247.
33. **Van Rie, J., S. Jansens, H. Höfte, D. Degheele, and H. Van Mellaert.** 1990. Receptors on the brush border membrane of the insect midgut as determinants of the specificity of *Bacillus thuringiensis* delta-endotoxins. Appl. Environ. Microbiol. 56:1378–1385.
34. **Van Rie, J., W. H. McGaughey, D. E. Johnson, B. D. Barnett, and H. Van Mellaert.** 1990. Mechanism of insect resistance to the microbial insecticide *Bacillus thuringiensis*. Science 247:72–74.
35. **Visser, B., T. van der Salm, W. van den Brink, and G. Folkers.** 1988. Genes from *Bacillus thuringiensis* entomocidus 60.5 coding for insect-specific proteins. Mol. Gen. Genet. 212:219–224.
36. **Wolfersberger, M. G., P. Luthy, A. Maurer, P. Parenti, V. F. Sacchi, B. Giordana, and G. M. Hanozet.** 1987. Preparation and partial characterization of amino acid transporting brush border membrane vesicles from the larval midgut of the cabbage butterfly (*Pieris brassicae*). Comp. Biochem. Physiol. A 86:301–308.
37. **Yamamoto, T., and G. K. Powell.** 1993. *Bacillus thuringiensis* crystal proteins: recent advances in understanding its insecticidal activity. p. 3–42. *In* L. Kim (ed.), Advanced engineered pesticides. Marcel Dekker, Inc., New York.
38. **Zahler, A. M., K. M. Neugebauer, W. S. Lane, and M. B. Roth.** 1993. Distinct functions of SR proteins in alternative pre-mRNA splicing. Science 260:219–222.

Appendix A: Production of Plant derived CryIF powder and Plant Control Powder

Materials and Methods

Preparation of Cry1Fa2-enriched maize-derived protein

Cry1Fa2 toxin was partially purified from leaves collected at the V-10 stage of growth from inbred line 5XH751 x TC1360 [BC2] Source #E9710039. Leaf material was also collected at the V-10 stage of growth from the isogenic inbred line, 5XH751 to serve as control material and was handled in the same manner. Leaves were frozen at -20°C until processed. Unless noted otherwise, all subsequent processing steps were performed at 4°C. Frozen tissue, corresponding to 3.6kg (Bt-maize leaf) and 2.3kg (control-maize leaf) was homogenized with 3 volumes of Tissue Extraction Buffer in a stainless steel commercial blender (Waring Co., New Hartford, CN) fitted with a double blade.

Extracts were clarified first by filtering through cheesecloth then by the slow addition of 10% (w/v) polyethylenimine to a final amount corresponding to 0.15% (w/v) while stirring. After continuing stirring an additional 15min, the material was further clarified by centrifugation (10k x g x 30min) in a Sorvall RC-5B in a GS-3 rotor (Dupont, Wilmington, DE). The supernatant was recovered by filtration through Miracloth (Calbiochem Co., San Diego, CA). The supernatant was brought to 40% saturation with ammonium sulfate by slowly adding the crystalline reagent while stirring. Following addition of the ammonium sulfate, the material was allowed to stir an additional 30min. The precipitate was collected by centrifugation as described above (10k x g x 30min) and resuspended in one-tenth to one-fortieth the original extract volume in Tissue Dialysis Buffer. The material was then dialysed overnight against 500mM ammonium bicarbonate to remove residual ammonium sulfate. After dialysis the material was lyophilized (Freezemobile 6 Lyophilizer, Virtis Co., Gardiner, NY) and stored under vacuum at room temperature.

The Cry1Fa2 maize material (5.37g) was designated as Batch #1568-022B and the control maize material (2.40g) was designated as Batch #1568-022A .

Buffers and Solutions

Unless indicated otherwise, all reagents listed below were obtained from generally available commercial sources.

Tissue Extraction Buffer
50mM 3-[cyclohexylamino]-1-propanesulfonic acid (CAPS), pH 10
0.1M NaCl
2mM disodium ethylenediamine tetracetic acid
2mM dithiothreitol
0.5mM Pefabloc® SC (4-[2-aminoethyl] benzenesulfonylfluoride, HCl) (Boerhinger Mannheim, Indianapolis, IN)

Tissue Dialysis Buffer
50mM $NaHCO_3$, pH 8.0
150mM NaCl

Appendix B: Certificate of Analysis for Plant CryIF lot 1568-022B

Mycogen
5501 Oberlin Drive
San Diego, California 92121
619 453 8030

AMENDED Certificate of Analysis

This analysis certificate supersedes the analysis reported 9-11-98, which is reproduced below for clarity

Date:	September 24, 1998
Test Substance:	Powder extracted from corn tissue, cryIF enriched
Lot Number:	1568-022B
Analysis Date:	August 27, 1998 - September 1, 1998, three determinations
Analysts:	Pioneer Hi-Bred
Test Method:	ELISA using truncated cryIF as standard
Lot Analysis Result:	**0.158% cryIF**
Reason for Change:	**Double accounting for a conversion factor resulting in 10-fold error in result, Per phone conversation on 9/17/98**

[redacted]

9/24/98
Date

Manager, Biochemistry

Certificate of Analysis

Date:	September 11, 1998
Test Substance:	Powder extracted from corn tissue, cryIF enriched
Lot Number:	1568-022B
Analysis Date:	August 27, 1998 - September 1, 1998, three determinations
Analysts:	Pioneer Hi-Bred
Test Method:	ELISA using truncated cryIF as standard
Lot Analysis Result:	1.58% cryIF

[redacted] Certificate Statement

Date

Manager, Biochemistry

THIS IS AN EXACT COPY OF
THE ORIGINAL DOCUMENT
[redacted] DATE 9/24/98

Appendix C: Production of Microbial CryIF toxin

Preparation of Bacterial Cry1F Toxin

Preparation of Bacterial Cell Paste

Pseudomonas fluorescens strain MR872 was grown in 10L and 1500L fermentors (BioLafitte) according to Mycogen Fermentation Department Protocol (WNB-01; BH-22). Post-fermentation, the cell culture broth was centrifuged for 20 mins at 4°C in a Sorvall RC2B centrifuge equipped with a Composite rotor (Model KA-12.500) at 8,500 rpm (10,000xg) in polypropylene screw cap bottles (500ml). The supernatant above the cell pellet was poured off and discarded. The bottles were refilled and centrifuged up to three times to accumulate large cell pellets. Production dates spanned 3/14/95 to 6/30/98. The bottles containing cell pellets were stored at –70°C.

Preparation of Washed Protein Inclusions

Bacterial protein inclusions were prepared using lysozyme and DNAase to digest the bacterial cell walls and subsequently released nucleic acids (Notebook# - Page#: 1599-003). One bottle of frozen cell pellet was thawed at 37°C (Batch 01). The pellet (249.4 gm wet weight) was washed by suspending to ~400 ml in distilled water using a Polytron homogenizer at low speed. The suspended pellet was centrifuged as above, and the supernatant discarded. The pellet was suspended in wash buffer (1599-002) as above, centrifuged and the supernatant discarded. The washed cell pellet was suspended in 1250 ml of wash buffer using the Polytron and 750 mg. of lysozyme (Sigma L-6876) disolved in 30 ml of wash buffer added. The cell suspension was warmed for 4 mins in a 37°C water bath and then transferred to 4°C for 1 hour. The digested cell suspension was made 60 mM in $MgCl_2$ by the addition of 36.5 ml of 2.5 M $MgCl_2$. Deoxyribonuclease I (375 mg.; Sigma DN-25) was added in 10 ml of wash buffer and the cell suspension warmed at 37°C until the viscosity of the suspension was reduced. The suspension was then transferred to 4°C overnight. The digested cell suspension was transferred to 500 ml screw cap centrifuge bottles and centrifuged as above. The inclusion pellet was suspended in wash buffer and centrifuged. The supernatant was discarded and the washed procedure was repeated five times. The washed pellet (56 gm wet weight) was suspended in distilled water and shell frozen in a dry ice/ethanol bath. The frozen inclusion pellet was stored at –70°C.

Digestion of Washed Inclusion Protein

The washed inclusion protein was digested with bovine trypsin (1599-024). The frozen inclusion protein (56 gm wet weight) was thawed at 30°C and diluted to 560 ml in 100 mM Tris-Cl pH 9.0. The diluted protein was incubated at 30°C for 30 minutes. Trypsin (56 mg, Sigma T-8642) was added and the mixture was incubated for 3 hours at 30°C with occasional mixing. At the conclusion of the incubation period, the digestion mixture was stored on ice.

Diafiltration of the Digested Inclusion Protein
The trypsin-treated inclusion protein was diafiltered using an Amicon Hollow Fiber Concentrator equipped with three HIP30 (Amicon, MWCO 30kD) filters mounted in series. The digested protein was combined with two other batches of digested protein (Batch 03 and 14) to a final volume of 1650 ml. Diafiltration was performed using 12 liters of 1 mM Tris-Cl pH 9.0. A final conductivity of 0.022 mMHO was measured in the diafiltration effluent. The diafiltered protein solution was stored at 4°C.

Lyophilization of the Diafiltered Protein Solution
The diafiltered protein solution was aliquoted (500 ml) to 1 L lyophilization flasks, and shell frozen in a dry ice/ethanol bath. The frozen protein solution was lyophilized until the temperature of the flask was approximately equal to room temperature. When all of the batches had been lyophilized, the dry protein powder was transferred to a 18L glass carboy and mixed by rotation and vigorous shaking of the carboy. The powder was then aliquoted to clean, amber 1L screw cap jars and stored at 4°C.

Batch Handling
The above is a description of the handling of Batch 01. A total of 29 batches were processed from 5/26/98 through 8/18/98. Fermentation was performed in three batches (Batch #970819R, 950314G, 980630C) by Lawrence Chew, Kristin Fencil, and Brad Heath. Preparation of protein was performed by David Treiber and Mark Knuth. Quantitation was performed by Josh Russell.

Lot Creation
Two lots were created. The first lot is identified as 1599-39 and the second lot is identified as 1599-45.

[redacted] 10/14/98

Appendix D: Certificate of Analysis of Microbial Lot 1599-39

Mycogen
5501 Oberlin Drive
San Diego, California 92121
619 453 8030

Certificate of Analysis

Date:	August 30, 1998
Test Substance:	Bacterial Cell Protein, CryIF (truncated)
Lot Number:	1599-39
Analysis Date:	August 28, 1998
Analysist:	[redacted]
Test Method:	SDS-PAGE with coomassie staining using BSA as standard
Lot Analysis Result:	11.4% cryIF

[redacted] Ph.D.
Manager, Biochemistry

8/30/98
Date

Appendix E: Certificate of Analysis of Control Plant Extract Lot 1568-022A

Mycogen
5501 Oberlin Drive
San Diego, California 92121
619 453 8030

AMENDED Certificate of Analysis

This analysis certificate supersedes the analysis reported 9-11-98, which is reproduced below for clarity

Date:	September 24, 1998
Test Substance:	Powder extracted from corn tissue, control
Lot Number:	1568-022A
Analysis Date:	August 27, 1998 - September 1, 1998, three determinations
Analysts:	Pioneer Hi-Bred
Test Method:	ELISA using truncated cryIF as standard
Lot Analysis Result:	**below limit of detection, less than 0.00004% cryIF**
Reason for Change:	**Double accounting for a conversion factor resulting in 10-fold error in result, Per phone conversation on 9/17/98**

[redacted]

9/24/98
Date

Manager, Biochemistry

Certificate of Analysis

Date:	September 11, 1998
Test Substance:	Powder extracted from corn tissue, control
Lot Number:	1568-022A
Analysis Date:	August 27, 1998 - September 1, 1998, three determinations
Analysts:	Pioneer Hi-Bred
Test Method:	ELISA using truncated cryIF as standard
Lot Analysis Result:	below limit of detection, less than 0.0004% cryIF

Certificate Statement

[redacted]
Date

Manager, Biochemistry

THIS IS AN EXACT COPY OF
THE ORIGINAL DOCUMENT
BY [redacted] DATE 9/24/98

Appendix F. Documents relevant to Deviations to Protocol MYCO98-001.

MYCOGEN CORPORATION
CHANGE IN PROTOCOL FORM # myco98001-1

Check One

- [ ] Amendment
- [x] Deviation

Protocol Number: MYCO 98-001 (PHASE 9.6)
Test Material: EVERGREEN CORN, BT
Study Director: [redacted]

Nature of Amendment/Deviation*:

BIO RAD'S PVDF TRANSFER MEMBRANE WAS USED INSTEAD OF IMMOBILON-PSQ TRANSFER MEMBRANE AND A 0.1% COOMASSIE BLUE STAIN WAS USED INSTEAD OF A 1% PONCEAU S IN 1% ACETIC ACID

Reason for Amendment/Deviation*:

THE IMMOBILON-PSQ TRANSFER MEMBRANES WERE OUT OF DATE COOMASSIE BLUE IS CONSIDERED A MORE SENSITIVE STAIN

Impact on Study*:

NO IMPACT ON STUDY

*Attach Additional Sheets If Necessary

Approved By: [redacted] Date: 9/21/98
Study Director

Approved By: [redacted] Date: 9.18.98
Quality Assurance

Acknowledged By: [redacted] Date: 9.18.98
Study Sponsor

H:\Hays\MYFILES\QA\ProtChngFrm.doc

THIS IS AN EXACT COPY OF THE ORIGINAL DOCUMENT
BY [redacted] DATE 9/24/98

[redacted]

From: [redacted]
Sent: Wednesday, September 09, 1998 5:17 PM
To: [redacted]
Subject: Cold Room

[redacted]

As a result of the electrical work which was carried out this last weekend as part of the construction project, some of the the cold room electricals were not operating. It is not clear how long the cold room was inoperative, but the temperature was discovered to be higher than ambient (around 23o C.) at about 8:00 AM on Tuesday 9/8/98. Physical plant corrected the situation and the room began cooling again by about 9:00 AM. By noon the room air had returned to about 8-10oC. The Cry1F protein samples were stored in the cold room during this entire episode. Because the protein exists as a lyophilized powder in tightly closed glass jars, I think the protein sample should be intact.

[redacted]

[redacted]

From: [redacted]
Sent: Wednesday, September 30, 1998 11:32 AM
To: [redacted]
Cc: [redacted]
Subject: RE: Bioequivalency Data

Reading from the protocol the following points seem relevant,

10.2.4 ...at least 3 tests

10.2.5 ...48±5 wells with 1 larva per well, = minimum of 43

10.3.2 ... eight doses plus untreated, at least 2 > 50%, at least 2 <50% an at least 5 0-100% mortality, and '[d]ue to the experimental nature of this study, it may be necessary to override these criteria.'

11.2.1 percent mortality corrected by Abbott equation

11.2.2 probits, then lc50, lc90 and slope

11.2.3 control not to exceed 15%, repeat unless authorize by SD

You're fulfilling 10.2.4, so it is not at issue. To me, the importance of 10.2.5 will only come out after seeing probits. If POLO can get reasonable estimates with the number of larvae per treatment then we need to reconsider, however, if it can, then a deviation on 10.2.5 seems reasonable.

The same logic holds on 10.3.2, where the concept of 8 doses was there mainly to help you get the 5. If the data meet the criteria in 10.3.2, especially the number in the under and over 50% criteria, then I feel that a deviation is not even necessary.

So, why don't you run a few data sets through POLO. You can send me the unaudited files if that is convenient and I can look too (I have polo)

Does this help?

[redacted]

-----Original Message-----
From: [redacted]
Sent: [redacted] 10:04 AM
To: [redacted]
Cc: [redacted]
Subject: Bioequivalency Data

[redacted]

I have a question about the bioequivalency data. We are at a point where we might reasonably stop testing certain insect/toxin combinations. However, some of the data might not fulfill the testing criteria in the protocol. For example, right now we have three runs of Dipel, bacterial Cry1F and plant Cry1F against Fall armyworm in the books. None of the treatments have the eight rates and at least 43 larvae per dose per run but the data look good.

Do you want to see the (unaudited) data and decide for yourself? Would you prefer that I do some probit analysis and see how good the fits are? Or do you have an alternative suggestion?

[redacted]

8-25-98

[redacted]

Please modify the initial sample preparation portion of your elisa as follows to best handle the cryIF plant enriched and control samples.

Concepts behind change request:

- these samples are being dissolved, not extracted
- results ultimately need to be presented in terms of weight of original powder, not in terms of extractable protein

These concepts necessitate the following changes.

1. Sample buffer pH needs to be at 9-10. The exact pH is not so critical, just that it is within the range. This is necessary to dissolve the toxin contained in the samples. The samples are poorly soluble at ~ neutral pH, as is the sample extraction buffer the procedure calls for. We use carbonate (50mM carbonate/bicarbonate, pH 9.6 buffer) or tris (1 mM Tris-Cl pH 9.0) to attain this. The exact nature is not so critical so long as you document it. Upon gentle mixing and within several minutes the samples should dissolve into solution. Subsequent dilution in your customary buffer is then fine.
2. Since these materials are fluffy, it is likely that the normal dilution practice of adding liquid to the powder might not yield a quantitative estimation of the solution due to volumetric effects. It would be better to dissolve the powder in a way that the final volume of the solution (powder + diluent) is known. We do this either in volumetric flasks, or in graduated cylinders, or in careful liquid handling where the material is drawn into a glass pipette and the volume determined.
3. You can still use the Bradford on material to set the concentrations in the ELISA, but we need to be able to get back to known weight of material in known volume of diluent, so steps 1 and 2 are necessary to do that. The last liquid samples were sent in at 5 mg powder/ml and everything seemed to work well in the elisa assay.

I hope this isn't too inconvenient. Call if you have any questions.

[redacted]

8/25/98

1998-10-07 P01:02

[redacted]

EXACT COPY

[redacted]

From: [redacted]
Sent: Tuesday, September 01, 1998 3:35 PM
To: [redacted]
Subject: check -in

[redacted]

[redacted]s back from Japan and the three of us can probably get together soon to discuss thoughts on the bacterial F elisa. Will the end of this week work for you? Suggest a time if you want to do it by phone, or would rather to it by email.

Will you be able to provide the two vials of anti-F for the column coupling? We will need the antibody in to set the schedule with our external sequencing lab. If there are any issues please let me know so we can figure out how to work through them.

Lastly, I assume the materials all arrived safely? Tell Jacque that the level of detail she sent with the last practice batches was sufficient. Whenever the corn and control plants get rerun then I'll need a copy of the data to make a CoA for Ricerca on those materials.

Hope things are going well for you, and thanks for all your help

[redacted]

Study Number: 98-08-RA-GLP-009 Pioneer Confidential
8.0 TEST METHODS – BIOLOGICAL

Experimental Design: Bioassays will be conducted using the proteins described in this against European corn borer larvae (ECB) to determine the LC_{50} rate of the protein. Neonate larvae are placed on insect diet incorporated with a series of rates of the protein. The larvae are allowed to grow and develop on the diet. After seven days, percent mortality is determined for each of protein and the living larvae are counted. The LC_{50} is then calculated from this data.

8.1 Test System Description

Species, strain, substrain: Ostrinia nubilalis (European corn borer)

Source of test systems: European corn borer (ECB) neonate larvae will be obtained from eggs produced by the Pioneer ECB production laboratory.

Test stage: All insects will be neonate larvae.

Number of replicates: 14 wells per concentration

Procedure for test system identification: Micro plates will be labeled with the species, test material, dose level, date of initial exposure and study number. Justification for Test System Selection: The insect species selected has demonstrated the required sensitivity to the test substances.

Rearing Conditions: ECB eggs will be placed in an incubator maintained at 27+ 1 C for three days.

8.2 Dosing Description

8.2.1 Route of Administration and Justification: The proteins to be tested are mixed with the PBS buffer to obtain a stock solution of known concentration. A series of dilutions are then prepared from the stock solution. The stock solution volume and the PBS buffer volume are dispensed into 15ml tubes. The incorporated diet is transferred to the microplates. Two hundred microliters (200ul) of incorporated diet is dispensed into each well. The bioassay procedure is standard for evaluating the toxicity of stomach active insecticides to lepidopteran larvae.

8.2.2 Dose Levels: Five doses (ug Cry1F/ml diet) plus an untreated control at
l target rates of 0, 0.02, 0.04, 0.08, 0.16, 0.32.

Duration of exposure: 7 days.

8.2.3 Test Material Formulation: Dilutions of test materials will be made using PBS and Southland Multiple Species Diet. Two hundred (200) microliter aliquots of incorporated diet will be pipetted into each well.

8.2.4 Description of Diet and Water: Southland Multiple Species Diet will be prepared according to Pioneer SOP PHICP-SP-0007/01. The ingredients are not know to contain contaminants which interfere with the conduct of the study. Deionized water will be used throughout the project.

8.2.5 Exposure conditions: After the insect larvae have been added, each tray will be covered with CDI perforated cover cut into 3 pieces and maintained at 26+ 1 C and approximately 75+ 5% relative humidity.

8.2.6 Method Used to Control Bias: Neonate larvae will be selected without regard to sex for testing based on activity and size. Larvae that are inactive or obviously non-neonate will not be used. The bioassay reference standard will be use to validate that the assay is working.

8.3 Test System Observations:
Mortality will be assessed following 7 days incubation. Surviving larvae will be counted and percent mortality/protein rate will be determined.

8.4 Test System Disposition:

No specimens will be collected. At the end of the observation period, the test insects will be autoclaved or frozen and disposed of as laboratory waste according to Federal, State and local regulations.

9.0 PROPOSED STATISTICAL TESTS Biochemical Assays: Means and standard deviations will be generated for each sample. CVs will be determined on each sample and correlation coefficients determined for each standard curve. Bioassays: The LC_{50} is determined by plotting the protein rate (x axis) vs. % mortality (y axis) using linear regression. Using the slope formula and JMP software (SAS Institute), the LC_{50} is calculated as $y = m(x) + b$ where $y= 0.50$ for 50% mortality, m = slope of the best fit line calculated by the software, b = y intercept (also calculated by the software). The LC_{50} is obtained by solving for x.

Bioassay Data Acceptance/Rejection Criteria: Acceptable control mortality will not exceed 15%. Should control mortality exceed 15% for a given run, the experiment shall be repeated, unless otherwise authorized by the study director. If more than one larva has been inadvertently added to any well, that well will not be scored.

Reagents
PBS
8.0g sodium chloride
1 15g sodium phosphate, dibasic
0.2g potassium phosphate, monobasic
0 2g potassium chloride

Dissolve into 800 ml double deionized water using a magnetic stir plate. Check the pH (should be 7.4 ± 0.01). Bring the final volume to 1000ml. Store at 20-25C for 3 months or at 4C for a maximum of 6 months. Discard if there is any visible evidence of fungal growth.

Southland Multiple Species Diet:

dry diet 166.0g
agar water 540.0ml
cooling water 390 Oml
low melt agar 15.0g

Thoroughly sanitize blender and. all materials required for preparing and dispensing diet by rinsing with a 20% bleach solution and triple rinsing with DI water. Add agar and agar water to the flask and cover tightly with aluminum foil. Autoclave at 121'C for a minimum of 20 minutes. Add the hot agar solution and cooling water to the blender container and blend, at low speed. Add dry diet powder to the blender container and place lid on the blender, Blend on high speed for a minimum of 1 minute. Pour the diet into a beaker containing a magnetic stir bar. Place the beaker on the hot/stir plate (hot plate should be set on 2.5 and the stir bar should be stirring at a slow speed). Allow the diet to remain on the stir plate until dispensed.

Appendix G. Documents relevant to Amendments to Protocol MYCO98-001.

[redacted]

From: [redacted]
Sent: Wednesday, September 02, 1998 12:11 PM
To: [redacted]
Subject: phone message 9/2/98

[redacted]

Just picked up your voice mail of 9/2/98 regarding the typo in the Equivalency Protocol MYCO98-001. This email will be printed and signed by me (and a copy faxed to you) indicating that Section 10.4 will be amended to read. Test Material Formulation: Dilutions of test materials will be made using 0.1% bovine serum albumin (BSA).... which is a change from the original of 1%.

The reason is typo oversight, and there is no material effect on the study if performed as indicated above at 0.1%.

Thanks for the voice mail.

[redacted]

Signature 9/2/98 DATE

10-12-98

Amendment to MYCO098-001

10.2.2 Amended to read "All insects will be purchased from French Ag Res., Lamberton, MN."

10.1 Amended to read "Neonate insect larvae will be exposed to the treated diet and mortality will be assessed after six days for all species."

10.3.2 Dose levels amended as indicated in the table attached as "Bioequivalency Rate Structures" identified from Ricerca.

These are anticipated amendments clarifying the details of the protocol and do not adversely affect the conduct of the study.

[redacted] 10/12/98

Page 1 of 2

Table attached

Amendment 10

Page 282

Dipel

Rate ng Powder/cm2	TBW	FAW	ECB	CEW	BCW
0	x	x	x	x	x
0.1			x	x	
0.3	x	x	x	x	x
1	x	x	x	x	x
3	x	x	x	x	x
10	x	x	x	x	x
30	x	x	x	x	x
100	x	x			x

Bacterial Cry1F

Rate ng Powder/cm2	TBW	FAW	ECB	CEW	BCW
0	x	x	x	x	x
0.1		x	x		
0.3	x	x	x		
1	x	x	x	x	x
3	x	x	x	x	x
10	x	x	x	x	x
30	x	x	x	x	x
100	x	x	x	x	x
300	x			x	x
500		x	x		x
1000	x			x	x
3000				x	

Plant Cry1F

Rate ng Powder/cm2	TBW	FAW	ECB	CEW	BCW
0	x	x	x	x	x
3	x	x	x		
10	x	x	x	x	x
30	x	x	x	x	x
100	x	x	x	x	x
300	x	x	x	x	x
1000	x	x	x	x	x
3000	x	x	x	x	x
5000	x	x	x	x	x
10000				x	x

[redacted]

From: [redacted]
Sent: Thursday, September 17, 1998 7:21 A[redacted]
To: [redacted]
Subject: IMPORTANT

Importance: High

9-17-98

RE: GLP equivalency work

[redacted]

The folks at Pioneer just informed me that their ELISA data have been miss reported by a factor of 10 too high. This means that it would appear that the CoA for the plant material (1568-022B) should read 0.158% F instead of 1.58%. Likewise the control plant material (1568-022A) would be below limits of detection at 0.00004% instead of that reported.

Does this detrimentally affect any of your work? I have them reconfirming everything again, but this is where we stand as of late Wednesday.

Call, email or FAX your questions,

[redacted]

[redacted]

[redacted]

Appendix H: Copy of Protocol MYCO98-001.

[redacted]

MYCOGEN PROTOCOL NUMBER: MYCO98-001

TITLE: Equivalency of Microbial and Maize Expressed Cry1F Protein; Characterization of Test Substances for Biochemical and Toxicological Studies

---

1. OBJECTIVE:

1.1 The objective of this study is to determine whether activity of maize expressed Cry1F protein differs from bacterially expressed Cry1F protein. The biological and biochemical assay data generated during this study will also be used to provide characterization information for biochemical and toxicological studies performed to support the registration of transgenic plants expressing this protein.

2. INTRODUCTION:

2.1 The subject protein is an insecticidal delta-endotoxin originally identified in *Bacillus thuringiensis* subspecies *aizawai* strain PS81I. The designation of this toxin is Cry1F. This strain of *B. thuringiensis* produces several types of endotoxins, so the gene for this Cry1F protein was isolated and expressed in *Pseudomonas fluorescens* for the purpose of obtaining large quantities of substantially homogeneous Cry1F protein independent of other contaminating endotoxins.

The protein is of commercial interest because it has activity on the European corn borer (ECB, *Ostrinia nubilalis*), one of the major pests of maize in North America and western Europe. Delta-endotoxins generally are synthesized as larger protoxins, which are enzymatically cleaved to smaller active toxins.

3. SPONSOR:

[redacted]
Product Registration Manager
Mycogen Corporation
5501 Oberlin Drive
San Diego, CA 92121

4. EXPERIMENTAL DESIGN:

The equivalency and characterization will be conducted using two types of analyses, biochemical and biological. The biochemical tests performed will include the following assays: Molecular weight determination by SDS PAGE; immunoreactivity by Western analysis and ELISA; glycosolation by electrophoresis

and staining; and N-terminal sequencing by Edman analysis. The biological activity will be determined on sensitive and insensitive insects. Additionally, *in vitro* digestibility in simulated gastric fluid will be determined.

5. TESTING FACILITIES:

5.1 Isolation/Biochemical Assays:
Mycogen Corporation
5501 Oberlin Drive
San Diego, CA 92121

Preparation of the batch will be the responsibility of the sponsor and will be documented. The process will be briefly described in the final report.

5.2 Biochemical - ELISA:
Pioneer Hi Bred Intl.
7300 NW 62nd Ave.
Johnston, IA 50134

5.3 Biochemical - N-Terminal:
UCSD Biology Dept.
Protein Sequencing Facility
9500 Gilman Dr., Mail Code 0322
La Jolla, CA 92093

5.4 Biological:
Ricerca, Inc.
Department of Analytical and Biological Sciences
7528 Auburn Rd.
Painesville, OH 44077

6. PERSONNEL:

6.1 Study Director:
[redacted]
Manager, Biochemistry
Mycogen
[redacted]

6.2 Principal Biological Investigator: [redacted]

Senior Research Entomologist
Ricerca (440) 357-3352
Fax (440) 354-4662

6.3 Biochemical Investigators:

6.3.1 SDS Page, Western assay, glycosolation, *in vitro* digestibility:

Principal Biochemical Investigator
[redacted]
Mycogen
5501 Oberlin Drive
San Diego, Ca
[redacted]

6.3.2. ELISA:

Biochemical Investigator
[redacted]
Pioneer Hi-Bred Intl.
[redacted]

6.3.3 N-Terminal:

Biochemical Investigator
[redacted]
UCSD
[redacted]

7. TEST AND CONTROL SUBSTANCE IDENTIFICATION:

7.1 Test substances:

7.1.1 Bacterially derived Cry1F Protein:

7.1.1.1 Source: frozen cell paste of *P. fluorescens* MR872,

7.1.1.2 Fermentation Lot Numbers: 970819R, 980722H and 980630C

7.1.1.3 Production date: 8/19/97, 7/22/98 and 6/30/98, respectively.

7.1.1.4 Stability: Stability of the material is being determined by the sponsor under a separate project.

7.1.2 Maize derived Cry1F Protein:

7.1.2.1 Source: stored frozen leaves

7.1.2.2 Plant event: Event RL0020 / TC1365 in corn genotype 5XH751

7.1.2.3 Stability: Stability of the material is being determined by the sponsor under a separate project.

7.2 Control substance:

7.2.1 Non *Bacillus thuringiensis* Purified Cry1F truncated protein 1.4 mg/ml and bovine serum albumin

7.2.1.1 Source and Lot numbers will be recorded in the raw data.

7.2.2 Plant extract from control maize not expressing Cry1F protein

7.2.2.1 Source: stored frozen leaves

7.2.2.2 Plant genotype: Inbred genotype 5XH751

7.2.2.3 Stability: Stability of the material is being determined by the sponsor under a separate project.

7.2.3 Buffer control for bacterially derived cry1F protein

7.2.3.1 Composition, source and lot numbers will be recorded in the raw data.

7.3 Reference Substance for Biological Assay

7.3.1 DiPel® 2X

7.3.1.1 Source and Lot numbers will be recorded in the raw data.

7.4 Storage conditions of test, control and reference substances

7.4.1 Test materials will be stored according to label requirements. Storage conditions will be monitored and records will be maintained of actual storage conditions.

7.5 Test, control and reference substance handling

7.5.1 Records of receipt, use, and disposition will be maintained for all test materials. Actual amounts used will be recorded and maintained with the study raw data. Prior to disposing of any test materials, the study director will be contacted.

8. PROPOSED EXPERIMENTAL INITIATION AND COMPLETION DATES:

8.1 Experiment Initiation Date: 08/26/1998

8.2 Experiment Completion Date: 09/23/1998

8.3 Proposed Report Date: 09/30/1998

9. TEST METHODS - BIOCHEMICAL:

The following methods will be used during the study. Procedures used including any modifications, will be completely documented in the raw data and will be described in the final report.

9.1 Sodium dodecyl sulfate polyacrylamide electrophoresis (SDS PAGE) will be conducted essentially as described by Laemmli (1970).

9.2 Colorimetric test for total protein content will be conducted essentially as described by Lowry (1951) using bovine serum albumin (BSA) as a standard.

9.3 Antibody reactivity in solution will be performed by ELISA to estimate the concentration of toxin in each preparation using authentic Cry1F as a standard. Samples will be evaluated for expression levels of the Bt protein according to the ELISA protocol described in SOP #PHIAR-SM-0001/01 and PHIAR-SM-0004/01. Notebook records will reflect the conditions for analysis of each sample including sample dry weight used, sample preparation, total protein analysis and final ELISA analysis. Each sample will be analyzed 3 times (duplicate wells/analysis) on 3 different days. Sample interpolated values will only be used from standard curves which pass the quality control (QC) criteria described in the SOP and the duplicate sample wells/analysis must have a % coefficient of variation (CV) of ≤15%. If the curve does not pass the QC criteria and/or the % CV of the sample wells is >15%, the sample must be re-analyzed. The Bt protein concentration will be reported as pg Cry1F protein per a specified ug quantity of total tissue soluble protein. For each sample, the mean Cry1F protein/ug total tissue soluble protein and the standard deviation of the mean will be calculated based on the 3 values obtained from the 3 individual analyses which came from assays in which the standard curve and %CV of the sample wells passed the QC criteria.

9.4 Western blotting of toxin following SDS Page will be conducted essentially as described by Towbin *et al.*, 1979.

9.5 The susceptibility of Cry1F protein to proteolytic degradation will be evaluated in simulated mammalian gastric fluid (SGF) containing pepsin. The method used will be documented in the raw data and described in the final report.

9.6 N-terminal analysis will be conducted by separating the protein of interest by SDS PAGE essentially as described by Laemmli (1970). Following western transfer to Immobilon-Psq, the proteins will be localized on the membrane by staining with 1% Ponceau S in 1% acetic acid. The bands of interest will be excised from the membrane and subjected to Edman degradation coupled with automated gas-phase sequencing.

9.7 Glycosylation by electrophoresis and staining will be conducted using Immuno-Blot Kit for Glycoprotein Detection (BioRad #170-6490).

10. TEST METHODS - BIOLOGICAL:

10.1 Experimental Design: Test materials will be formulated and dispensed onto the surface of artificial diet in 96 well microtiter plates. Neonate insect larvae will be exposed to the treated diet and mortality will be assessed after four to six days, depending on species. Comparisons will be made to determine if the spectrum of activity (i.e., rank order of toxicity) differs between expression systems.

10.2 Test System Description

10.2.1 Species, strain, substrain:

10.2.1.1 *Ostrinia nubilalis* (European corn borer)
10.2.1.2 *Heliothis virescens* (Tobacco budworm)
10.2.1.3 *Spodoptera frugiperda* (Fall armyworm)
10.2.1.4 *Agrotis ipsilon* (Black cutworm)
10.2.1.5 *Helicoverpa zea* (Corn earworm)

10.2.2 Source of test systems: The European corn borer and the Black cutworm will be received from French Agricultural Research, Lamberton, MN. As soon as the source of the remaining insects is determined, the protocol will be amended to include the source.

10.2.3 Test stage: All insects will be neonate larvae.

10.2.4 Number of replicates: At least three of each test substance/species combination.

10.2.5 Number per replicate: 48±5 wells with one neonate larva per well.

10.2.6 Procedure for test system identification: Microtiter plates will be labeled with the species, test material, dose level, date of initial exposure and study number.

10.2.7 Justification for Test System Selection: The five insects selected differ in their sensitivity to the test substances.

10.2.8 Rearing Conditions: Eggs will be purchased from commercial sources and shipped by overnight delivery service. Upon arrival at Ricerca, they will be placed in an incubator maintained at 10 ± 5 °C until needed for testing. To obtain neonate larvae, eggs will be surface sterilized using a solution of bleach (4 ml commercial laundry bleach in 100 ml deionized water containing surfactant).

10.3 Dosing Description

10.3.1 Route of Administration and Justification: Test materials will be administered using diet surface overlay techniques in standardized 96 well microtiter plates. The bioassay procedure is standard for evaluating the toxicity of stomach active insecticides to lepidopteran larvae.

10.3.2 Dose Levels: Eight doses (ng Cry1F/cm^2) plus an untreated control consisting of 0.1% BSA in phosphate buffer. Concentrations will be chosen so at least two produces >50% mortality, two produce <50% mortality and at least five produce between 0 and 100% mortality. Range finding bioassays will be performed to determine the actual dose levels to use for each test material for the definitive study. The dose levels selected for each species and test material will be included in the protocol by amendment. Due to the experimental nature of this study, it may be necessary to override these criteria.

10.3.3 Duration of exposure: Four to six days. Exact time of exposure for each species will be determined during range finding bioassays. A protocol amendment will be issued prior to initiation of the definitive test to include exposure duration.

10.4 Test Material Formulation: Dilutions of test materials will be made using 1% bovine serum albumin (BSA) in phosphate buffer. Fifteen (15) microliter aliquots of each test solution concentration will be pipetted on the surface of diet in each well (total of 48/plate) and allowed to air dry.

10.5 Description of Diet and Water: Insect diet will be prepared (Attachment A). The ingredients are not known to contain contaminants which interfere with the conduct of the study. Deionized water will be used throughout the project. If a particular species will not perform

appropriately on this diet, the diet actually used will be included in the protocol by amendment.

10.6 Exposure conditions: After the diet has dried and insect larvae have been added, each tray will be heat sealed with mylar film and placed in an environmental chamber maintained at 25 ± 4 °C and approximately 60 ± 10% relative humidity.

10.7 Method Used to Control Bias: Sheets containing approximately 1000 to 2000 eggs will be placed in a container for hatching. Neonate larvae will be selected without regard to sex for testing based on activity and size. Larvae that are inactive or obviously non-neonate will not be used.

10.8 Test System Observations:

10.8.1 Mortality will be assessed after four to six days, depending on species. Optimal exposure conditions for each species will be determined during range-finding bioassays and will be included in the protocol by amendment.

10.8.2 Larvae which are not stunted and respond to a light touch with a probe will be classified as alive.

10.9 Test System Disposition:

10.9.1 No specimens will be kept. At the end of the observation period, the test insects will be autoclaved or frozen and disposed of as laboratory waste according to Federal, State and local regulations.

11. PROPOSED STATISTICAL TESTS:

11.1 Biochemical Assays: Means and standard deviations will be generated as appropriate. No other statistical tests will be performed on the data, except as noted in section 9.3. Correlation coefficients for standard curves will be determined, where appropriate.

11.2 Bioassays:

11.2.1 Percent mortality will be calculated at each dose level. Abbott's formula will be used to correct for control mortality.

11.2.1.1 Abbott's formula:

$$\% \text{ Corrected Mortality} = \frac{100 * (\% \text{ mortality}_{\text{treatment}} - \% \text{ mortality}_{\text{control}})}{100 - \% \text{ mortality}_{\text{control}}}$$

11.2.2 Probit analysis using POLO-PC (LeOra Publishing, Berkeley, CA) or equivalent software will be used to calculate LC_{50}, LC_{90}, and slopes.

11.2.3 Bioassay Data Acceptance/Rejection Criteria:

11.2.3.1 Acceptable control mortality will not exceed 15%. Should control mortality exceed 15% for a given run, the experiment shall be repeated, unless otherwise authorized by the study director.

11.2.3.2 If more than one larva has been inadvertently added to any well, that well will not be scored.

11.3 Other statistical tests may be performed at the discretion of the study director. A protocol amendment will be issued describing any additional tests planned for use during the study.

12. RECORDS TO BE MAINTAINED:

12.1 All manually recorded raw data, calculations and results will be recorded in ink. Errors will be lined out with a single solid line to not obscure the original entry. Errors will also be initialed and dated by the recording scientist and a reason for the error will accompany each change. All raw data including measurements, results, and laboratory notebooks in support of this study will be retained. Automated data collection shall comply with the GLP regulations. The software version shall be recorded in the raw data.

12.2 The records to be maintained include but are not limited to the following:

12.2.1 Test and reference substance receipt, use, storage conditions and disposal

12.2.2 Description of equipment used including manufacturer, model number, operating conditions and software name and version.

12.2.3 Equipment maintenance, calibration/standardization/testing records

12.2.4 Standard, buffer and wash solution preparation; reagent names, manufacturers, lot numbers, and grade/purity

12.2.5 All observations and measurements

12.2.6 Correspondence and memorandum relating to study conduct or interpretation

12.2.7 Protocol and any protocol amendments

12.2.8 Protocol and/or SOP deviations

13. ARCHIVES:

13.1 The protocol, raw data and a copy of the final report shall be archived at the termination of the study. Each test site shall be responsible for archiving their facility records. Study raw data shall be transferred to Mycogen for permanent archival. Test sites shall maintain a copy of the protocol and raw data generated at their facility for a period of three years.

14. REGULATORY REQUIREMENTS/GOOD LABORATORY PRACTICES:

14.1 This protocol will be followed in accordance with FIFRA Guidelines for Good Laboratory Practices (GLP) 40 CFR 160. The study will be conducted in accordance with each test site's current Standard Operating Procedures.

15. OWNERSHIP OF STUDY RECORDS:

15.1 All information, data, and the like generated by, and all inventions and copyrights resulting from the Services hereunder, shall be the sole and exclusive property of the Sponsor.

16. PROTOCOL AMENDMENTS:

16.1 This study may be modified during its progress. All modifications or additions will be documented in the form of an amendment that describes what is being changed or added, the reason for the change, and the date the change took effect. All amendments will be signed by the Study Director prior to application in the study. When an amendment is required without sufficient time for the issuance of a written amendment, that change may be affected verbally by the study director and followed by a written amendment. In this case, the effective date of the written amendment will be the date of the verbal change. Correspondence concerning changes to the protocol shall be documented in the raw data.

17. GOVERNMENT INSPECTION:

17.1 The testing site shall notify the Sponsor of any private or Governmental request for confidential information relating to this study. The Sponsor shall have the right to participate in the testing site's response to any such request.

18. QUALITY ASSURANCE:

18.1 The study shall be inspected and records maintained by the QAU of each respective organization involved in the study. The QAU at each test site shall report their findings to the Study Director and to Mycogen's Quality Assurance Unit. Mycogen's QAU will assure test site QA findings are forwarded to Mycogen management. Mycogen will provide the QA function for the biochemical assay performed at UCSD.

19. SAFETY AND HEALTH:

19.1 Laboratory personnel will practice proper safety and health habits according to facility Standard Operating Procedures or policies.

20. PRINCIPAL INVESTIGATOR REPORTS:

20.1 The principal investigators shall prepare reports for their portions of the study to be submitted to the study director for inclusion in the final report. These reports shall minimally contain the following:

20.1.1 Compliance Statement signed by the Principal Investigator

20.1.2 Inspection Statement from Quality Assurance Unit including the dates the study was inspected, the phase inspected, the name of the inspector, the dates the findings were reported to the study director, the dates the inspections were reported to the principal investigator and internal management

20.1.3 Names of the personnel involved in the study

20.1.4 Experimental initiation and termination dates

20.1.5 Identification and description of the test material

20.1.6 Presentation of all pertinent data in tabular format when appropriate

20.1.7 Description of test procedures

20.1.8 Deviations, if any, from this protocol and SOPs

21. Final Report:

21.1 The final report will be formatted according to EPA notice PR 86-5. The final report will include all requirements as described in the Good Laboratory Practice Standards section 160.185, including:

21.1.1 Inspection Statement from Quality Assurance Unit including the dates the study was inspected, the phase inspected, the name of the inspector, the dates the findings were reported to the study director and management

21.1.2 GLP Compliance Statement signed by the Study Director

21.1.3 Names of the personnel involved in the study

21.1.4 Dates of study experimental initiation and termination

21.1.5 Identification and description of the test material

21.1.6 Presentation of all pertinent data in tabular format when appropriate

21.1.7 Description of test procedures

21.1.8 Deviations, if any, from this protocol and their possible effect on the study outcome

21.1.9 Summary and analysis of the data and conclusions

21.1.10 Principal investigator reports

22. REFERENCES:

22.1 Laemmli, U.K. (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature (London)* **227**:680-685.

22.2 Lowry, O.H., Roseborough, N.J., Farr, A.L. and Randall, R.J. (1951) Protein measurement with the Folin-phenol reagent. *J. Biol. Chem.* **193**:265-272.

22.3 Towbin, H., Staehlin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* **76**:4350-4354.

22.4 Methods in Enzymology "Enzyme Structure Part I" (C.H.W. Hirs and S.N. Timasheff, eds.) , Vol. 91, Academic Press, New York, 1983.

TITLE: Equivalency of Microbial and Maize Expressed Cry1F Protein; Characterization of Test Substances for Biochemical and Toxicological Studies

23. SIGNATURES:

Acknowledged by:

[redacted] 980820

Principal Biochemical Investigator: Date

[redacted] 980824

Principal Biological Investigator Date

[redacted] 08/25/98

Study Director Date

Approved by:

[redacted] 08/25/98

Sponsor Date

Reviewed by:

[redacted] 8-25-98

Mycogen Quality Assurance Date

Attachment A
Diet Preparation

The ingredients, their sources and their proportions for a standard, one liter batch are listed in the table below:

Name	Source	Catalog No.	Amount Used
Nutri-Soy Flour	Bioserv	1500	30.12 g
Ground Wheat Germ	Bioserv	1661	25.68 g
Wesson Salt	ICN Biomedicals	902851	6.96 g
Sugar	Local Grocery	----	30.12 g
Methyl Paraben	Bioserv	7685	720 mg
Sorbic Acid	ICN Biomedicals	102937	720 mg
Vitamins	Bioserv	6265	6.96 g
Agar	Bioserv	7060	16.40 g
Deionized Water	City of Painesville, OH	----	682.4 ml

1. Mix agar with deionized water in a glass beaker.
2. Cover beaker with lid. To aid the dissolution of the agar, use either a microwave oven (batches less than 1 liter) or an autoclave (batches greater than 1 liter).
3. Heat the agar solution until just below the boiling point, stirring occasionally.
4. Cool the solution, stirring as necessary.
5. After agar solution is cooled, add it to the remaining ingredients according to manufacturer's instructions.
6. This mixture may be stored under refrigerated conditions for up to seven days.
7. Different size batches may be prepared by adjusting proportions accordingly.

INVESTIGATOR'S REPORTS

Appendix I: Report of [redacted], Mycogen, Biochemical Equivalency

Qualitative Analyses and Comparison of Cry1F as Obtained from Maize and Bacterial Origin

Table of Contents

Summary

A synthetic gene encoding a truncated version of Cry1Fa2 protein derived from *Bacillus thuringiensis* subspecies *aizawai* strain PS81I, has been introduced into *Zea mays*. Although the DNA sequence of the cry1Fa2 transgene differs from the wild-type gene in codon usage patterns, to enable higher levels of expression and thereby greater efficacy, the truncated Cry1Fa2 protein encoded therefrom is nearly identical to the Cry1Fa2 toxin produced by PS81I. In this study, Cry1Fa2 as obtained from maize leaf tissue (eucaryotic source) was compared functionally, biochemically and immunologically with Cry1Fa2 obtained from bacteria (prokaryotic sources). In every instance examined, Cry1Fa2 as obtained from maize was found to be substantially similar to Cry1Fa2 produced in the bacterium.

Introduction

One of the major pests of maize in North America and western Europe is the European corn borer (ECB), *Ostrinia nubilalis*. Owing to the cryptic feeding behavior of this insect and its susceptibility to *Bacillus thuringiensis* (Bt) Cry1Fa2, the gene encoding this toxin was selected as a maize transgene. The cry1Fa2 gene was initially derived from Bt subsp. *aizawai*, strain PS81I. A genetically truncated version of this gene was resynthesized to reflect biases in *Zea mays* codon usage and inserted into the *Zea mays* genome under the constitutive control of the ubiquitin (UBI) promoter. UBI, originally isolated from maize, drives constitutive expression in monocots in most tissues.

The genetically truncated Cry1Fa2 protein (*ca.* 68.2kDa) recovered from transformed maize tissue differs in only one amino acid from full-length Cry1Fa2 protoxin produced by Bt subsp. *aizawai* strain PS81I following solubilization and proteolytic processing in the susceptible insect gut (Figure 1). The one amino acid substitution, a leucine for phenylalanine substitution at amino acid residue 604, is conservative, *ie.*, both are hydrophobic amino acids with non-polar side chains.

The immunologic reactivity of this transgene product, as gauged by an antibody derived from bacterially expressed Cry1Fa2, is virtually indistinguishable from the parent antigen. Although multiple molecular weight species appear to react with the antibody, N-terminal sequence analysis reveals that each of these polypeptides are derived from the *ca.* 68.2kDa transgene product. The origin of these various molecular weight species is most probably the result of partial proteolysis by resident maize proteases. Further evidence for proteolysis-generated products of the maize Cry1Fa2 is provided by examination of the amino acid sequence adjacent to that determined for the N-terminus of each of the polypeptides. In each instance, protease sensitive residues are encountered in the immediate N-terminal direction of the exposed residue.

Studies designed to determine whether maize-derived Cry1Fa2 is glycosylated were apparently negative. Although the Cry1Fa2 toxin contains over 90 canonical sites for O- and N-glycosylation, no evidence for such modifications were obtained.

Finally, an *in vitro* digestion technique, simulating mammalian ventricle enzymic decomposition of Cry1Fa2 by incubation with pepsin-HCl was investigated. These studies demonstrate that Cry1Fa2 is completely proteolyzed to amino acids and small peptides within 5min at molar ratios approximating 1:100 (pepsin:Cry1Fa2) and that this same molar ratio effects nearly complete proteolysis of the Bt toxin within one minute. The fraction of the peptide bonds hydrolyzed has previously been shown to correlate with independent *in vivo* determinations of the digestibilities of proteins and other foodstuffs.

Materials and Methods

Protein Assay

Protein content of the maize and bacterial Cry1Fa2 extracts were determined using a Modified Lowry Protein Assay (Pierce Chem. Co., Rockford, IL). The modification involves a change in the composition of the original Lowry alkaline cupric sulfate - tartrate reagent (Lowry, *et al.*, 1951) by the addition of potassium iodide. Various dilutions of the sample extracts are diluted in water and 0.2ml of the diluted material is mixed with 1.0ml of the modified Lowry alkaline cupric sulfate - tartrate reagent. Following a 10min incubation at room temperature, 0.1ml of 1.0N Phenol Reagent is added and incubated at room temperature for 30min. The samples are then read at 750nm in a UV-Vis Scanning Spectrophotometer (Shimadzu UV-2101PC). The protein concentrations for the unknown samples were calculated from a standard curve (absorbance vs. standard concentration). The standard curve was fit by and the sample Cry1Fa2 concentration iterated by using the quadratic equation. Bovine serum albumin (no. 23209, Pierce Chem. Co., Rockford, IL) served as the standard.

SDS PAGE

Sodium dodecyl sulfate polyacrylamide electrophoresis (SDS PAGE) was performed essentially as described by Laemmli (1970) using Bio-Rad Laboratories (Hercules, CA) reagents and molecular weight standards (Broad Range, no. 161-031). Sample preparations were mixed 1:2 with Sample Buffer and boiled 5min. Samples were subjected to SDS PAGE using 10% polyacrylamide gels (Bio-Rad Laboratories, Hercules, CA), unless noted otherwise in the Figure legends. Electrophoresis was conducted at 200V at 100mA for approximately 35min. The proteins were visualized by staining in a solution of Coomassie Brilliant Blue R-250 (Bio-Rad Laboratories, Hercules, CA). Notebook refs. 1569-012 to 1569-040.

Western blotting

Western blotting of protein from SDS electrophoretograms was conducted as described (Towbin *et al.*, 1979) with minor variations. Kaleidoscope prestained molecular weight standards were used (no. 161-0324, Bio-Rad Laboratories, Hercules, CA). Following electrophoretic transfer of the proteins from the SDS gel to Protran nitrocellulose (Schleicher & Schuell, Keene, NH) in the presence of Towbin Transfer Buffer, the unoccupied binding sites on the membrane were blocked 1hr with a 1% (w/v) solution of Rad-Free blocking powder (Schleicher & Schuell, Keene, NH) in PBS. The membranes were probed with a 1:1000 dilution of rabbit polyclonal anti-Cry1Fa2 antibody, lot number 091796-P (Pioneer HiBred, Johnston, IA), by incubating 1hr at room temperature in the blocking solution described above. Immunoblots were developed by incubation with an alkaline phosphatase-conjugated goat anti-rabbit (no. A3687, Sigma Chemical Co., St. Louis, MO.) at a 1:5000 dilution in the blocking solution described above. Following extensive rinsing in PBS, the western blot was colorimetrically developed by a 10min incubation

in 10ml nitro-blue tetrazolium chloride/5-bromo-4-chloro-3'-indolyphosphate *p*-toluidine salt (1-Step™ NBT/BCIP, Pierce Chem. Co., Rockford, IL).

N-terminal Sequence Analysis

The maize-expressed Cry1Fa2 was concentrated before blotting by affinity chromatography using rabbit polyclonal anti-Cry1Fa2 antibody, lot number 091796-P (Pioneer HiBred, Johnston, IA) coupled to Sepharose 4B using cyanogen bromide-activated Sepharose 4B (Sigma Chemical Co., St. Louis, MO). SDS PAGE and western blotting were as described for western blotting with the following exceptions. Transblot PVDF (no. 162-0180, Bio-Rad Laboratories, Hercules, CA) was used in place of nitrocellulose and CAPS Transfer Buffer was used in place of Towbin Transfer Buffer. Following transfer of the proteins from the SDS electrophoretogram (100V x 1hr), the membrane was rinsed in several changes of distilled water. The proteins of interest were localized on the membrane by staining with 0.1% (w/v) Coomassie Brilliant Blue R-250 in 50% (v/v) methanol for 5min at room temperature followed by partial destaining in several changes of 50% (v/v) methanol containing 10% (v/v) acetic acid. The membrane was allowed to dry. The bands of interest were excised from the membrane and subjected to Edman degradation coupled with automated gas-phase sequencing (Matsudaira, 1987) at the UCSD Protein Sequencing Facility (La Jolla, CA).

Glycoprotein Detection

Detection of potential glycosylation of Cry1Fa2 *in planta* was enabled by first subjecting the plant- and bacterially derived Cry1Fa2-containing extracts to SDS PAGE and transfer Protran nitrocellulose (Schleicher & Schuell, Keene, NH) as described in the preceding sections. Following transfer of the proteins to the membrane including ovalbumin as a positive glycoprotein standard, the nitrocellulose was washed with PBS at room temperature for 10min with gentle agitation. The membrane was immersed in 10ml of Reagent A and incubated in the dark at room temperature for 20min with gentle agitation. The membrane was washed three times with PBS at room temperature for 10min each with gentle agitation. The membrane was immersed in 10ml Biotinylation Solution for 60min at room temperature with gentle agitation. The membrane was washed three times with TBS at room temperature for 10min each with gentle agitation. The nitrocellulose was incubated in 10ml Blocking Solution overnight at 4°C. The membrane was washed three times with TBS at room temperature for 10min each with gentle agitation. The membrane was immersed in 10ml Conjugate Solution and incubated at room temperature for 60min with gentle agitation. The membrane was washed three times with TBS at room temperature for 10min with gentle agitation. Carbohydrate was localized by immersing the membrane in the alkaline phosphatase detection reagent, 10ml nitro-blue tetrazolium chloride/5-bromo-4-chloro-3'-indolyphosphate *p*-toluidine salt (1-Step™ NBT/BCIP, Pierce Chem.

Co., Rockford, IL) for 10min at room temperature. The reaction was stopped by rinsing in water and drying the membrane.

Pepsin-HCl digestion

Crystallized pepsin A [EC 3.4.23.1] from porcine gut mucosa was obtained from Sigma Chemical Co., St. Louis, MO (no. P6887). The assay conditions were as follows. 90ul 0.15M HCl was added to a 1.5ml eppendorf polypropylene microfuge tube. 10ul of a 5mg per ml solution of trypsinolyzed Cry1Fa2 as obtained from bacterial origin was added and mixed well. For the time-course of digestion, 10ul of a 50ug per ml solution of pepsin was added and allowed to incubate 0, 1, 2, 4, 10, 15, 30 and 60min. For the concentration-dependence of digestion experiments, 10ul of a pepsin solution serially diluted from a 5.0mg per ml stock pepsin solution (0.5mg, 50ug, 5.0ug, 0.5ug, 50ng and 5.0ng per ml) were added and incubated 5min at room temperature. These amounts correspond to a 1:1, 1:10, 1:100, 1:1000, 1:10000: 1:100000 and 1: 1000000 ratio of pepsin:Cry1aF2 (w/w) or 1:1.9, 1:18.8, 1:188, 1:1883, 1:18830, 1:188300 and 1:1883000 molar molar ratio of pepsin:Cry1aF2.

Since pepsin is unstable above pH 6.0, the reactions were terminated by the addition of 2M Tris base to a final concentration of 167mM, imparting a final pH of 7.0 to the sample. Sample preparations were mixed 1:2 with Sample Buffer and boiled 5min. Samples were subjected to SDS PAGE using 10% polyacrylamide gels (Bio-Rad Laboratories, Hercules, CA). Electrophoresis was conducted at 200V at 100mA for approximately 35min. The proteins were visualized by staining as described in the SDS PAGE section above.

Buffers and Solutions
Unless indicated otherwise, all reagents listed below were obtained from generally available commercial sources.

Towbin Transfer Buffer
25mM Tris-Cl, pH ~ 8.3
192mM glycine
20% methanol

CAPS Transfer Buffer
10mM CAPS, pH 11
10% methanol

Phenol Reagent
1.0N Folin-Ciocalteu reagent

Tris Buffered Saline (TBS)
50mM Tris, pH 7.2
27mM sodium chloride

Phosphate Buffered Saline (PBS)
9mM sodium phosphate, pH 7.2
27mM sodium chloride

Glycoprotein Reagent A
100mM sodium acetate, pH 5.5
10mM sodium periodate
5mM Na_4EDTA

Glycoprotein Hydrazide Solution
Vial B contents (Bio-Rad Labs. Immuno-blot Kit for Glycoprotein Detection #170-5490)
500ul dimethylformamide

Glycoprotein Solution 7
100mM sodium acetate, pH 5.5
5mM Na_4EDTA

Glycoprotein Biotinylation Solution
2ul Glycoprotein Hydrazide Solution
10ml Glycoprotein Solution 7
prepared immediately before use

Glycoprotein Blocking Solution
0.5g vial C (Bio-Rad Labs. Immuno-blot Kit for Glycoprotein Detection #170-5490)
100ml TBS

Glycoprotein Conjugate Solution
5ul Streptavidin-alkaline phosphatase (Bio-Rad Labs. Immuno-blot Kit for Glycoprotein Detection #170-5490, Regeant D)
10ml TBS

Results

Lowry Protein Determination

Control Corn	4.67mg protein per 5mg substance
Cry1Fa2 Corn	4.08 mg protein per 5mg substance
PHB	4.09 mg protein per 5mg substance
Powder 1	4.27 mg protein per 5mg substance

Notebook refs. 1569-052 & 1569-056.

SDS PAGE

SDS PAGE of maize-expressed Cry1Fa2 and trypsinolyzed Cry1Fa2 as obtained from bacterial origin is depicted Figure 2. The bacterial-derived material shows a predominance of the *ca.* 66kDa molecular weight species with minor polypeptides and proteins of approximately 100, 45, 20 and 15kDa. Computer-assisted analysis of the Cry1Fa2 trypsin resistant core results in a molecular weight prediction of 65912Da. Thus, the principal electrophoretic species observed corresponds to trypsinolyzed Cry1Fa2. The extracts corresponding to Cry1Fa2-maize and corresponding control in lanes 4 and 6 of Figure 2 appear indistinguishable. Each contain predominate proteins of *ca.* 100kDa and several minor bands randging in molecular weight from 40-80kDa. Notebook refs. 1569, pp. 035-37.

Western blotting

The immunoreactivity of maize- and bacterially-derived Cry1Fa2 is shown in Figure 3. Maize-expressed Cry1Fa2 appears as a doublet band at approximately 65kDa. Neither of these proteins are evident by coomassie blue staining of SDS PAGE (Figure 2). Examination of the amino acid sequence of Cry1Fa2 as illustrated in Figure 1 shows the difference between the natural N-terminus of Cry1Fa2 and the resulting N-terminus following proteolysis by *cf.* trypsin (Gill *et al.*, 1992). The net difference between the two N-termini is 27 amino acids. This corresponds to 3220Da which is the approximate difference in the estimated molecular weight of the two reactive polypeptide species observed in lane 4 of Figure 3. Both preparations of bacterially-derived, trypsinolyzed Cry1Fa2 contain a number of Cry1Fa2 antibody reactive polypeptides. The most prominent of these have apparent molecular masses equivalent to 65, 25 and 20kDa. The 65kDa species are the principle coomassie-stained bands seen in lanes 8 and 10 of Figure 2 and correspond to the molecular weight predicted for trypsinolyzed Cry1Fa2 (*ca.*, 66kDA). Notebook refs. 1569, pp. 018-023.

Glycoprotein Detection

Detection of potential carbohydrate structure in Cry1Fa2 as expressed *in planta* was assessed by the periodate oxidation/ biotinylation/ streptavidin-alkaline phosphatase/ colorimetric method.

The Cry1Fa2 protein toxin expressed in maize has a number of potential N-acetylglycosylation sites specific to the consensus sequences asparagine-X-serine/threonine or asparagine-X-cysteine (Miletich and Broze, 1990), Figure 4A. Possible sites for O-glycosylation, serine or threonine residues, are also indicated in Figure 4A.

The detection technique used here demonstrated that there was no apparent post-translational modification of Cry1Fa2 in maize-derived extracts involving carbohydrates. Figure 4B shows that while the bacterially-derived Cry1Fa2 materials are essentiallly bereft of reactivity, the maize-derived materials exhibit a number of discrete molecular weight species. None of these co-migrate with mobilities concordant with the immunoreactive species obtained in the western blot (Figure 3).

With respect to the apparent lack of N-acetyglycosylation it should be emphasized that presence of the consensus tripeptides, asparagine-X-serine/threonine or asparagine-X-cysteine is not sufficient to conclude that an asparagine residue is a suitable site for glycosylation since folding of the protein plays an important role in the regulation of N-glycosylation (Pless and Lennarz, 1977). In like manner, the specificities of O-glcosylation are presently unknown. The specificity seems to be modulated by sequence context, secondary structure and surface accessibility (Hansen, *et al.*, 1998). Notebook refs. 1569, pp. 025-037.

N-terminal Sequencing

Maize-expressed Cry1Fa2 and trypsinolyzed Cry1Fa2 as obtained from bacterial origin was blotted (Figure 5) and subjected to N-terminal analysis. In each case, multiple signals were obtained for each round of the Edman reaction. Bolded entries are those that agree with the predicted sequence. The abbreviation Xaa refers to an unknown assignment in the Edman reaction.

Cry1Fa2 Corn, upper band Gly- (Leu, Val)-(Asp, Thr)

Cry1Fa2 Corn, lower band (Ser, Gly, Glu)-(Phe, Val, Thr)-(Gly, Arg)-Xaa-Leu PHB Xaa-Leu-Leu-Arg-(Pro, Phe, Xaa)-(Val, Leu)-(Asp, Pro)-(Gly, Ile)

Powder 1 (Ser, Phe)-(Thr, Leu)-(Gly, Leu)-(Ser, Arg)-(Glu, Leu)-(Pro, Phe)-(Val, Leu)

Examination of the first 50 amino acid residues of Cry1Fa2,

NH_2 — MENNIQNQCVPYNCLNNPEVEILNEERSTGRLPLDISLSLTRFLLSEFVP — COOH

shows that the protease sensitive site at the carboxyl group of residue 27 is immediately followed by a serine (S) residue (underline) . Modification of the serine amino acid side chain often leads to partial blocking and lower yields per cycle in the Edman reaction. It is believed that there was significant N-terminal blocking in the samples, particularly the maize samples, leading to relatively low yields through the early rounds of the Edman reaction. There also appears to be evidence of further processing at the C-terminus of arginyl residue number 42. This is apparent by lining up the double amino acid assignments of Powder 1. The non-bolded residues correspond to tryptic processing at arginyl residue 42. This additional protease sensitive site contributes to the raggedness of the Cry1Fa2 N-terminus regardless of origin (bacterial or maize). Notebook refs 1568, pp. 028-034 and 1569, pp. 041-044

In vitro Digestibility

Pepsin, an acidic protease, is the principal protease of vertebrate gastric juice. Its inactive, zymogen form, pepsinogen, is produced by the gut mucosa. Following its activation to pepsin and discharge into the acidic environment of the gut, pepsin preferentially cleaves proteins at the carboxylic groups of aromatic amino acids (*cf.*, phenylalanine and tyrosine). In vertebrates, a second round of digestion occurs as the gut contents empty into the duodenum where they are subjected to a higher pH of ~6.5 and various pancreatic enzymes. These include lipase, amylase and the serine protease, trypsin. Pancreatic lipase catalyzes the hydrolysis of emulsified esters of glycerol and long chain fatty acids (Brindley, 1985). α-Amylase hydrolyzes internal [α]-1,4-glucan links in certain polysaccharides (Takeshita and Hehre, 1975). Trypsin, a pancreatic serine protease, catalyzes the hydrolysis at the carbonyl group of amino acids with positively charged side chains and also exhibits esterase and amidase activities (Brown and Wold, 1973). Pancreatin, a preparation of enzmes prepared from mammalian pancreas contains all three of these enzymes as well as ribonucleases and other proteases.

An *in vitro* methodology which mimics *in vivo* vertebrate upper gastrointestinal digestion was developed. The digestibility of bacterially derived Cry1Fa2 protein toxin was determined by exposing suspensions of the protein at room temperature to a simulated gastric juice, containing either varying amounts of pepsin for a set incubation period or a set amount of pepsin incubated for varying periods of time.

Ratios of Cry1aF2:pepsin (w/w) tested ranged from 1:1 to 1:1000000. This corresponds to a molar ratio of Cry1aF2:pepsin ranging from approximately 1:2 to 1:1883000. Cry1Fa2 was completely proteolyzed to amino acids and small peptides within 5min at molar ratios approximating 1:100 (Cry1Fa2:pepsin) as shown in Figure 6B. In separate experiments, a molar

ratio of 1:100 (Cry1Fa2:pepsin) effected nearly complete proteolysis of the Bt toxin within one minute (Figure 6A). The doublet corresponding to a molecular weight of ~40000kDa probably represents the zymogen form, pepsinogen (41000Da, Amon and Perlmann, 1963) and pepsin A (35000Da, Bovey and Yanari, 1960). Owing to the observed completeness of the reaction, the inclusion of a second compartmentalization reaction simulating the passage and resonance within the small intestinal juice, containing pancreatin and sodium chloride, pH 6-7, was omitted. Notebook refs. 1569, pp. 051- 058 and 060-061.

References

Amon, R. and Perlmann, G. (1963) The amino acid composition of chromatographically purified pepsinogen. *J. Biol. Chem.* **238**:653.

Bovey, F. and Yanari, S. (1960) Pepsin, *in* The Enzymes. (Boyer, P., Lardy, H., and Myrback, K., eds.), Vol. IV, Academic Press, NY, 63.

Brindley, D.N. (1985) Metabolism of triacylglycerols *in* Biochemistry of lipids and membranes (Vance, D.E. and Evans, J.E., eds.), Benjamin/Cummings Publishing Co., CA, 213.

Brown, W. and Wold, F. (1973) Alkyl isocyanates as active-site-specific reagents for serine proteases. Reaction properties. *Biochem.* 12, 828.

Convents, D., Houssier, C., Lasters, I. and Lauwereys, M. (1990) The *Bacillus thuringiensis delta* endotoxin: Evidence for a two domain structure of the minimal toxin fragment. *J. Biol. Chem.* **263**:1369.

Gill, S.S., Cowles, E.A. and Pietrantonio, P.V. (1992) The mode of action of *Bacillus thuringiensis* endotoxins. *Annu. Rev. Entomol.* **37**:615.

Hansen, J.E., Nielsen, J., Lund, O., Rapacki, K. and Brunak, S. (1998) O-GLYCBASE Version 3.0: a revised database of O-glycosylated proteins. *Nucleic Acids Research*, **25**:278.

Höfte, H., Greve, H., Seurinck, J., Jansens, S., Mahillon, J., Ampe, C., Vandekerckhove, J., Vanderbruggen, H., Van Montagu, M., Zabeau, M. and Vaeck, M. (1986) Structural and functional analysis of a cloned delta endotoxin of *Bacillus thuringiensis* Berliner. *Eur. J. Biochem.* **161**:272.

Laemmli, U.K. (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature (London)* **227**:680.

Lowry, O.H., Roseborough, N.J., Farr, A.L. and Randall, R.J. (1951) Protein measuremant with the Folin-phenol reagent. *J. Biol. Chem.* **193**:265.

Matsudaira, P. (1987) Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* **262**:10035

Miletich, J.P. and Broze, G.J. Jr. (1990) Beta protein C is not glycosylated at asparagine 329. The rate of translation may influence the frequency of usage at asparagine-X-cysteine sites. *J. Biol. Chem.* **265**:11397.

Pless, D.D. and Lennarz, W.J. (1977) Enzymatic conversion of proteins to glycoproteins. *Proc. Natl. Acad. Sci. U.S.A.* **74**:134.

Ryle, A. (1970) The porcine pepsin and pepsinogens, *in* Methods in Enzymology (Perlmann, G., and Lerand, L., eds.), Vol. XIX, Academic Press, NY, 316.

Takeshita, M. and Hehre, E. (1975) The capacity of α-amylases to catalyze the non-hydrolytic degradation of starch and glycogen with formation of novel glycosylation products. *Arch. Biochem. Biophys.* **169**:627

Towbin, H., Staehlin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* **76**:4350.

Figures

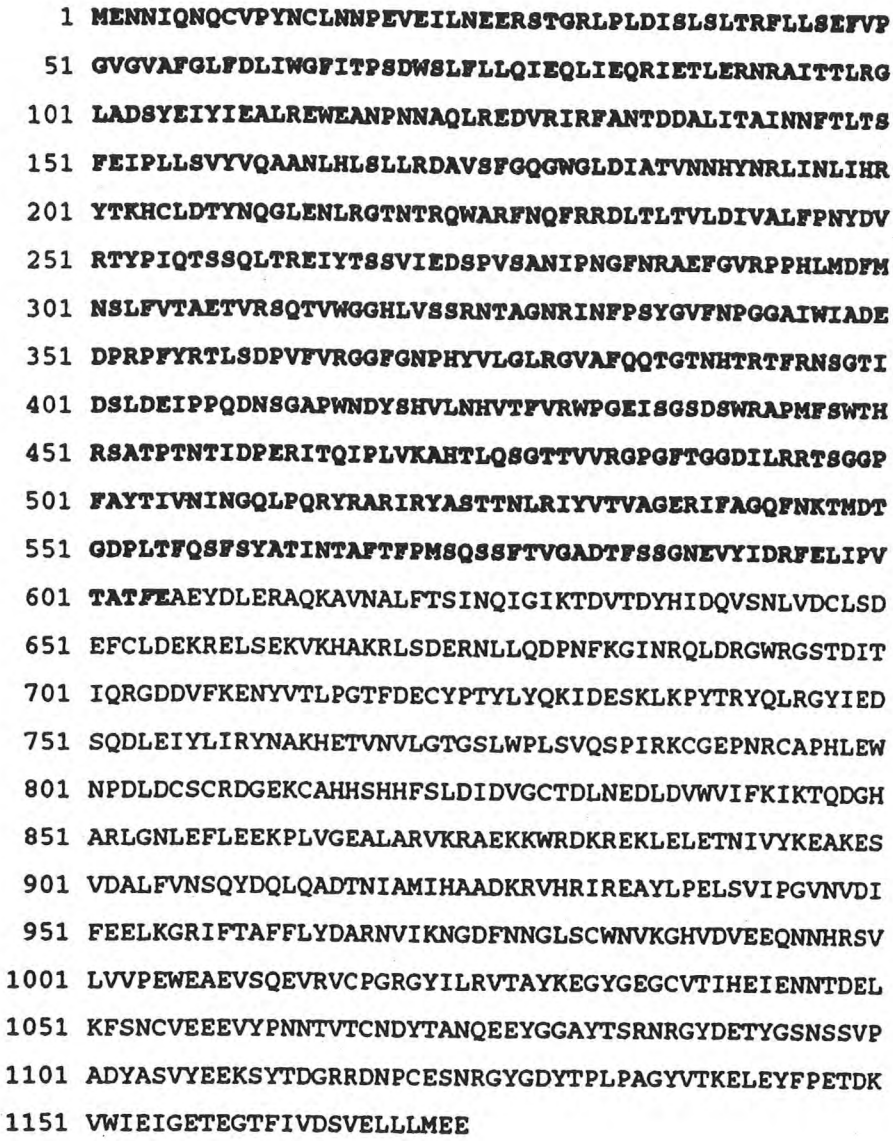


Figure 1. Amino acid sequence of MR872 and synthetic *cry1F* gene products. The amino acid sequence encoded by plant-optimized synthetic *cry1F* gene transformed into maize is indicated in bold (amino acids 1-605). The MR872 *cry1F* gene encodes amino acids 1-1174. Treatment with trypsin under alkaline conditions results in an N-terminal truncation near amino acid residue 28 (C-terminal of arginine 27) and a C-terminal truncation at the carboxyl group of the arginyl residue encountered at amino acid residue number 612. One amino acid difference exists between the gene expression products of MR872 and the synthetic maize gene. This is a leucine (L) for phenylalanine (F) substitution at amino acid residue 604 (shown in italics). This conservative amino acid substitution was the result of the introduction of a restriction site for cloning purposes.

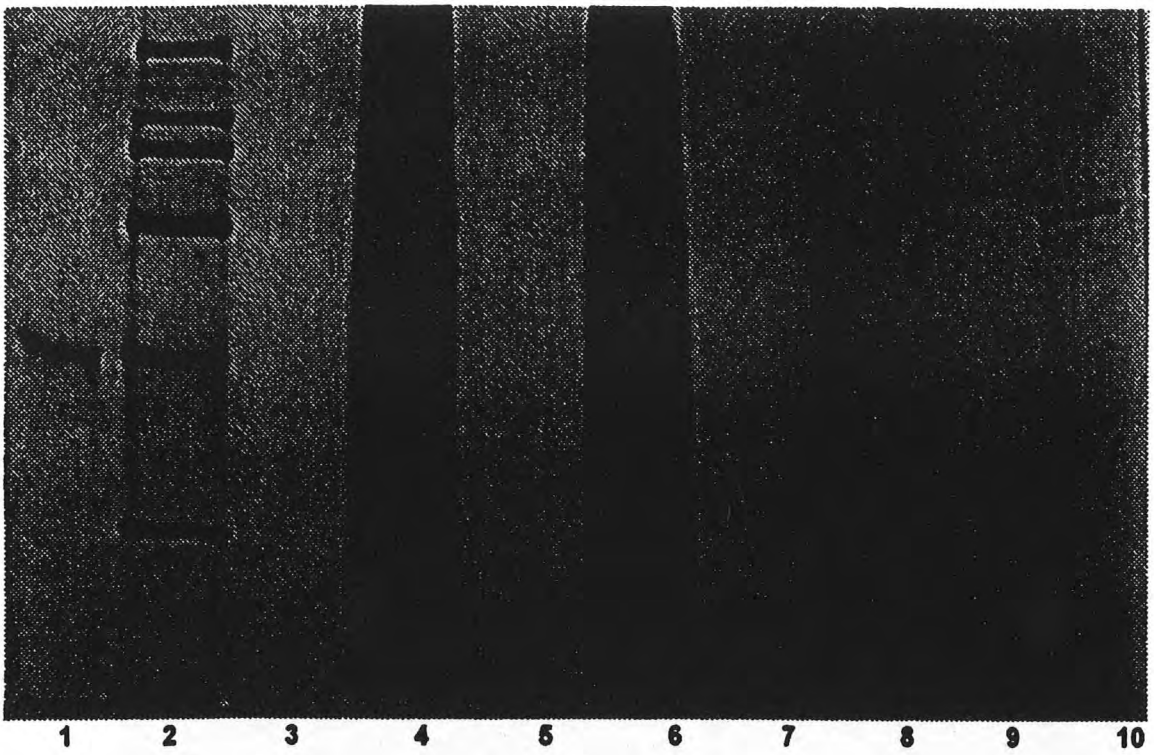


Figure 2. SDS PAGE comparison of maize-expressed Cry1Fa2 and trypsinolyzed Cry1Fa2 as obtained from bacterial origin. SDS PAGE was performed as described in Materials and Methods. Lane 1, ovalbumin (1.0ug); Lane 2, molecular weight markers; Lane 3, blank; Lane 4, maize protein extract from V-10 stage of growth from inbred line 5XH751 x TC1360 [BC2] Source #E9710039 (12.5ug); Lane 5, blank; Lane 6, maize protein extract from V-10 stage of growth from the isogenic inbred line, 5XH751 (12.5ug); Lane 7, blank; Lane 8, PHB powder from MR872 trypsinolyzed gene expression product (5.0ug); Lane 9, blank; Lane 10, powder #1 from MR872 trypsinolyzed gene expression product (5.0ug).

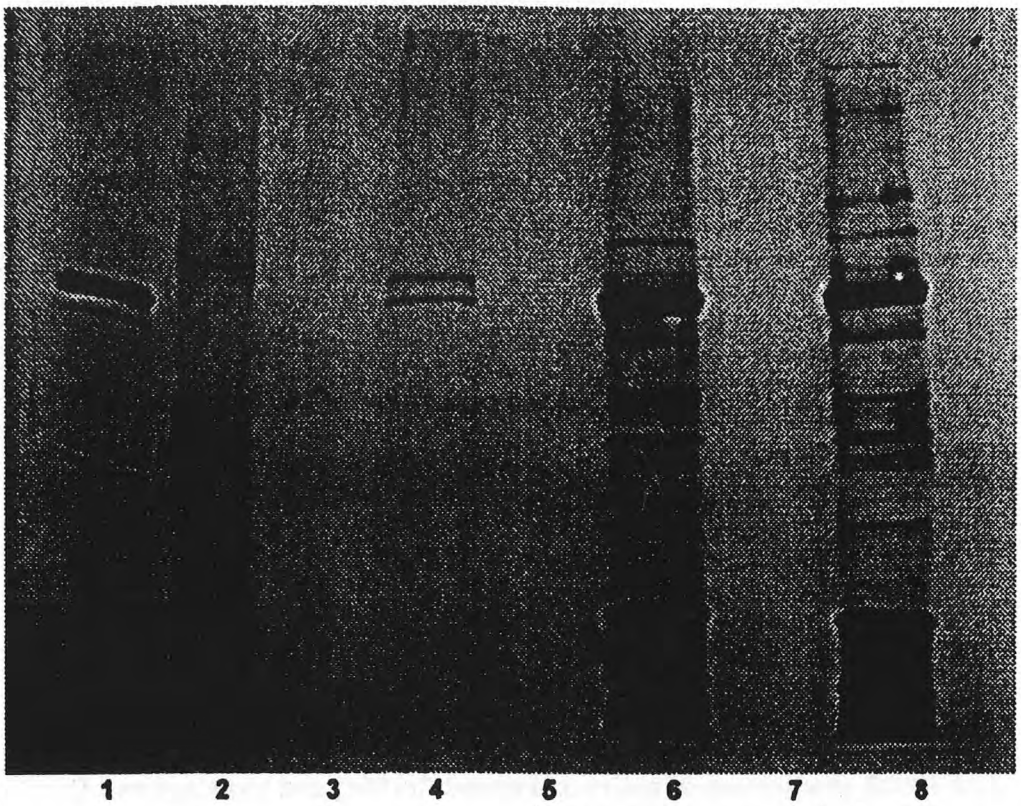


Figure 3. Comparison of the Cry1Fa2 immunoreactivity of maize-expressed Cry1Fa2 and trypsinolyzed Cry1Fa2 as obtained from bacterial origin. SDS PAGE and western blotting were performed as described in Materials and Methods. Lane 1, Cry1Fa2 standard (0.7ug); Lane 2, pre-stained molecular weight markers; Lane 3, blank; Lane 4, maize protein extract from V-10 stage of growth from inbred line 5XH751 x TC1360 [BC2] Source #E9710039 (12.5ug); Lane 5, blank; Lane 6, PHB powder from MR872 trypsinolyzed gene expression product (5.0ug); Lane 7, blank; Lane 8, powder #1 from MR872 trypsinolyzed gene expression product (5.0ug).

1 MENNIQNQCVPYNCLNNPEVEILNEERSTGRLPLDISLSLTRFLLSEFVP

51 GVGVAFGLFDLIWGFITPSDWSLFLLQIEQLIEQRIETLERNRAITTLRG

101 LADSYEIYIEALREWEANPNNAQLREDVRIRFANTDDALITAINNFTLTS

151 FEIPLLSVYVQAANLHLSLLRDAVSFGQGWGLDIATVNNHYNRLINLIHR

201 YTKHCLDTYNQGLENLRGTNTRQWARFNQFRRDLTLTVLDIVALFPNYDV

251 RTYPIQTSSQLTREIYTSSVIEDSPVSANIPNGFNRAEFGVRPPHLMDFM

301 NSLFVTAETVRSQTVWGGHLVSSRNTAGNRINFPSYGVFNPGGAIWIADE

351 DPRPFYRTLSDPVFVRGGFGNPHYVLGLRGVAFQQTGTNHTRTFRNSGTI

401 DSLDEIPPQDNSGAPWNDYSHVLNHVTFVRWPGEISGSDSWRAPMFSWTH

451 RSATPTNTIDPERITQIPLVKAHTLQSGTTVVRGPGFTGGDILRRTSGGP

501 FAYTIVNINGQLPQRYRARIRYASTTNLRIYVTVAGERIFAGQFNKTMDT

551 GDPLTFQSFSYATINTAFTFPMSQSSFTVGADTFSSGNEVYIDRFELIPV

601 TATFE

Figure 4A. Amino acid sequence of the synthetic *cry1F* gene product showing sites for potential glycosylation. The amino sequences corresponding to possible sites for N-glycosylation, asparagine (N) - X - (serine (S) or threonine (T)), where X is any amino acid, are underlined. Putative sites for O-glycosylation, serine (S) or threonine (T), are italicized.

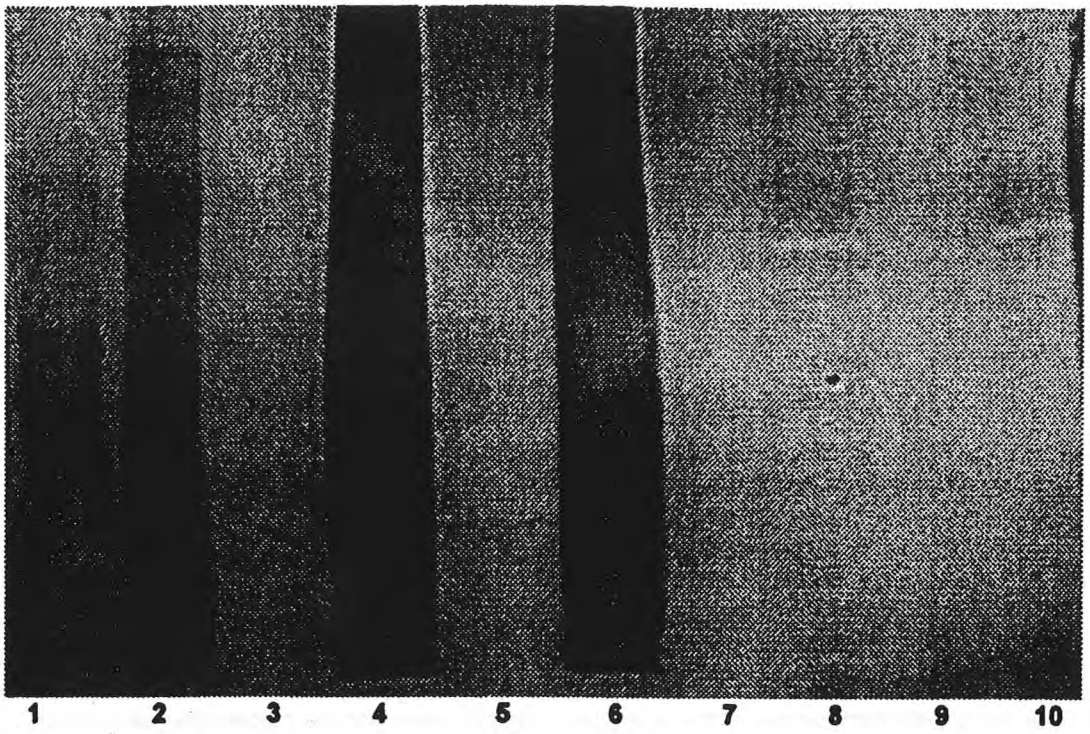


Figure 4B. Total carbohydrate detection of maize-expressed Cry1Fa2 and trypsinolyzed Cry1Fa2 as obtained from bacterial origin. SDS PAGE and blotting to nitrocellulose were performed as described in Materials and Methods. Lane 1, ovalbumin (1.0ug); Lane 2, molecular weight markers; Lane 3, blank; Lane 4, maize protein extract from V-10 stage of growth from inbred line 5XH751 x TC1360 [BC2] Source #E9710039 (12.5ug); Lane 5, blank; Lane 6, maize protein extract from V-10 stage of growth from the isogenic inbred line, 5XH751 (12.5ug); Lane 7, blank; Lane 8, PHB powder from MR872 trypsinolyzed gene expression product (5.0ug); Lane 9, blank; Lane 10, Powder #1 from MR872 trypsinolyzed gene expression product (5.0ug).

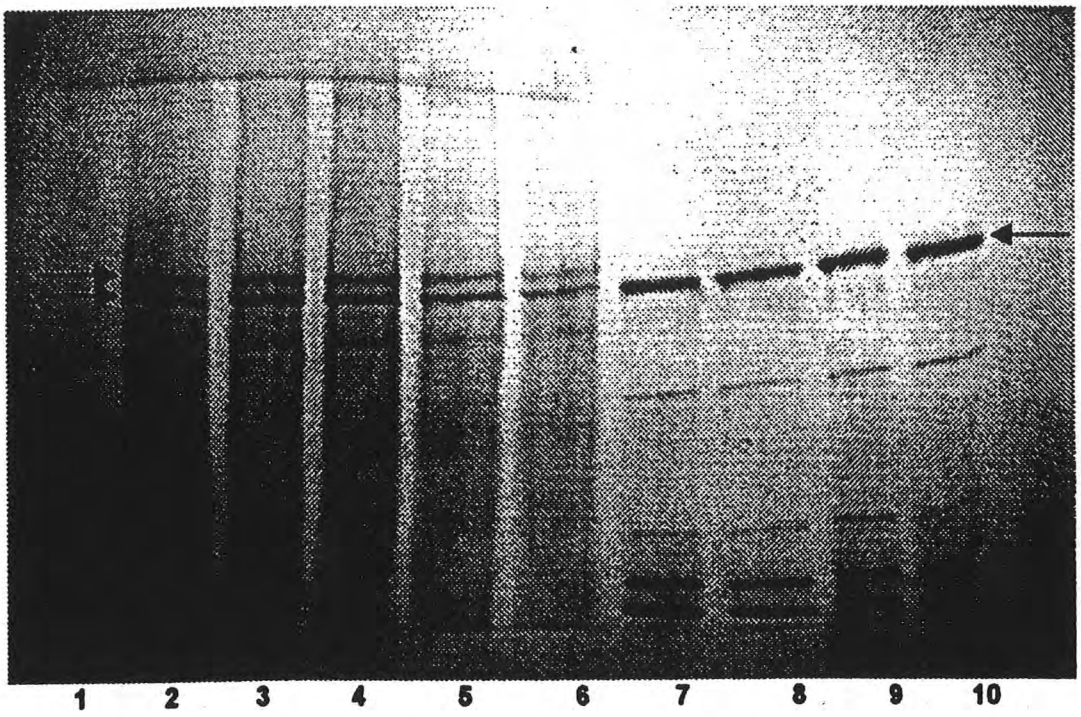


Figure 5. Comparison of the N-termini of maize-expressed Cry1Fa2 and trypsinolyzed Cry1Fa2 as obtained from bacterial origin. SDS PAGE, PVDF blotting, staining, destaining and N-terminal sequencing were performed as described in Materials and Methods. Lane 1, pre-stained molecular weight markers; Lanes 2 – 6, partially purified maize-expressed Cry1Fa2 (~0.5 ug per lane); Lanes 7 and 8, MR872 trypsinolyzed gene expression product, powder PHB (~1.0 ug per lane); Lanes 9 and 10, , MR872 trypsinolyzed gene expression product, powder 1 (~1.0 ug per lane). Arrows indicate bands that were sequenced (the two arrows on the left of the electrophoretagram correspond to the upper and lower bands of the maize-derived Cry1Fa2 referred to in the text).

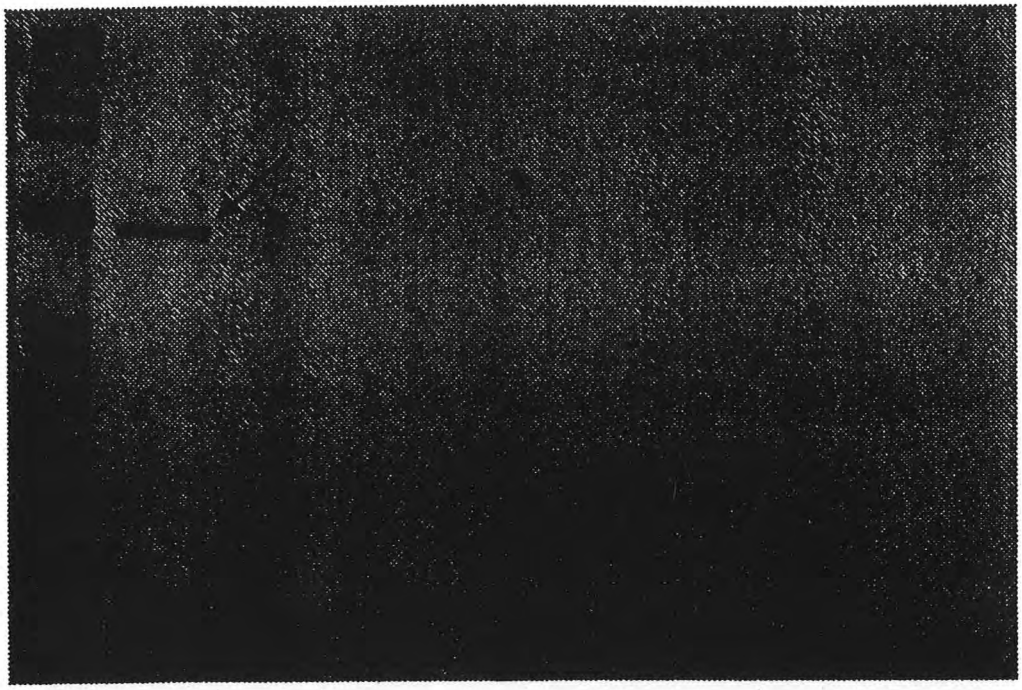


Figure 6A. Time course of the *in vitro* digestibility of trypsinolyzed Cry1Fa2 as obtained from bacterial origin. Pepsin-HCl digestion and SDS PAGE were performed as described in Materials and Methods. Lane 1, molecular weight markers; Lanes 2 – 9 each contain trypsinolyzed Cry1Fa2 (2.1ug) and pepsin (21ng) incubated for the indicated time periods, at room temperature. Lane 2, 0 min; Lane 3, 1 min; Lane 4, 2 min; Lane 5, 4 min; Lane 6, 10 min; Lane 7, 15 min; Lane 8, 30 min; Lane 9, 60 min. The single arrow indicates the electrophoretic species corresponding to Cry1Fa2.

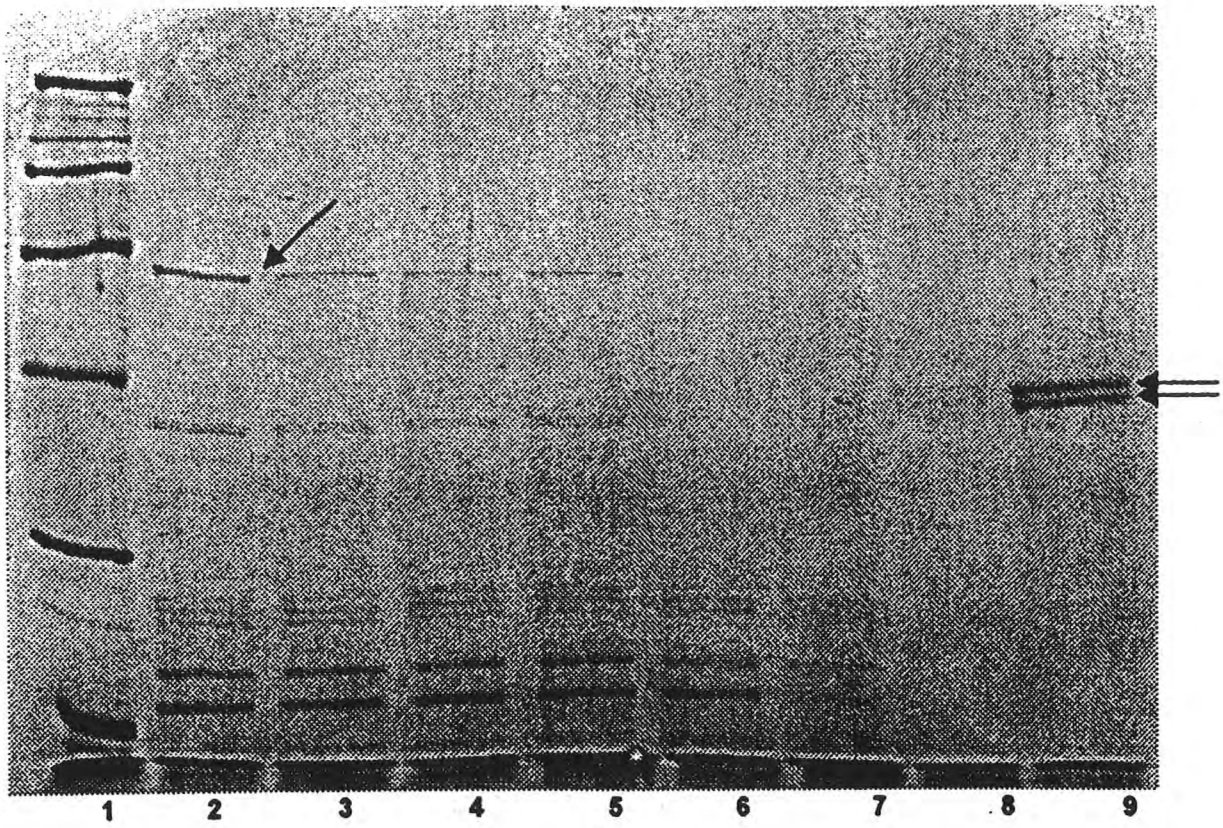


Figure 6B. Concentration dependence of the *in vitro* digestibility of trypsinolyzed Cry1Fa2 as obtained from bacterial origin. Pepsin-HCl digestion and SDS PAGE were performed as described in Materials and Methods. Lane 1, molecular weight markers; Lanes 2 – 9 each contain trypsinolyzed Cry1Fa2 (2.1ug) and the indicated amounts of pepsin, incubated for 5min at room temperature. Lane 2, no added pepsin; Lane 3, 2.1pg pepsin; Lane 4, 21pg pepsin; Lane 5, 210pg pepsin; Lane 6, 2.1ng pepsin; Lane 7, 21.0ng pepsin; Lane 8, 210ng pepsin; Lane 9, 2.1ug pepsin. The single arrow indicates the electrophoretic species corresponding to Cry1Fa2 and the double arrow (right) corresponds to pepsin A.

Signature Page

[redacted]

OCT 19 '98

Date

10-19-98

Date

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 238, No. 2, February 1963
Printed in U.S.A.

The Amino Acid Composition of Chromatographically Purified Pepsinogen*

RUTH ARNON† AND GERTRUDE E. PERLMANN

From The Rockefeller Institute, New York 21, New York

(Received for publication, September 21, 1962)

The amino acid composition of pepsinogen prepared by salt precipitation (1) has been reported by Van Vunakis and Herriott (2). The procedure for the purification of this protein, however, can be appreciably simplified with the use of chromatography on cellulose columns. Since such preparations are currently used in our laboratory for chemical, physical, and immunological studies, a need for precise data on the amino acid distribution of these preparations led to the following study. Five pepsinogen samples were analyzed. As will be shown in this report, the results obtained point to a uniformity of our preparations and, in part, agree well in their composition with those of Van Vunakis and Herriott (2), thus warranting their use for investigations to be reported (3).

EXPERIMENTAL PROCEDURE

Materials—Pepsinogen R-I and R-II were prepared in this laboratory, whereas lot Nos. 6005, 6006, and 6012 were preparations of the Worthington Biochemical Corporation, obtained by the procedure outlined below.

Methods—For amino acid analyses, samples of 10 mg of the protein were hydrolyzed in a vacuum with 2 ml of 6 N hydrochloric acid at 110 ± 1° for 22, 48, and 72 hours. The hydrolysates were evaporated in a rotating evaporator at 50° and the residue, after several washings with distilled water, was dissolved in 0.2 N citrate buffer of pH 2.2 and the volume of the solution adjusted to 5.0 ml. In view of variations in the salt content of the individual pepsinogen preparations, the concentrations of the hydrolysates were derived from nitrogen analyses by the Pregl micro-Kjeldahl method with a nitrogen factor for conversion to dry weight based on separate nitrogen, moisture, and ash determination.[1] The amino acid composition was determined according to Spackman, Stein, and Moore (4). Separate samples were oxidized with performic acid and after hydrolysis used for the estimation of cysteic acid (5).

Tryptophan analyses were carried out on unhydrolyzed pepsinogen solutions, according to the method of Goodwin and Morton (6), with correction for extraneous absorption by substituting into the equation the tyrosine content derived from the chromatographic analyses.

Electrophoresis was performed on 0.7% protein solutions in a sodium phosphate buffer of pH 6.8 and 0.1 ionic strength cell of 11-ml capacity in the apparatus described by Lo worth (7). A phosphate sodium chloride buffer of pH 7.7 0.15 ionic strength was used for ultracentrifugation.

Potential pepsin activity refers to the activity of pepsin solutions which had been activated at pH 2.0 and 37° fo minutes and were subsequently assayed with the aid of hemoglobin method at pH 2.0 (8).

Pepsinogen Purification—Pepsinogen R-I and R-II were fied from commercial crude pepsinogen (Pentex) by a modi tion of the procedure of Ryle (9). As suggested by Ryle, c matography was carried out on DEAE-cellulose columns. manipulations, however, were done at 5° and elution was formed with a gradient from 0 to 0.45 M sodium chloride in M phosphate buffer of pH 6.9 to insure a pH more alkaline t that at which activation to pepsin is initiated (2).[2] Fig. 1 sh the results of the procedure. Samples were removed at step for spectrophotometric analysis, for measurement of tein concentration (12), and assay of pepsinogen activity. active material collected from the tubes in the center of main peak (see Fig. 1) was concentrated by negative pres dialysis against 0.002 M phosphate buffer and was lyophili No further purification was achieved if rechromatographe DEAE-cellulose and Sephadex G-50 columns. The preparat obtained have a potential pepsin activity corresponding to of the Worthington crystalline pepsin. On free electropho the protein migrates as a single component with the mob of $u = -10.0 \times 10^{-5}$ volts cm^{-2} sec^{-1}. The protein is homogeneous in the ultracentrifuge with a sedimentation stant $s_{20} = 3.2 \times 10^{-13}$ to $3.3_5 \times 10^{-13}$ corresponding molecular weight of 41,000 ± 1,000 as determined by sedi tation equilibrium and assuming a partial specific volum 0.750.

RESULTS

In Table I are presented the analytical data obtained fr 22-, 48-, and 72-hour hydrolysate for the five pepsinogen pr rations. It is of interest that the composition of these pre tions is identical within the limit of error of the method. the case with most proteins, serine, threonine, and tyrosine partly decomposed during hydrolysis, and in the final cal tions a correction has been applied which was obtained extrapolation to zero time. It should be noted, however,

* This work was supported in part by Grant A-2449 of the United States Public Health Service, National Institutes of Health.

† On leave of absence from the Weizmann Institute of Science, Rehovoth, Israel.

[1] We are indebted to Mr. T. Bella for these determinations.

[2] Similar procedures have also been described by Van Vu and Leikhim (10) and by Liener (11).

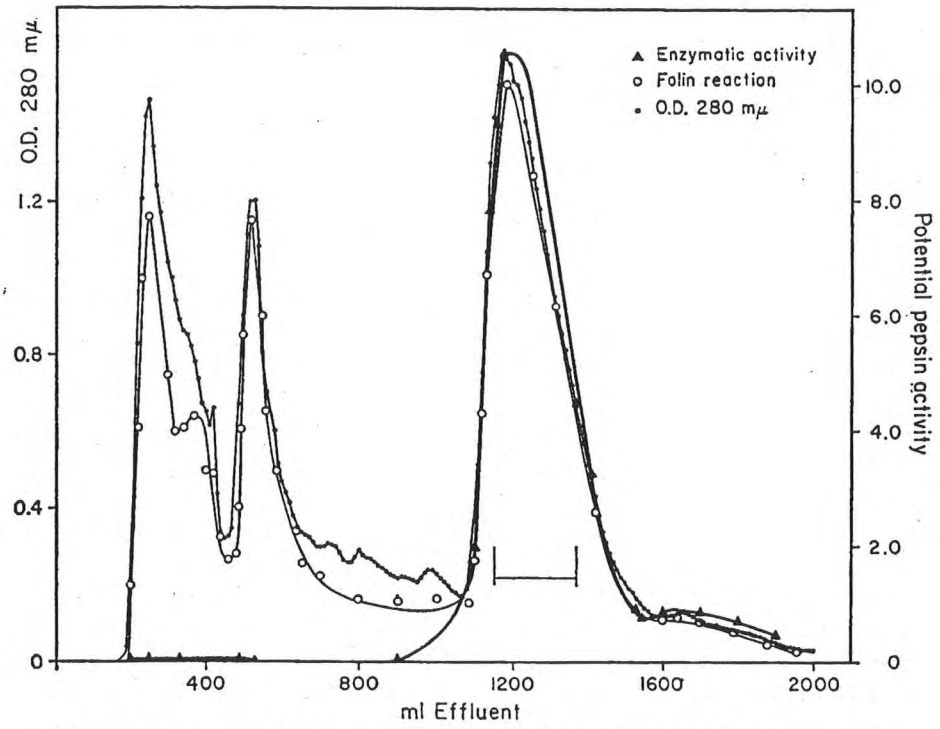


FIG. 1. Chromatographic separation of pepsinogen on DEAE-cellulose columns

TABLE I

Amino acid composition of pepsinogen as determined by chromatography of acid hydrolysates on columns of Amberlite IR-120

Amino acid	Amino acid per 100 g of protein														
	R-I			R-II			6005			6006			6012		
	22 hr	48 hr	72 hr	22 hr	48 hr	72 hr	22 hr	48 hr	72 hr	22 hr	48 hr	72 hr	22 hr	48 hr	72 hr
Lysine	3.79	3.75	3.72	3.61	3.84	3.63	3.88	3.83	3.83	3.74	3.92	3.96	3.74	3.74	3.76
Histidine	1.26	1.13	1.14	1.12	1.20	1.15	1.18	1.20	1.20	1.20	1.24	1.28	1.18	1.22	1.19
Arginine	1.86	1.81	1.81	1.78	1.93	1.80	1.93	1.80	2.02	1.98	2.03	2.11	1.87	1.90	1.86
Aspartic acid	13.72	13.74	13.28	14.55	16.15	14.77	14.95	15.15	15.20	14.22	14.94	14.75	15.0	15.35	15.72
Threonine	7.09	6.72	6.30	7.07	7.29	6.59	7.30	7.28	7.00	7.08	7.06	6.58	7.47	7.24	6.61
Serine	9.52	8.78	7.72	9.92	9.31	7.70	10.38	9.49	8.70	9.97	9.31	8.05	9.95	9.03	7.05
Glutamic acid	10.53	10.46	10.10	10.22	11.15	10.40	11.05	11.10	11.30	10.59	10.92	10.93	10.90	11.10	11.20
Proline	4.93	5.24	4.99	5.04	5.52	5.11	5.31	5.21	5.30	4.99	5.20	5.24	5.22	5.33	5.38
Glycine	6.31	6.20	6.06	6.28	6.78	6.39	6.50	6.49	6.53	6.18	6.51	6.51	6.88	6.90	6.99
Alanine	4.40	4.31	4.29	4.28	4.60	4.36	4.54	4.57	4.58	4.29	4.54	4.52	4.58	4.63	4.70
Half-cystine				1.60						1.74			1.76		
Valine	6.40	6.72	6.57	6.58	7.34	7.03	6.77	7.21	7.20	5.91	6.81	6.89	6.22	7.31	7.63
Methionine	1.48	1.38	1.39	1.45	1.42	1.47	1.55	1.71	1.57	1.50	1.31	1.56	1.56	1.54	1.61
Isoleucine	7.30	7.65	7.52	7.50	8.20	7.79	7.56	8.00	8.07	7.14	7.92	7.87	8.07	8.39	8.54
Leucine	10.18	10.10	9.88	10.20	10.85	10.28	10.60	10.70	10.78	10.40	10.56	10.56	10.80	10.94	11.2
Tyrosine	7.10	7.10	6.55	7.15	7.07	6.98	7.37	7.32	7.19	6.97	6.80	6.81	7.70	7.61	7.25
Phenylalanine	5.83	5.94	5.80	5.90	6.54	6.20	5.89	5.99	6.01	5.94	6.34	6.31	6.17	6.26	6.37
Tryptophan[a]	2.99			3.175			3.06			2.715			3.025		
Nitrogen content (percentage)	14.4			14.96			14.46								
Molar extinction coefficient ($\epsilon \times 10^3$)	51,300			52,100			50,270								

[a] Obtained from ultraviolet measurements.

TABLE II

Amino acid composition of pepsinogen[a]

Amino acid	Amino acid per 100 g of protein	Amino acid residues per 100 g of protein	N as percentage of total N	Calculated No. of residues for molecular weight 41,000	No. of residues per molecule	
					This paper	Van Vunakis and Herriott (2)
Lysine	3.78	3.22	4.59	10.65	11	12
Histidine	1.19	1.05	2.24	3.14	3	4
Arginine	1.96	1.76	4.39	4.62	4	3
Aspartic acid	14.77	12.76	10.79	45.6	46	46
Threonine[b]	8.12	6.89	6.64	27.9	28	28
Serine[b]	12.03	9.99	11.18	47.0	47	53
Glutamic acid	10.80	9.47	7.13	30.2	30	32
Proline	5.20	4.83	4.39	18.5	19	20
Glycine	6.51	4.96	8.46	36.2	36	36
Alanine	4.48	3.58	4.92	20.8	21	27
Half-cystine[c]	1.72	1.46	1.39	5.9	6	(6)
Valine[d]	6.84	6.11	6.01	24.9	25	27
Methionine[e]	1.50	1.32	0.98	4.3	*ca.* 4	(5)
Isoleucine[d]	7.83	6.91	5.96	25.2	25	64
Leucine	10.54	9.09	7.82	32.9	33	
Tyrosine[b]	7.59	6.83	4.08	17.3	17	20
Phenylalanine	6.10	5.43	3.59	15.2	15	20
Tryptophan	2.99	2.76	2.89	6.1	(6)	6
Total	97.87	97.45				

[a] Average of five analyses.
[b] The values for threonine, serine, and tyrosine were obtained by extrapolation of 22-, 48-, and 72-hour hydrolysates to zero time.
[c] Determined as cysteic acid, after performic acid oxidation.
[d] The values at 72 hours were taken for the calculations.
[e] Methionine corrected for 5% loss during chromatography.

the destruction of serine and threonine during hydrolysis of pepsinogen exceeds that usually observed with most proteins.

A summary of the analytical results is presented in Table II which lists the average value for each amino acid derived from the data of Table I. Those of half-cystine, however, were obtained in independent analyses of the oxidized protein. If these results are expressed as number of residues per molecular weight of 41,000 (Table II), they lie close to an integral value. For comparison, the amino acid composition obtained by Van Vunakis and Herriott (2) is given. Except for differences in a few nonpolar amino acids, the results agree reasonably well. Furthermore, it should be pointed out that the sum of the total amino acid residue weights is in close agreement with the value of 41,000 obtained independently by ultracentrifugation.

DISCUSSION

The results of the amino acid composition of five pepsinogen preparations purified by chromatography on DEAE-cellulose columns point to a uniformity of these preparations. The analytical results, if taken together with the physicochemical properties, therefore, warrant their use for the immunological studies to be reported (3), as well as for the optical rotatory measurements that have been presented elsewhere (13). Although traces of a faster sedimenting component corresponding to 1 to 2% caused a slight asymmetry of the ultracentrifuge patterns, there are reasons to believe that this material represents associated protein and not a contaminant.

Comparison of the amino acid composition of our preparations with that reported by Van Vunakis and Herriott (2) indicates a fair agreement, particularly if taken into account that their analysis was carried out with the aid of Dowex 50-X4 chromatography.

SUMMARY

The amino acid composition of five pepsinogen preparations, purified by chromatography on diethylaminoethyl cellulose columns, agree with each other within the experimental error of the amino acid determination. The results are compared with those of Van Vunakis and Herriott.

Acknowledgments—The authors are greatly indebted to Dr. Stanford Moore for making available the oxidation procedure for the determination of cysteic acid. Our sincere thanks go to Dr. David Yphantis for the ultracentrifuge measurements, to Dr. Samuel E. Allerton for the electrophoresis experiments, and to Miss Catherine de Watteville and Miss Jo-Ann Fine for their assistance in the amino acid analyses.

REFERENCES

1. HERRIOTT, R. M., *J. Gen. Physiol.*, **21**, 501 (1938).
2. VAN VUNAKIS, H., AND HERRIOTT, R. M., *Biochim. et Biophys. Acta*, **23**, 600 (1957).
3. ARNON, R., AND PERLMANN, G. E., *J. Biol. Chem.*, in press.
4. SPACKMAN, D. H., STEIN, W. H., AND MOORE, S., *Anal. Chem.*, **30**, 1190 (1958).

5. Moore, S., *J. Biol. Chem.*, **238**, in press.
6. Goodwin, T. L., and Morton, R. A., *Biochem. J.*, **40**, 628 (1946).
7. Longsworth, L. G., *Ind. Eng. Chem., Anal. Ed.*, **18**, 219 (1946).
8. Anson, M. L., in J. H. Northrop, M. Kunitz, and R. M. Herriott (Editors), *Crystalline enzymes*, Ed. 2, Columbia University Press, New York, 1948, p. 305.
9. Ryle, A., *Biochem. J.*, **75**, 145 (1960).
10. Van Vunakis, H., and Leikhim, E., *N. Y. State Dept. Health, Ann. Rep.*, **72** (1958).
11. Liener, I. E., *Biochim. et Biophys. Acta*, **37**, 522 (1960).
12. Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J., *J. Biol. Chem.*, **193**, 265 (1951).
13. Perlmann, G. E., and Harrington, W. F., *Biochim. et Biophys. Acta*, **54**, 606 (1961).

)0 has also shown that much of
r moved by the aminopeptidase
the aminopeptidase has been
n the *N*-terminal structure of
'h e studies indicate that leucine
ining information concerning the
ally active proteins to their hor-

otein Structure" (A. Neuberger, ed.),

A , 207 (1958).

Chapter 4

Pepsin

F. A. BOVEY AND SAM S. YANARI

Central Research Laboratories, Minnesota Mining and Manufacturing Company, St. Paul, Minnesota

TABLE OF CONTENTS

I. HISTORICAL BACKGROUND

Pepsin is the principal enzyme of gastric juice. It is distinctive among enzymes for having a very low isoelectric point and a very low pH optimum. It was perhaps the first individual enzyme to be recognized (Spallanzani, 1783) and named [Schwann, 1825 (*1*)] and the second enzyme to

1. A. L. Gillespie, "The Natural History of Digestion." London, 1898.

be crystallized (*2*). Extensive studies of pepsin in both crude and crystallized form have been carried out, and these have been summarized in earlier reviews (*3–6*). The precursor of pepsin in the stomach, pepsinogen, was discovered by Langley (*7*); its existence had been postulated earlier (*8*).

II. Occurrence

Pepsin probably occurs in the gastric juice of all vertebrates. Purified pepsins from swine (*2*), beef (*9*), sheep (*4*), and chicken (*10*) exhibit the same substrate specificity, whereas the specificity of salmon pepsin (*11*) is quite different (*12*).

Pepsins have also been purified from whale (*13*), tuna (*14*), shark (*15*), halibut (*16*), and codfish (*17*). Limited studies (*14*) indicate that the pepsins of fish may differ from those of warm-blooded animals in specificity and in being stable to higher pH values. Most studies in the literature have been carried out with swine pepsin. That swine and beef pepsins are different proteins is indicated by the fact that saturated solutions of swine pepsin will still dissolve crystalline beef pepsin (*9*).

III. Isolation, Purification, and Assay

Pepsin may be purified from gastric juice (*3*), from crude commercial preparations (*3*), or from pepsinogen obtained from gastric mucosa (*18*). The first crystalline pepsin was obtained from swine by a salt fractionation procedure (*2*) which has since been somewhat modified (*18, 19*).

2. J. H. Northrop, *J. Gen. Physiol.* **13**, 739 (1930).
3. J. H. Northrop, M. Kunitz, and R. M. Herriott, "Crystalline Enzymes," 2nd ed. Columbia Univ. Press, New York, 1948.
4. M. Bergmann and J. S. Fruton, *Advances in Enzymol.* **1**, 63 (1941).
5. E. L. Smith, "The Enzymes," 1st ed., Vol. I, Part 2, p. 793, 1951.
6. H. Neurath and N. M. Green, in "The Proteins" (H. Neurath and K. Bailey, eds.), Vol. II, Part B, p. 1057. Academic Press, New York, 1954.
7. J. N. Langley, *J. Physiol. (London)* **3**, 246 (1882).
8. M. Ebstein and P. Grutzner (1854), quoted in ref. 1.
9. J. H. Northrop, *J. Gen. Physiol.* **16**, 615 (1933).
10. R. M. Herriott, Q. M. Bartz, and J. H. Northrop, *J. Gen. Physiol.* **21**, 575 (1938).
11. E. R. Norris and D. W. Elam, *JBC* **134**, 443 (1940); *Science* **90**, 399 (1939).
12. J. S. Fruton and M. Bergmann, *JBC* **136**, 559 (1940).
13. T. Saito and Y. Ishihara, *Nippon Kogei Kagaku Kaisha* **30**, 426 (1956) [*C. A.* **51**, 18035 (1957)].
14. E. R. Norris and D. W. Elam, *JBC* **204**, 673 (1953).
15. G. P. Sprissler, Thesis, Catholic University of America, 1942.
16. N. Eriksen, Thesis, University of Washington, 1943.
17. J. Labarre, J. L. Trembley, L. Trochu, and A. P. Dussault, *Rev. can. biol.* **10**, 140 (1951).
18. R. M. Herriott, V. Desreux, and J. H. Northrop, *J. Gen. Physiol.* **24**, 213 (1940).
19. J. St. L. Philpot, *BJ* **29**, 2458 (1935).

A preparation low in salt is obtained by crystallization from 20% alcohol (*20*). Commercial crystalline pepsins often contain significant nonprotein contaminants, probably peptides arising from autolysis (*21*). Heirwegh and Edman (*22*) removed peptide contaminants from pepsin preparations by passing a 0.5% solution of pepsin at pH 2.9 through a Dowex 50 column at low temperature. Mitz and Schleuter (*23*) have employed absorption on the weak anion exchanger DEAE-cellulose (*24*) to obtain a stable, dry product.

The most widely used assay for peptic activity employs acid-denatured hemoglobin as substrate (*25*), followed by measurement of material soluble in 5% trichloroacetic acid (TCA) by means of either optical density at 280 mμ (*3*) or the Folin phenol reagent (*25*). The clotting of milk has been employed as an assay (*26*). Proteolysis by pepsin at pH 3.0 can be followed with a sensitive pH-stat in a manner analogous to the assay of chymotrypsin by Ottesen (*27*). Since the pH of the assay is below the pK_a' of the α-carboxyl groups of the released polypetides (*28*), a net uptake of protons can be observed [Eq. (1)]. This method is less sensitive than the

$$\text{R—NHCO—R}' \xrightarrow[\text{H}^+]{\text{H}_2\text{O}} \text{RNH}_3^+ + \text{R}'\text{COOH} \qquad (1)$$

TCA-soluble method (*25*) but is convenient for the rapid assay of a limited number of samples. Berger *et al.* (*29*) have devised a sensitive, indirect assay for pepsin based on its rate of inactivation of ribonuclease. Synthetic substrates are in many ways preferable to proteins as substrates for the study of proteolytic enzymes, but highly susceptible synthetic substrates for pepsin have been discovered only relatively recently (*30*).

IV. Physical and Chemical Properties of Native Pepsin

A. Molecular Weight and Shape

In Table I are given values of the ultracentrifugal sedimentation constant, $s_{20,w}$, and the diffusion coefficient, $D_{20,w}$, (the subscripts referring to water solutions at 20°) reported for native pepsin by various authors,

20. J. H. Northrop, *J. Gen. Physiol.* **30**, 177 (1946).
21. J. Steinhardt, *JBC* **129**, 135 (1939).
22. K. Heirwegh and P. Edman, *BBA* **24**, 219 (1957).
23. M. A. Mitz and R. J. Schleuter, *JACS* **81**, 4024 (1959).
24. E. A. Peterson and H. A. Sober, *JACS* **78**, 751 (1956).
25. M. L. Anson, *J. Gen. Physiol.* **22**, 79 (1938).
26. R. M. Herriott, *J. Gen. Physiol.* **21**, 501 (1938).
27. M. Ottesen, *ABB* **65**, 70 (1956).
28. J. T. Edsall and J. Wyman, "Biophysical Chemistry," Vol. I. Academic Press, New York, 1958.
29. A. Berger, H. Neumann, and M. Sela, *BBA* **33**, 249 (1959).
30. L. E. Baker, *JBC* **193**, 809 (1951).

TABLE I

SEDIMENTATION-DIFFUSION RESULTS FOR NATIVE PEPSIN

$s_{20,w} \times 10^{13}$	Ref.	$D_{20,w} \times 10^{7}$	Ref.	$V_{20,w}$ cc.-g.$^{-1}$	Ref.	M	f/f_0	a/b[a]
3.3	(32)	9.0	(33)	0.75	(31)	35,500	1.08	2.50
3.0	(34)	9.69	(2)	(est.)		33,400[b]	1.13	3.25
3.20	(35)	8.71	(35)			35,700	1.11	3.00
3.16	(36)	11.2	(36)			27,000	—	—

[a] Assuming an unhydrated, prolate ellipsoid.
[b] Employing 8.71×10^{7} for $D_{20,w}$.

together with molar volume $V_{20,w}$, and values of the molecular weight, M, calculated from them; also shown are the frictional ratio, f/f_0, and the axial ratio, a/b (31). Crystalline pepsin usually appears to be quite homogeneous in the ultracentrifuge. The molecular weight appears to be about 35,000, in satisfactory agreement with the results of other methods (see below). If we assume a degree of hydration of 0.25 g./g. of protein (roughly a monomolecular layer), then, from the Oncley sedimentation-constant diagram (37), a/b is found to be nearly unity, indicating a nearly spherical form. Most globular proteins have higher f/f_0 ratios, but insulin, ribonuclease, and cytochrome c are comparable to pepsin (31).

Polson (38) found the reduced specific viscosity, η_{sp}/c, to be 0.042 dl-g.$^{-1}$ for a 3% pepsin solution at pH 4.6. Jirgensons (39) reported 0.052 dl-g.$^{-1}$ for η_{sp}/c for 1% solutions at pH 4.6. Edelhoch (35) found the intrinsic viscosity $[\eta]$ to be 0.031 to 0.033 dl-g.$^{-1}$ in cacodylate buffers of ionic strength 0.01 to 0.100; at still lower ionic strength, higher values were obtained owing to the electroviscous effect. Bovey and Yanari (40) observed a value of 0.031 in 0.02 M KCl at pH 4.0. For impenetrable spheres of any size and having a specific volume of 0.75, $[\eta]$ is 0.019 dl-g.$^{-1}$; for an unhydrated prolate ellipsoid, a value of 0.033 dl-g.$^{-1}$ corresponds to an axial ratio of 3.9 (41); for a degree of hydration of 0.25 g./g. of protein, the

31. J. T. Edsall, in "The Proteins" (H. Neurath and K. Bailey, eds.), Vol. I, Part B, Chapter 7. Academic Press, New York, 1953.
32. J. St. L. Philpot and I.-B. Eriksson-Quensel, *Nature* **132**, 932 (1933).
33. A. G. Polson, Thesis, University of Stellenbosch, 1937; quoted in ref. 31, p. 640.
34. J. Steinhardt, *JBC* **123**, 543 (1938).
35. H. Edelhoch, *JACS* **79**, 6100 (1957).
36. H. A. Dieu, *Bull. soc. chim. Belges* **65**, 603 (1956).
37. J. L. Oncley, *Ann. N. Y. Acad. Sci.* **41**, 121 (1941); see ref. 31, p. 648.
38. A. Polson, *Kolloid-Z.* **88**, 51 (1939).
39. B. Jirgensons, *ABB* **39**, 261 (1952).
40. F. A. Bovey and S. S. Yanari, unpublished results.
41. J. W. Mehl, J. L. Oncley and R. Simha, *Science* **92**, 132 (1940).

Oncley viscosity diagram (42) gives an axial ratio of 2.5, which is again lower than that calculated in this way for most globular proteins.

By measurement of light scattering, Yasnoff and Bull (43) calculated a molecular weight of 37,600 at pH 3.0. Kronman and Stern (44) found a value of 34,800 ± 870 at pH 4.5 and 5.0. Dieu (36) reported 35,000. Edelhoch (35) observed an apparent decrease of molecular weight between pH 6.0 and 7.0 (see Section IX) but reported only an approximate value of 33,000 to 37,000 at lower pH.

Dieu and Bull (45) calculated 34,400 from measurement of film pressures of pepsin monolayers. Northrop (46) obtained 35,000 by osmotic pressure measurements. Bernal and Crowfoot (47) calculated a value of 40,000 from X-ray diffraction. Probably the most accurate estimate is based on the amino acid analysis. With certain assumptions discussed in Section VI, this gives a value of 36,422, somewhat higher than the average of the results of physical measurements, but within the probable experimental error of these methods.

There appears to be no evidence for the existence of native pepsin in polymerized form.

B. Electrophoresis; Isoelectric Point

When examined by electrophoresis, native pepsin preparations often appear somewhat heterogeneous, in contrast to the apparent homogeneity shown in sedimentation. Tiselius *et al.* (48) observed a sharp band of active protein followed by a slower, more diffuse boundary due to inactive components; when the latter were removed, the specific activity was increased by 31 to 69% in various preparations. The purified protein migrated as an anion from pH 4.6 down to pH 1. Herriott *et al.* (49) found purified pepsin to be an anion in 0.1 N HCl; on standing at 30°, autolysis modified it so that it became less negatively charged. Hoch (50) reported that even "pure" pepsin exhibited heterogeneity on prolonged electrophoresis at pH 3.9 and 5.9, but this heterogeneity probably arose mainly from autolysis during electrophoresis, for at pH 8, where pepsin is inactive, the published pattern shows no detectable heterogeneity in the main peak

42. Ref. 38; see ref. 31, p. 691.
43. D. S. Yasnoff and H. B. Bull, *JBC* **200**, 619 (1953).
44. M. J. Kroman and M. D. Stern, *J. Phys. Chem.* **59**, 969 (1955).
45. H. A. Dieu and H. B. Bull, *JACS* **71**, 450 (1949).
46. J. H. Northrop, *J. Gen. Physiol.* **13**, 767 (1930).
47. J. D. Bernal and D. Crowfoot, *Nature* **133**, 794 (1934).
48. A. Tiselius, G. E. Henschen, and H. Svensson, *BJ* **32**, 1814 (1938).
49. R. M. Herriott, V. Desreux, and J. H. Northrop, *J. Gen. Physiol.* **23**, 439 (1940).
50. H. Hoch, *Nature* **165**, 278 (1950).

(Section VII.B). Edelhoch (*55*) reported that the native pepsin employed in his studies migrated as a single peak at all pH values.

From the data of Perlmann (*51*) it can be concluded that the isoelectric point is only slightly less than pH 1. The isoionic point, calculated from the amino acid composition (Section VI.A) on the assumption that all carboxyl groups have a pK_a' of 4, is 3.1. The considerable discrepancy is probably due largely to adsorbed anions, but it is possible that there may be one or two abnormally strong carboxyl groups (Section VIII.D).

V. The Formation of Pepsin from Pepsinogen

Pepsinogen was first obtained in crystalline form by Herriott and Northrop (*26, 52*). Osmotic pressure measurements indicate a molecular weight of 42,000 ± 3000 (*26, 53*). It is considerably less acidic than pepsin, having an isoelectric point of about 3.7 (*26*). Unlike pepsin, it is stable in dilute alkali, pH 7 to 9, and use of this may be made to determine pepsinogen in the presence of pepsin (*54*). It is destroyed on 24-hour standing in 50% aqueous ethanol (*55*) and is stable for only short periods below pH 6 (*26, 53*), under both of which conditions pepsin is stable.

The formation of pepsin from pepsinogen is catalyzed by hydrogen ion and also by pepsin itself. The reaction, therefore, proceeds autocatalytically below pH 5. It has a maximum at about pH 2 and is first-order with respect to both pepsin and pepsinogen (*56*). During the reaction, a molecule of *pepsin inhibitor* is split off; it contains twenty-nine amino acids and has a molecular weight of 3242 (*57*). It has been obtained in crystalline form (*58*). In addition, five smaller peptides are formed, having an aggregate molecular weight of about 4000 (*57, 59*). The inhibitor remains bound to the enzyme above pH 5 (*54, 58*), and it is for this reason that the process is autocatalytic only below this pH. During the course of complete conversion of pepsinogen, nine peptide bonds are broken (*56*). Six are accounted for by production of the inhibitor and peptides. The remainder arise from attack of pepsin on the inhibitor in later stages of the reaction, a process having a maximum rate at pH 3.5. The order in which the first six bonds are broken is unknown. From amino acid sequence studies, it appears (Sections VI.A, VI.B, Fig. 1) that pepsin forms the *C*-terminal

51. G. E. Perlmann, *Advances in Protein Chem.* **10**, 23 (1955).
52. R. M. Herriott and J. H. Northrop, *Science* **83**, 469 (1936).
53. Ref. 3, p. 81.
54. R. Ege and P. Menck-Thygesen, *Biochem. Z.* **264**, 13 (1933).
55. H. Tauber and I. S. Kleiner, *Z. physiol. Chem.* **220**, 205 (1933).
56. R. M. Herriott, *J. Gen. Physiol.* **22**, 65 (1938).
57. H. Van Vunakis and R. M. Herriott, *BBA* **22**, 537 (1956).
58. R. M. Herriott, *J. Gen. Physiol.* **24**, 325 (1941).
59. H. Van Vunakis and R. M. Herriott, *BBA* **23**, 600 (1957).

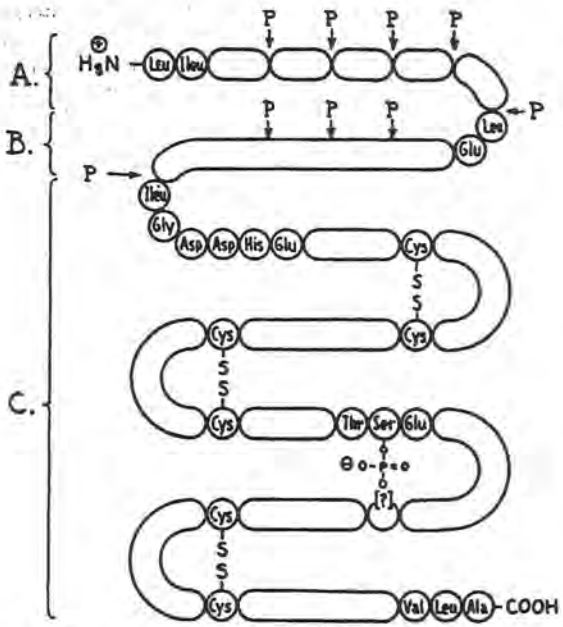


FIG. 1. The structure of pepsinogen. The principal attack by pepsin is at points marked *P*, releasing miscellaneous peptides (*A*), pepsin inhibitor (*B*), and pepsin (*C*). The undetermined sequences are actually much larger in relation to the known sequences than is indicated. The location of the phosphoserine residue with respect to the disulfides is also uncertain.

segment of the pepsinogen molecule, which is a single chain; that the inhibitor is in the middle; and that the five peptides constitute the *N*-terminal part. The breakage of one critical bond is therefore sufficient to detach the free pepsin molecule. Most of the carboxyl groups, but only a minority of the basic groups, are situated at the pepsin end of the molecule.

VI. Structure of Pepsin and Pepsinogen

A. Amino Acid Composition

The amino acid composition of pepsin was first determined by Brand by microbiological assay (*60*). Pepsin (*61*), pepsinogen (*59*), and pepsin inhibitor (*57*) have been analyzed by column chromatography. In Table II, the compositions of these three proteins are compared. The pepsin analysis has been corrected to show six half-cystines (Section VI.C) instead of the reported four (*61*). The pepsinogen analysis shows fewer residues of aspartic acid, glycine, threonine, and tyrosine than the total of pepsin

60. E. Brand, quoted in ref. *3*, p. 26.
61. O. O. Blumenfeld and G. E. Perlmann, *J. Gen. Physiol.* **42**, 553 (1959).

TABLE II

AMINO ACID COMPOSITION OF PEPSINOGEN, PEPSIN, AND PEPSIN INHIBITOR

Amino Acid	Pepsinogen (59)	Pepsin (61)	Inhibitor (57)
Aspartic acid	46	44	4
Glutamic acid	32	27	2
Glycine	36	38	1
Alanine	27	18	2
Valine	27	21	2
Isoleucine		27	
Leucine	64	28	5
Serine	53	44	2
Threonine	25	28	1
Half-cystine	(6)	6	0
Methionine	5	5	
Proline	20	15	3
Hydroxyproline	—	0.1	—
Phenylalanine	20	14	1
Tyrosine	16	18	1
Tryptophan	(6)	6	
Histidine	4	1	
Lysine	12	1	4
Arginine	3	2	1
Amide NH_3	39	36	
Phosphate	1	1	
Total amino acid residues/mole	402	343	29
Sum of amino acid residue weights	42,926	36,422	3242

plus inhibitor and, therefore, is probably too low for these residues at least. The composition of the miscellaneous peptides is correspondingly doubtful. They appear to average about 800 in molecular weight.

Pepsin is unusual in having seventy-one dicarboxylic amino acid residues (thirty-six free carboxyls) and only four basic residues in a total of 343. The proline and serine content are exceptionally high. If the fifteen proline residues were distributed statistically along the chain, only relatively short helical segments could exist, as Perlmann has pointed out (*62*). This question is discussed in Section VI.E.

B. Number of Chains; N- and C-Terminal Sequences

End-group analysis show that both pepsin and pepsinogen consist of only one chain. The *N*-terminal residues of pepsinogen, pepsin, and pepsin

62. G. E. Perlmann, *Proc. Natl. Acad. Sci. U. S.* **45**, 915 (1959).

inhibitor, determined by reaction with fluorodinitrobenzene (*63*), are found to be leucine, isoleucine, and leucine, respectively (*59*). That the *N*-terminal residue of pepsin is isoleucine has been confirmed by Heirwegh and Edman (*22*). The *N*-terminal sequences, determined by digesting the dinitrophenyl (DNP) derivatives for short periods with 6 *N* HCl and identifying the DNP amino acids and DNP peptides (*59*), are different for all three proteins, and so neither pepsin nor inhibitor can occupy the *N*-terminal position. Williamson and Passman (*64*) determined the *N*-terminal hexapeptide sequence of pepsin, using the Edman phenylisothiocyanate degradation (*65*). They reported the *N*-terminal residue to be leucine instead of isoleucine.

The *C*-terminal sequence of pepsin has been determined by Van Vunakis and Herriott (*66*) and by Williamson and Passman (*67*), using carboxypeptidase A. For both molecules, the sequence is Val·Leu·Ala; pepsin, therefore, constitutes the *C*-terminal part of the pepsinogen molecule.

The sequence in the neighborhood of the phosphoserine residue of pepsin has been determined by Flavin (*68*) by chromatography of the phosphopeptides resulting from acid hydrolysis. The most probable sequence is Thr·Ser-phos·Glu, and there was no evidence for phosphoseryl peptides containing any other sequence (see also Section VI.D).

The amino acid sequence results are summarized in Fig. 1.

C. Disulfide Bonds

Brand (*60*) reported 1.7 cystine and 0.5 cysteine residues in pepsin. Blumenfeld and Perlmann (*61*) reported four half-cystine residues and no cysteine. Kern (*69, 70*), using monothioglycol, $HOCH_2CH_2SH$, as reducing agent, found that two disulfide bonds can be readily reduced in both pepsin and pepsinogen and that there is a third which is very difficult to reduce. No free sulfhydryl groups were found. After reduction of one disulfide bond, enzymic activity was hardly affected, but, as reduction continued, activity decreased. After 2 hours, reaction had leveled off with two bonds reduced, and activity was nearly gone. The molecule is apparently initially active with two bonds reduced but is unstable and undergoes fairly

63. F. Sanger, *BJ* **39**, 507 (1945).
64. M. B. Williamson and J. M. Passman, *JBC* **199**, 121 (1952); **222**, 151 (1956).
65. P. Edman, *Acta Chem. Scand.* **4**, 283 (1950).
66. H. Van Vunakis and R. M. Herriott, quoted in C. B. Anfinsen and R. R. Redfield, *Advances in Protein Chem.* **11**, 1 (1956).
67. M. B. Williamson and J. M. Passman, *BBA* **15**, 246 (1954).
68. M. Flavin, *JBC* **210**, 771 (1954).
69. H. L. Kern, Ph.D. Thesis, Johns Hopkins University, 1953.
70. See also R. M. Herriott, *in* "The Mechanism of Enzyme Action" (W. D. McElroy and B. Glass, eds.), p. 37 *et seq.* Johns Hopkins, Baltimore, 1954.

rapid secondary changes with loss of activity. Reduction of two disulfide bonds with monothioglycol causes no decrease in molecular weight and produces no TCA-soluble peptides. These bonds thus serve to maintain loops in a single chain, as shown in Fig. 1. Kern reported an increase in intrinsic viscosity from 0.057 dl-g.$^{-1}$ (but see lower values reported in Section IV.A) to 1.0 dl-g.$^{-1}$ when native pepsin is reduced, whereas reduction of alkali-denatured pepsin (seen Section IX.B) resulted in an increase from 0.13 to approximately 0.23 dl-g.$^{-1}$, a result not readily explainable.

Bovey and Yanari (*71*) studied the reduction of pepsin by sodium sulfite at pH 6 to 7 in Tris buffer (*72*). Both native and alkali-denatured pepsin have one disulfide bond which can be reduced by 0.045 *M* sulfite; two can be reduced in denatured pepsin with 0.20 *M* sulfite, or with 0.045 *M* sulfite if four peptide bonds are previously broken by trypsin; three bonds can be reduced by 0.045 *M* sulfite if twenty-two peptide bonds are previously broken by chymotrypsin. The environment and degree of "masking" of these three disulfide bonds thus appear to differ very markedly. Accompanying the reduction of one disulfide bond by sulfite at pH 6.7 and 0°, there was an increase in intrinsic viscosity from 0.032 to 0.063 dl-g.$^{-1}$ and in $[\alpha]_D^{25°}$ from −66° to −75°; there was no loss of enzymic activity, confirming the observation of Kern (*69*). In the presence of 8 *M* urea, three disulfide bonds are reduced by sodium sulfite (approximately 0.05 *M*) (*73*) and there is extensive unfolding (*71*) of the protein which is not reversed on dialyzing away the sulfite to re-form the bonds (*71*).

D. Phosphate Binding

Perlmann (*51*) has reported that the phosphate group of pepsin (Section VI.A) is not removed by prostate phosphatase, which was believed to be specific for the monophosphate linkage, but is completely removed by potato phosphatase, which is thought to possess both phosphodiesterase and phosphomonoesterase activity. Removal of phosphate does not affect enzymic activity but raises the isoelectric point from less than 1 to 1.7. Comparison of the pH-mobility curves of the untreated and dephosphorylated pepsins, coupled with observation of the nonactivity of prostate monophosphatase, led Perlmann to conclude that the phosphorus is bound as a diester linking two portions of the chain. Similar conclusions were drawn concern-

71. F. A. Bovey and S. S. Yanari, *Abstr. Am. Chem. Soc. 134th Meeting*, 33C. Chicago (September, 1958).
72. I. M. Kolthoff and W. Stricks, *JACS* **72**, 1952 (1950); R. Cecil, *BJ* **60**, 496 (1955); **64**, 22 (1956); **68**, 18P (1957); I. M. Kolthoff, A. Anastasi, and B. H. Tan, *JACS* **80**, 3235 (1958).
73. J. R. Carter, *JBC* **234**, 1705 (1959).

ing α-and β-casein (*51*, *74*, *75*). However, the specificity of prostatic monophosphatase is not well established (*76*). Enzymic and chemical studies of other workers indicate that the phosphate in α- and β-casein is in a monoester linkage (*77*, *78*). In addition, we have seen (Section VI.B) that only one phosphorus-containing amino acid sequence can be isolated from the hydrolysis of pepsin. The nature of the binding of the phosphate in pepsin must therefore be regarded as still somewhat uncertain.

E. Conformation

Viscosity and sedimentation measurements indicate that the pepsin molecule is compact and probably not far from spherical (Section IV.A). The optical rotation—$[\alpha]_D^{20°} = -64.5°$ (*39*); $[\alpha]_D^{25°} = -65.0°$ (*79*); $[\alpha]_{600\,m\mu}^{25°} = -63.3°$ (*62*)—and rotatory dispersion constant—$\lambda_c = 216$ mμ (*39*, *62*)—suggest that there is relatively little α-helical content, a conclusion which is strengthened by the observation that the optical rotation is unchanged in concentrated urea solution (*39*, *62*, *80*) and that enzymic activity is retained (*34*, *81*). Apparently, hydrogen bonding between peptide carboxyl and —NH— groups, normally considered necessary to the stability of α-helices, is not so important in maintaining the conformation of pepsin. Guanidine hydrochloride (5 *M*) does cause some unfolding, however (*39*). Little change in tyrosyl ultraviolet absorption is observed when pepsin is acidified from near neutrality to a pH less than 1, so tyrosyl-carboxylate ion-dipole bonds (*82–84*) likewise do not appear to be important. Hydrophobic bonding (*85*) may be the chief means by which native pepsin maintains its integrity (*62*, *86*); 63% of the amino acid residues are of the nonpolar type and, therefore, could contribute to such bonding (*62*).

The optical rotation of pepsinogen, $[\alpha]_D^{23°} = -62°$, is slightly less negative than that of pepsin (*26*).

74. G. E. Perlmann, *BBA* **13**, 452 (1954).
75. G. E. Perlmann, *Nature* **174**, 273 (1954).
76. G. Schmidt, "Methods in Enzymology," Vol. II, p. 523, 1955.
77. T. Hofman, *BJ* **69**, 139 (1958).
78. E. B. Kalan and M. Telka, *ABB* **79**, 275 (1959).
79. S. S. Yanari and F. A. Bovey, unpublished observations.
80. W. Kauzmann and R. B. Simpson, *JACS* **75**, 5154 (1953).
81. G. E. Perlmann, *ABB* **65**, 210 (1956).
82. J. L. Crammer and A. Neuberger, *BJ* **37**, 302 (1943).
83. H. A. Scheraga, *BBA* **23**, 196 (1957).
84. D. B. Wetlaufer, J. T. Edsall, and B. R. Hollingworth, *JBC* **233**, 1421 (1958).
85. W. Kauzmann, in "The Mechanism of Enzyme Action" (W. D. McElroy and B. Glass, eds.), p. 70. Johns Hopkins, Baltimore, 1954.
86. S. S. Yanari and F. A. Bovey, *Federation Proc.* **18**, 356 (1959).

TABLE III

SYNTHETIC SUBSTRATES FOR PEPSIN

	Substrate[a]	Rate	Ref.
1	Ac-Phe—Phe	*****	(30)
2	Cbz-Phe—Phe amide	0	(30)
3	Cbz-Tyr—Phe	*****	(30)
4	Cbz-O-AcTyr—Phe	0	(30)
5	Cbz-Glu—Phe	***	(87)
6	Cbz-Glu—D-Phe	*	(87)
7	Cbz-Glu(NH_2)—Phe	**	(87)
8	Ac-Phe—Tyr	******	(30)
9	Ac-D-Phe—Tyr	0	(30)
10	Ac-Phe—diiodotyr	********	(30)
11	Ac-D-Phe—diiodotyr	0	(30)
12	Cbz-Tyr—Tyr	**	(87)
13	Ac-Tyr—Tyr	*****	(30)
14	Ac-D-Tyr—Tyr	0	(30)
15	Ac-dehydrotyr—Tyr	0	(30)
16	Cbz-Glu—Tyr	***	(87, 91)
17	Cbz-D-Glu—Tyr	0	(87)
18	Cbz-Glu—Tyr Et ester	***	(81)
19	Cbz-Glu—Tyr amide	**	(91)
20	Cbz-Glu—Tyr-Gly amide	**	(4)
21	Cbz-Glu—Tyr-Gly	**	(87)
22	Cbz-Glu(NH_2)—Tyr amide	*	(87)
23	Cbz-Glu—Diiodotyr	0	(87)
24	Cbz-Gly-Glu—Tyr	***	(87)
25	Gly-Glu—Tyr	**	(87)
26	Glu—Tyr	*	(87)
27	Glu—Tyr-Gly amide	0	(87)
28	Tyr-Lys-Glu—Tyr	***	(89)
29	Cbz-Gly—Tyr	**	(87)
30	Cbz-Gly—Tyr-Gly amide	*	(4)
31	Cbz-Met—Tyr	***	(90)
32	Met—Tyr	**	(90)
33	Bz—Tyr	*	(87)
34	Chloroacetyl—Tyr	0	(87)
35	Cbz-Cys—Tyr	***	(88)
36	Cys—Tyr	**	(88)
37	Cbz-S-BzCys—Tyr	**	(88)
38	Cbz-Cystinyl—Tyr	**	(88)
39	Cystinyl—Tyr	*	(88)
40	Cbz-Tyr—Cys	**	(88)
41	Tyr—Cys	**	(88)
42	Cbz-Tyr—Cystine	*	(88)
43	Tyr—Cystine	*	(88)
44	Cbz-Phe—Glu	**	(87)
45	Cbz-Glu—Glu	*	(87)
46	Cbz-Glu—Gly	*	(87)
47	Cbz-Gly-Glu—Gly amide	0	(4)

[a] All residues are of the L-configuration unless otherwise specified.

VII. Enzymic Activity

A. Synthetic Substrates

The hydrolysis of synthetic substrates by pepsin has received some study, but the information obtained is for the most part qualitative. The limited number of kinetic studies will be considered in Section VIII.B. In Table III are shown those peptides on which the action of pepsin has been tested. The following conclusions as to specificity appear to be justified from this information:

1. Pepsin hydrolyzes only peptide linkages; it is not an esterase and will not attack amide linkages.
2. Amino acid residues must be of the L-configuration on both sides of the peptide bond; 5 versus 6, 8 versus 9, 10 versus 11, 13 versus 14, and 16 versus 17. Glycyl peptides show a low susceptibility; 29, 46, 47.
3. Acylation of the α-amino group increases the susceptibility of dipeptides; 25 versus 26, 31 versus 32, and 35 versus 36.
4. Aromatic residues greatly favor attack, especially if present on both sides of the susceptible bond; phenylalanyl is about as good as tyrosyl; 1, 3, 8, 10, 12, and 13. The effect of iodination of the tyrosyl residue requires further clarification. Iodination is reported to enhance susceptibility in one case, 10, and greatly decrease susceptibility in another, 23. *O*-acetylation of the tyrosyl residue, 4, or substitution of dehydrotyrosine, 15, abolishes activity.
5. Peptides of cysteine and cystine are hydrolyzed at moderate rates if coupled to aromatic amino acid residues, 35–43.
6. Glutamyl peptides are hydrolyzed if coupled to aromatic residues, which preferably should be *C*-terminal; 5, 16–27, 44–47. Glutaminyl peptides are less susceptible than the corresponding glutamyl peptides.

The relative rates of hydrolysis cannot be interpreted literally, since many of the values were determined under conditions which were not comparable. Differences in solubility in certain cases probably account for unexpected differences in hydrolytic susceptibilities, e.g., compare 1 and 2 versus 16 and 19.

B. Proteolysis and Autolysis; Active Fragments

The specificity rules deducible from the rather limited number of types

87. J. S. Fruton, M. Bergmann, and W. P. Anslow, Jr., *JBC* **127**, 627 (1939).
88. C. R. Harrington and R. V. Pitt-Rivers, *BJ* **38**, 417 (1944).
89. A. A. Plentl and I. H. Page, *JBC* **163**, 49 (1946).
90. C. A. Dekker, S. P. Taylor, and J. S. Fruton, *JBC* **180**, 155 (1949).
91. E. J. Casey and K. J. Laidler, *JACS* **72**, 2159 (1950).

of synthetic substrate discussed in the previous section do not necessarily apply to the peptic digestion of proteins and large polypeptides, which is often considerably more extensive than these rules would lead one to expect. We shall not review here the many studies of the proteolytic action of pepsin which have been carried out during the past century but will confine our discussion to that very limited group of investigations in which some knowledge of the bonds cut or the products formed was gained. Tiselius and Eriksson-Quensel (*92*) reported that, even after a considerable period of digestion of acid-denatured ovalbumin with pepsin, appreciable amounts of apparently unchanged substrate remained, together with very short peptides or amino acids; these results strongly suggest an "all-or-none" action. Desnuelle *et al.* (*93*) found that pepsin, in acting on horse globin, first breaks it into very large peptides and medium-sized peptides (about twelve residues); the larger ones are then reduced to the latter size. The determination of the *N*-terminal groups of these peptides showed that alanine, phenylalanine, leucine, and serine predominate. Finally these medium-sized peptides are broken down to tri- and tetrapeptides; at this stage, pepsin appears to show no selectivity. In the hydrolysis of horse serum albumin, pepsin appeared to show no selectivity even from the start, hexa- and pentapeptides being the first recognizable products. It was observed that trypsin and chymotrypsin were very much more specific and that their action ceased at a much earlier stage of degradation of the substrate. Sanger and his co-workers (*94*, *95*) studied the peptic digestion of the A- and B-chains of oxidized insulin. These investigations have provided the most complete information at present available concerning the specific points of attack on very large polypeptides; there is no secondary or tertiary organization in these substrates, however, and they contain the "unnatural" cysteic acid sulfonate groups. As in all these studies, the conditions were rather severe (2 mg. of enzyme per 50 mg. of substrate; 48 hours at 37°) so that fine discrimination of rates was impossible. The results are revealing, nevertheless, and are summarized in Fig. 2, in which rapidly attacked bonds are marked by P and more slowly attacked bonds by p. Of the bonds attacked, only the Phe·Phe and Phe·Tyr linkages have been shown to be susceptible in synthetic substrates. Unexpectedly high susceptibility is shown by Leu·Glu, Glu·AspNH_2, and Leu·Val bonds; of the two Leu·Val sequences of the B-chain, the right-hand one (17–18) is not attacked, possibly because prior attack of the adjacent Tyr·Leu bond produces a free α-amino group. A Val·CysSO_3H bond is attacked in

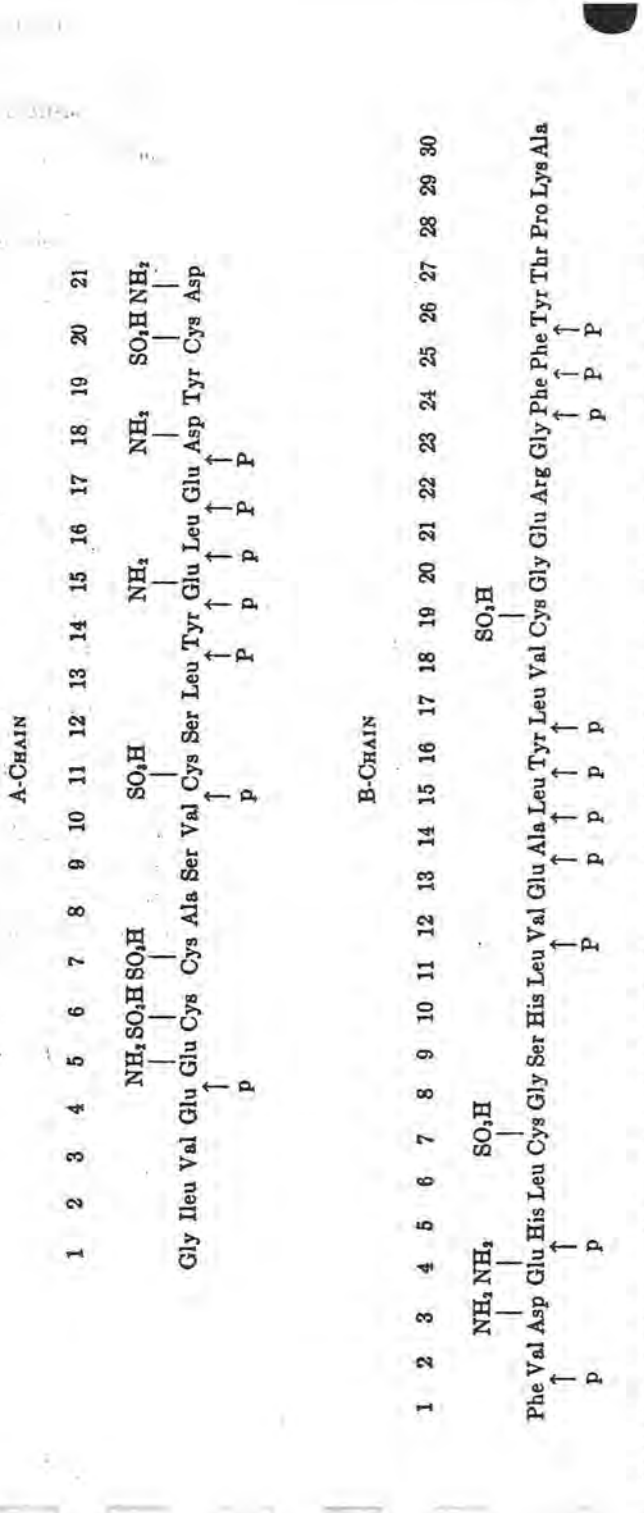


FIG. 2. The attack by pepsin on the A- and B-chains of oxidized insulin. Points of most rapid attack are marked by P; points of slow attack by p.

92. A. Tiselius and I.-B. Eriksson-Quensel, *BJ* **33**, 1752 (1939).
93. P. Desnuelle, M. Rovery, and G. Bonjour, *BBA* **5**, 116 (1950).
94. F. Sanger and H. Tuppy, *BJ* **49**, 481 (1951).
95. F. Sanger and E. O. P. Thompson, *BJ* **53**, 366 (1953).

the A-chain but not in the B-chain. There is rapid attack of Glu·AspNH$_2$ in the A-chain but no attack on the adjacent AspNH$_2$·Tyr, which one would expect to be much more susceptible. These results show that the primary structure of the chain can itself profoundly influence the selectivity of peptic action beyond the actual peptide bond attacked.

The action of pepsin on ribonuclease can be limited to the scission of one bond; this releases the *C*-terminal tetrapeptide and suffices to inactivate the enzyme (*96*). The peptic digestion of oxidized ribonuclease results in the rapid splitting of at least eight peptide bonds (*97*); these results are discussed by Anfinsen (*98*).

Pepsin has also been used in the elucidation of the structure of the corticotropins (*99–101*). White (*99*) isolated the *C*-terminal tetrapeptide Pro·Leu·Glu·Phe after the peptic digestion of corticotropin A. Cole *et al.* (*101*) showed that of the ten peptide bonds in the *C*-terminal sequence of α-corticotropin—Asp·Glu·Ala·Ser·Glu·Ala·Phe·Pro·Leu·Glu·Phe—all were hydrolyzed with the exception of the Pro·Leu bond. It is significant that pepsin hydrolyzes the peptide bond involving the —NH— group (*99*) of a prolyl residue but not that involving the carbonyl group (*101*) of a prolyl residue.

Viswanatha *et al.* have shown that acetyltrypsinogen can be activated by pepsin; maximal activity is reached when about five peptide bonds are split (*102*, *103*). Prolonged peptic digestion yields an active derivative about one-fourth the size of trypsin. The important implications of these findings are beyond the scope of this article.

Li has shown that peptic digestion of the follicle-stimulating hormone (*104*) and of ACTH (*105*) produces products which retain hormonal activity. Croxatto and Croxatto (*106*) and McGlory *et al.* (*107*) have shown that a polypeptide with pressor activity is released by incubation of pepsin with casein and other proteins.

96. C. B. Anfinsen, *JBC* **221**, 405 (1956); M. Sela and C. B. Anfinsen, *BBA* **24**, 229 (1957).
97. J. L. Bailey, S. Moore, and W. H. Stein, *JBC* **221**, 143 (1956).
98. C. B. Anfinsen, "The Enzymes," 2nd ed., Vol. 5, p. 000, in press.
99. W. F. White, *JACS* **75**, 4877 (1953).
100. C. H. Li, *Advances in Protein Chem.* **11**, 102 (1956).
101. R. D. Cole, C. H. Li, J. I. Harris, and N. G. Pon, *JBC* **219**, 903 (1956).
102. T. Viswanatha, R. C. Wong, and I. E. Liener, *BBA* **29**, 174 (1958); *Federation Proc.* **17**, 329 (1958).
103. T. Viswanatha and I. Liener, *BBA* in press (1960).
104. C. H. Li, *JACS* **72**, 2815 (1950).
105. C. H. Li, A. Tiselius, K. O. Pedersen, L. Hagdahl, and H. Carstensen, *JBC* **190**, 317 (1951).
106. H. Croxatto and R. Croxatto, *Science* **95**, 101 (1942); O. Alonso, R. Croxatto, and H. Croxatto, *Proc. Soc. Exptl. Biol. Med.* **52**, 61 (1943).
107. D. H. McGlory, N. S. Olsen, and L. Field, *JBC* **219**, 299 (1956).

Being both a protein and a proteolytic enzyme, pepsin is capable of digesting itself (*108*). This self-digestion or autolysis has already been alluded to (Section IV.B) as the probable cause of the electrophoretic inhomogeneity of many pepsin preparations. Ingram (*109*) observed that, after 24 hours at pH 4 and 45°, 40% of the tyrosine precipitated from a pepsin solution as the free acid. Four peptides were separated by paper chromatography; three were isolated and were all found to have *N*-terminal alanine. Molecular weights of 1080, 1430, and 1850 were assigned on this basis. These appear to parallel the dodecapeptides observed by Desnuelle *et al.* (*93*). Perlmann (*110*) has reported that when pepsin is incubated at 37° and pH 5.6 for 12 to 48 hours, and then dialyzed, the dialyzate contains fragments, presumably of considerably lower molecular weight than the original pepsin, which are enzymically active. The dialyzate shows a considerably altered specificity, however, having 64% of the original activity (per milligram of nitrogen) toward *N*-acetyl-L-phenylalanyl-L-diiodotyrosine, but only 3% toward hemoglobin. Apparently, the "activity site" is intact but the "affinity site" is considerably altered. This work has been confirmed (*111*). Autolysis in 8 *M* urea (*81*, *112*) produces several electrophoretically distinguishable components, but only that having the mobility of the original pepsin is active; its enzymic efficiency is somewhat greater than that of the original pepsin, but despite an apparent loss of tyrosine its molecular weight is only slightly reduced.

C. Inhibition and Inactivation; Chemical Derivatives of Pepsin

The naturally occurring pepsin inhibitor, split from pepsinogen by the action of pepsin, has been discussed in Section V. This polypeptide inhibits pepsin between pH 5 and 6. It was at one time thought (*55*) that it contained a relatively large proportion of arginine, and its inhibitory action appeared logical on the basis of strong complexing through coulombic attraction to the negatively charged pepsin. However, improved amino acid analysis (see Table II and Fig. 1) makes this conclusion somewhat less obvious. This idea appears to be supported, however, by the observation of Katchalski *et al.* (*113*) that poly-L-lysine combines rapidly with pepsin to form an inactive complex with a dissociation constant of approximately 0.5×10^{-4}, as measured by degree of inhibition. Poly-L-lysine is effective

108. Ref. 3, pp. 74, 81.
109. V. M. Ingram, *Nature* **167**, 83 (1951).
110. G. E. Perlmann, *Nature* **173**, 406 (1954).
111. M. Funatsu and K. Tokuyasu, *Proc. Japan Acad.* **35**, 139 (1959) [*C. A.* **53**, 16256g (1959)].
112. G. E. Perlmann and M. J. Mycek, *in* "Symposium on Protein Structure (Paris, 1957)" (A. Neuberger, ed.), p. 179. Methuen, London, and Wiley, New York, 1958.
113. E. Katchalski, A. Berger, and H. Neumann, *Nature* **173**, 998 (1954).

both at pH 1.7 and pH 6.0. No other effective competitive inhibitors for pepsin have been reported.

By treatment with ketene, both the amino and phenolic groups of pepsin can be acetylated. The *O*-acetyl groups may then be selectively hydrolyzed with weak acid. In this way, Herriott (*114, 115*) was able to demonstrate that *N*-acetylation has no effect on enzymic activity but that *O*-acetylation leads to progressive loss of activity. Hollander (*116*) has shown that acetylation of both amino groups and tyrosyl residues produces a loss of activity. Deamination of the free amino groups with nitrous acid has no effect on enzymic activity (*117*). The association of tyrosyl residues with activity is further supported by observations that activity is reduced by their diazotization (*117*) or iodination (*118, 119*). Philpot and Small (*117*) found that half of the tyrosines could not be diazotized even by a large excess of nitrous acid. Li (*119*) found that twelve of the tyrosines in the native protein could be iodinated as readily as free tyrosine, whereas the remainder could be iodinated only in the denatured molecule. A portion of the tyrosines are apparently bonded or "masked" in such away that their chemical reactivity differs markedly from that of the others. Herriott *et al.* have found (*120*) that sulfur mustard, $ClCH_2CH_2SCH_2CH_2Cl$, alkylates the carboxylate groups of pepsin at pH 5.5 to 6.0, leading to progressive loss of activity. This suggests, as seems reasonable on other grounds (Section VIII.D), that carboxyl groups are associated with the active center, although a general disruption of structure could lead to the same result. Herriott has shown (*121*) that a plot of the logarithm of remaining activity against equivalents of reagent bound per mole of pepsin gives straight lines for the reaction of pepsin with iodine, sulfur mustard, and ketene. Ninety per cent inactivation is produced by iodination (two atoms per ring) of half of the tyrosines and by alkylation of only one-third of the carboxyls.

VIII. Kinetics and Mechanism

A. pH Optimum

Northrop (*122*) found that pepsin was active on protein substrates over a pH region extending from below 2 to above 4, with a maximum at

114. R. M. Herriott and J. H. Northrop, *J. Gen. Physiol.* **18**, 35 (1934).
115. R. M. Herriott, *J. Gen. Physiol.* **19**, 283 (1935).
116. V. Hollander, *Proc. Soc. Exptl. Biol. Med.* **53**, 179 (1943).
117. J. St. L. Philpot and P. A. Small, *BJ* **32**, 542 (1938).
118. R. M. Herriott, *J. Gen. Physiol.* **20**, 335 (1937); **31**, 10 (1947).
119. C. H. Li, *JACS* **67**, 1065 (1945).
120. R. M. Herriott, M. L. Anson, and J. H. Northrop, *J. Gen. Physiol.* **30**, 185 (1946).
121. R. M. Herriott, in "The Mechanism of Enzyme Action" (W. D. McElroy and B. Glass, eds.), p. 41 *et seq.* Johns Hopkins, Baltimore, 1954.
122. J. H. Northrop, *J. Gen. Physiol.* **5**, 263 (1922).

1.8. Baker (*30*) observed an optimum at pH 2.0 for acetyl-L-phenylalanyl-L-tyrosine and acetyl-L-phenylalanyl-L-diiodotyrosine. Carbobenzoxy-L-glutamyl-L-tyrosine and carbobenzoxy-L-glutamyl-L-phenylalanine show higher pH optima of 3.8 and 4.4 (*87*), respectively; this is probably to be attributed to the effect of the γ-carboxyl group (Section VIII.D). Pope and Stevens (*123*) reported that crystalline pepsin gave a narrower pH optimum in its action on horse hemoglobin and diphtheria antitoxin than did a crude, amorphous pepsin preparation. Bull and Currie (*124*) found the optimum for the action of pepsin on ovalbumin to be at about pH 1; (see Section VIII.B for further discussion.

Many proteins are rapidly denatured at the low pH values at which pepsin's action is optimal. Christensen (*125*) found that for native ovalbumin and native β-lactoglobulin the optimum is at about pH 1 and the pH-activity curve is rather narrow, the rate being very slow at pH values exceeding 1.5 to 2.0, whereas for these proteins in the denatured state the optimum is at 1.6 to 1.8 and the curve is considerably broader, appreciable activity being exhibited at pH 3.4. The rate-determining step in the hydrolysis of these proteins may be the rate of acid denaturation, for it was found that the two rates were comparable. This supports the suggestion of Linderstrøm-Lang *et al.* (*126*) that proteolytic enzymes attack only the unfolded form of the protein. Sri Ram and Maurer (*127*) found a pH optimum of 2 for peptic hydrolysis of bovine serum albumin, a result which may be assumed to refer to the denatured protein.

B. Kinetics

Very little detailed study has been given to the kinetics of the action of pepsin. Bull and Currie (*124*) made a careful study of its action on ovalbumin and offered an interpretation of the form of the pH-activity curve on the alkaline side. They suggested that the enzyme-substrate complex has a pK_a' at about 2.1 and acquires a proton before going over to its activated form, which then yields the reaction products. A further interpretation of these results has been given by Laidler (*128*). The hydrolysis rates were extrapolated to zero time and therefore were assumed to refer to the native substrate, but, in view of the findings of Tiselius and Eriksson-Quensel (*92*) and Christensen (*125*), it is possible that acid denaturation of the substrate is still the rate-determining step. It is probably more than

123. C. G. Pope and M. F. Stevens, *Brit. J. Exptl. Pathol.* **32**, 314 (1951).
124. H. B. Bull and B. T. Currie, *JACS* **71**, 2758 (1949).
125. L. K. Christensen, *ABB* **57**, 163 (1955).
126. K. Linderstrøm-Lang, R. D. Hotchkiss, and G. Johansen, *Nature* **142**, 996 (1938).
127. J. Sri Ram and P. H. Maurer, *ABB* **70**, 185 (1957).
128. K. J. Laidler, "The Chemical Kinetics of Enzyme Action," p. 144 *et seq.* Oxford Univ. Press, London and New York, 1958.

a coincidence that the apparent enthalpy of activation for formation of the enzyme-substrate complex, 31.4 kcal., is nearly as large as the 35 kcal. (*129*) observed for the acid denaturation of ovalbumin.

Casey and Laidler (*91*) measured K_m and k_3 at 23.8° to 47.2° for the action of pepsin on carbobenzoxy-L-glutamyl-L-tyrosine and carbobenzoxy-L-glutamyl-L-tyrosine ethyl ester at pH 4.0, the activity optimum. Michaelis-Menten kinetics were found to be accurately obeyed. Table IV summarizes the data obtained, including values of ΔH_3^* and ΔS_3^*, the enthalpy and entropy corresponding to the reaction of the enzyme-substrate complex to yield the products, and ΔH^* and ΔS^*, the corresponding quantities associated with K_m, which is not known to be an equilibrium constant in these reactions. The fact that ΔS^* is positive, despite a possible negative contribution from k_3, may arise from a charge neutralization between the γ-carboxylate group of the substrate and a positively charged center on the enzyme; this effect may also be responsible for the considerably higher pH optimum, compared to that observed for other substrates.

Baker (*130*), working at 37° and pH 2.0 (the optimum), found values of K_m of 0.0024 and 0.0063 mole·l.$^{-1}$ for the action of pepsin on acetyl-L-phenylalanyl-L-tyrosine and acetyl-L-tyrosyl-L-tyrosine. The initial velocities were in accord with Michaelis-Menten kinetics with respect to initial substrate concentration, but at each substrate concentration the reaction was observed to be first-order in time, suggesting that one of the products of the reaction acted as a competitive inhibitor with a K_i about equal to K_m (*131*). Although there were some experimental difficulties, not fully explainable (*132, 133*), it appears to have been demonstrated that, whereas tyrosine has no inhibitory effect, acetyl-L-phenylalanine has a K_i for this reaction which is of the proper magnitude to account for at least part of this effect (*133*). It should be noted that possible transpeptidation (Section VIIIC) is not taken account of in this work; this, too, might account for part of the observed deviation from Michaelis-Menten kinetics.

C. Transpeptidation

Neumann *et al.* (*134*) have made observations of transpeptidations catalyzed by pepsin which are of great significance for the understanding of peptic action. They observed that, when *N*-acetyl-L-tyrosyl-L-tyrosine,

129. H. K. Cubin, *BJ* 23, 25 (1929).
130. L. E. Baker, *JBC* 211, 701 (1954).
131. G. E. Briggs and J. B. S. Haldane, *BJ* 19, 338 (1925).
132. N. M. Green, *Nature* 178, 145 (1956).
133. L. E. Baker, *Nature* 178, 145 (1956).
134. H. Neumann, Y. Levin, A. Berger, and E. Katchalski, "Proceedings of the International Symposium on Enzyme Chemistry, Tokyo-Kyoto (1957) (K. Ichihara, ed.), p. 129. Academic Press, New York, 1958; *BJ* 73, 33 (1959).

TABLE IV

ACTION OF PEPSIN ON SYNTHETIC SUBSTRATES
(pH 4.0, temperature = 31.6°, ref. *128*)

Substrate	k_3 (sec.$^{-1}$)	K_m (mole·l.$^{-1}$)	ΔH_3^* (kcal.)	ΔS_3^* (kcal.)	ΔH^* (kcal.)	ΔS^* (kcal.)
Cbz-L-glutamyl-L-tyrosine	0.00108	0.00189	20.7	−4.5	23.1	16.1
Cbz-L-glutamyl-L-tyrosine ethyl ester	0.00141	0.00178	17.2	−19.8	20.2	4.6

N-*p*-aminobenzoyl-L-tyrosyl-L-tyrosine, *N*-carbobenzoxy-L-glutamyl-L-tyrosine, and *N*-carbobenzoxy-L-phenylalanyl-L-tyrosine are incubated with pepsin, paper chromatograms show the formation of tyrosyltyrosine as well as the expected products of hydrolysis. Pepsin was found to be incapable of removing the *N*-acetyl and other *N*-terminal groups from these peptides, so the formation of tyrosyltyrosine cannot be explained in this way even for the first two of these peptides. The true explanation appears to be that transpeptidation occurs as indicated by Eq. (2), followed by

$$2\text{AcTyrTyr} \rightarrow \text{AcTyr} + \text{AcTyrTyrTyr} \quad (2)$$

attack on the tripeptide [Eqs. (3) and (4)]. Parallel reactions occur for the

$$\text{AcTyrTyrTyr} \rightarrow \text{AcTyr} + \text{TyrTyr} \quad (3)$$

$$\text{AcTyrTyrTyr} \rightarrow \text{AcTyrTyr} + \text{Tyr} \quad (4)$$

other dipeptides. It should be noted that breaking the *C*-terminal bond of the tripeptide regenerates the starting dipeptide. It can be shown that deviations from Michaelis-Menten kinetics may be expected if such a series of reactions occurs.

It is clear from these examples that a free α-amino group is not necessary for transpeptidation, but it was observed that, if the α-carboxyl group is blocked, as in *N*-carbobenzoxy-L-glutamyl-L-tyrosine amide, no transpeptidation occurs. Evidently, when the peptide linkage of the dipeptide is attacked by the enzyme, the amino group rather than the carboxyl group is activated and is transferred to the α-carboxyl group of another dipeptide molecule. Since transpeptidation could be observed at a pH as low as 1.5, the undissociated carboxyl group evidently acts as the acceptor. This activation of the amino group is in contrast to the action of many enzymes, including trypsin and chymotrypsin, in which it is well estab-

lished (*135–137*) that the carboxyl group is activated. With pepsin, it may be that the substrate is acylated by the enzyme (Section VIII.D).

D. Possible Mechanisms

The evidence obtained so far concerning the detailed action of pepsin is very meager, so that any proposed mechanism must be highly speculative. The following observations appear to be of chief importance:

1. The pH optimum is at strongly acid pH, suggesting the participation of carboxyl groups at the active site; one or more of these may be of abnormally low pK_a'. The participation of carboxyl groups is also supported by the observation of Herriott *et al.* (*120*) that alkylation with sulfur mustard reduces activity.
2. Glutamyl peptides have considerably higher pH optima than the other substrates tested.
3. The dependence of V_{max} on pH on the alkaline side of the pH optimum suggests that the enzyme-substrate complex, at least in the case of ovalbumin (*124*), may acquire a proton prior to or during activation.
4. The course of transpeptidations catalyzed by pepsin indicates that the amino group of the peptide linkage is activated (*134*), with possible acylation by the enzyme.

These facts are incorporated as well as appears possible at present into the mechanism shown in Fig. 3. The hypothetical positive receptor site (*128*) might be the lysine ε-ammonium group or one of the two guanidinium groups of arginine; if the former, mild acetylation, which does not affect pepsin's activity toward protein substrates (*114, 115*), should alter the pH optimum for glutamyl peptides.

Missing from this picture is an explanation of the acid branch of the pH-activity curve. No study of k_3 (or V_{max}) and K_m has been made in this region. Baker's work (*30*) was carried out at substrate concentrations too low to ensure that V_{max} was being observed, and so it cannot be decided at present whether the acid branch corresponds to protonation of groups in the enzyme, in the enzyme-substrate complex, or in both. In addition, possible acid inactivation of pepsin at very low pH (Section IX.B) must be taken into account. These questions remain for future investigation, but it is already evident that the mechanism of action of pepsin is very different from those of the other known proteolytic enzymes.

135. M. Bergmann and H. Fraenkel-Conrat, *JBC* **119**, 707 (1937).
136. C. S. Hanes, F. J. R. Hird, and F. A. Isherwood, *Nature* **166**, 288 (1950).
137. Y. D. Halsey, in "The Proteins" (H. Neurath and K. Bailey, eds.). Vol. II, Part B, p. 1135. Academic Press, New York, 1954.

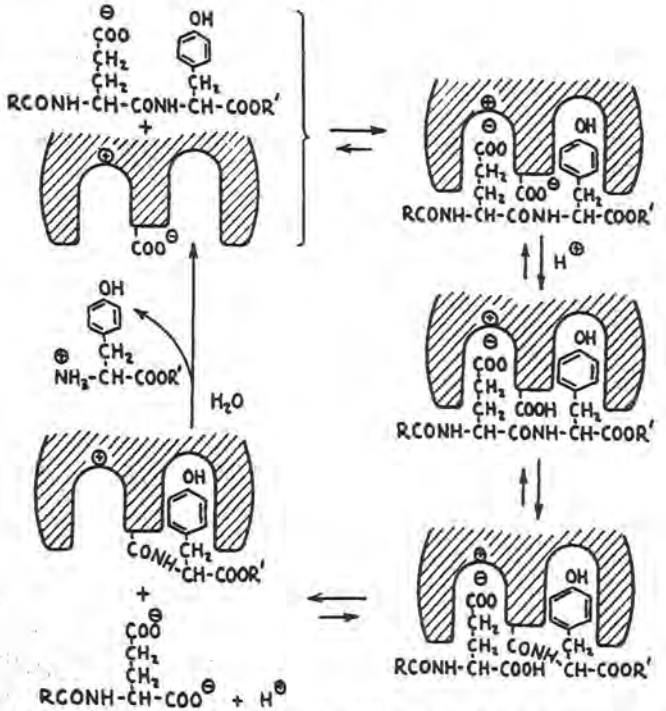


FIG. 3. A proposed mechanism of action of pepsin; N-acyl-L-glutamyl-L-tyrosine ester represented as substrate.

IX. Denaturation and Inactivation of Pepsin and Pepsinogen

A. Pepsinogen

Herriott (*26*) found that pepsinogen has considerably greater thermal stability than pepsin. It may be boiled for short periods and will recover its activity on cooling, although addition of salt will precipitate denatured protein from the hot solution. The thermal denaturation is thus reversible. It was found to be characterized by an enthalpy of 31 kcal. Pepsinogen is more stable to alkali than pepsin but undergoes reversible denaturation at pH values of 9 or greater with expulsion of two protons, being in this respect somewhat similar to β-lactoglobulin (*138*) and different from pepsin (see below), in which the denaturation and proton expulsion are essentially irreversible. In acid, pepsinogen undergoes conversion to pepsin (Section V).

B. Pepsin

1. *Alkali Denaturation*

The denaturation and inactivation of pepsin, particularly in the pH

138. C. Tanford, L. G. Bunville, and Y. Nozaki, *JACS* **81**, 4032 (1959).

region 6 to 7 or higher, where it is commonly termed "alkali denaturation," has been the subject of many studies because it presents several features of unusual interest. Most of these studies have been carried out at a pH near 6 and at room temperature. As already indicated, pepsin is stable at low pH but is sensitive to higher pH and is very rapidly inactivated at pH 6.5 to 7.0 and above. Tammann (*139*) found the reaction to be approximately first-order in time, although showing a deviation toward a higher order at longer times. Arrhenius (*140*) described the process on the assumption that it was first-order in time. Michaelis and Rothstein (*141*) reported the rate to be 3/2-order with respect to time. Northrop (*2*) found the inactivation to be first-order in time and to be paralleled by formation of acid-insoluble protein. Loughlin (*142*) measured the inactivation rate at several pH values but did not report the reaction order. Steinhardt (*143*) made a very careful study of the reaction over the pH interval 6.0 to 6.8 and at varying ionic strengths. Measurements at 15° and 25° gave an over-all activation energy of 63.5 kcal. The rate was reported to be first-order in concentration and time, but there were indications of deviations toward a higher order at high degrees of inactivation. The inactivation rate at pH 6.0 to 6.5 was found to vary as $[H^+]^{-5}$; at higher pH, there was deviation toward a dependence of somewhat lower order. A dependence on $[H^+]^{-5}$ means a 100,000-fold change in the denaturation rate over 1 pH unit, which to casual observation could easily appear as a discontinuity. Michaelis and Rothstein (*141*) and Buzzell and Sturtevant (*144*) observed the rate of inactivation to vary approximately as $[H^+]^{-4}$; Edelhoch (*145*) found it to vary as $[H^+]^{-3.4}$, and Yanari and Bovey (*79*) observed the same dependence for the accompanying release of protons (see below) in the pH region near neutrality. The inactivation rate is very sensitive to ionic strength, increasing at first rapidly (for example, forty-fold between 0.012 and 0.10) (*143*), then leveling off (*143*, *145*), and finally decreasing again at high ionic strength to nearly the value at low ionic strength (*144*).

Calorimetric studies of the alkali denaturation process have given results of considerable interest which are at present difficult to interpret.

139. G. Tammann, *Z. physik. Chem.* (*Leipzig*) **18**, 426 (1892).
140. S. Arrhenius, "Immunochemistry," p. 57. Macmillan, New York, 1907.
141. L. Michaelis and M. Rothstein, *Biochem. Z.* **105**, 60 (1920).
142. W. J. Loughlin, *BJ* **27**, 1779 (1933).
143. J. Steinhardt, *Nature* **138**, 74 (1936); *Kgl. Danske Videnskab-Selskab Mat.-fys. Medd.* **14**, 11 (1937).
144. A. Buzzell and J. M. Sturtevant, *JACS* **74**, 1983 (1952).
145. H. Edelhoch, *JACS* **80**, 6640 (1958).

Kistiakowsky *et al.* (*146*, *147*) measured the heat of denaturation as the difference between the heats of reaction of the native and denaturated protein with an amount of alkali sufficient to bring them both to the same pH; this quantity decreased from 85 kcal. (endothermic) at pH 4.3 to nearly zero at pH 6.8. Sturtevant *et al.* (*144*, *148*, *149*) measured the heat of denaturation directly and found it to obey a first-order law in time under conditions where the rate of inactivation showed a deviation toward a higher order, and to depend on $[H^+]^{-1.4}$ instead of $[H^+]^{-3.9}$, as the inactivation rate did. Thus, the measured heat effect and the inactivation process are not rigorously associated. They observed maxima in the heat effect as a function of pH. At 35° the maximum was at pH 6.4 and ΔH was about 50 kcal.; at 15° the maximum was at pH 7.1 and ΔH was about 15 kcal. On either side of the maxima, ΔH fell off to nearly zero within 0.2 to 0.4 pH unit, suggesting that outside this pH region the heat contents of the native and denatured pepsin are the same.

The acid-base titration curves of native and denatured pepsin show a hysteresis in the pH region 6 to 7 (*26*, *145*, *146*). There appear to be about six more protons bound to the native pepsin up to pH 7 than are taken up by the denatured pepsin on passing back through this region from high pH. Edelhoch (*145*) has observed that protons are released at a measurable rate during alkali denaturation. Their rate of release is approximately the same as that of the inactivation and shows the same complex dependence on ionic strength. Yanari and Bovey (*79*) confirmed this finding, observing a maximum of about two and one-half protons released during complete inactivation. These two processes are thus very closely associated.

Yanari and Bovey (*150*) found that, at near-neutral pH, the rate of inactivation of pepsin is second-order at higher pepsin concentrations and relatively low temperatures (e.g., 2% pepsin, 12°), in accord with the results of Buzzell and Sturtevant (*144*), whereas at higher temperatures and lower concentrations (e.g., 0.02% pepsin, 25°) the reaction is first-order. Under intermediate conditions, the rates can be analyzed to demonstrate competitive first- and second-order reactions. Casey and Laidler (*151*, *152*), working at considerably lower pH (4.83) and higher tem-

146. J. B. Conn, G. B. Kistiakowsky, and R. M. Roberts, *JACS* **62**, 1895 (1940).
147. J. B. Conn, D. C. Gregg, G. B. Kistiakowsky, and R. M. Roberts, *JACS* **62**, 2080 (1941).
148. M. Bender and J. M. Sturtevant, *JACS* **69**, 607 (1947).
149. J. M. Sturtevant, *J. Phys. Chem.* **58**, 97 (1954).
150. S. S. Yanari and F. A. Bovey, *Abstr. Am. Chem. Soc. 136th Meeting*, p. 59C. Atlantic City (September, 1959).
151. E. J. Casey and K. J. Laidler, *Science* **111**, 110 (1950); *JACS* **73**, 1455 (1951).
152. Ref. 128, p. 361 *et seq.*

perature (50° to 62°), and employing a 200-fold range of pepsin concentration, observed that at high concentration the reaction was nearly first-order in both time and concentration but tended toward higher orders at increasing dilution, the order approaching 4 to 5 at concentrations less than 0.006%. These observations may help to explain the somewhat discordant results of other workers but cannot themselves be fully understood and reconciled as yet. Apparently, a cooperation of pepsin molecules to form dimers or larger clusters may precede inactivation, as suggested by Laidler (*151, 152*). Perlmann (*62*) observed that at pH 4.6 loss of activity begins at 60° and is accompanied by increases in levorotation.

The decrease in sedimentation constant (*19*) (3.3*s* to 2*s*), increase in intrinsic viscosity [0.05 to 0.13 (*153*); 0.03 to 0.21 (*35*)], and increase in levorotation (*154–156*) which accompany alkali denaturation indicate a partial unfolding of the molecule. There is an accompanying blue shift of the ultraviolet spectrum (*145–157*) which may arise from the removal of tyrosyl and tryptophyl residues from predominantly hydrophobic sites to predominantly aqueous sites during the unfolding (*86*). Whereas the ultraviolet spectrum of native pepsin is insensitive to changes in pH in the region 0 to 5.5, alkali-denatured pepsin shows very marked changes (*157*).

Let us suppose, following Steinhardt (*143*) and Levy and Benaglia (*158*), that the pepsin molecule passes through several ionized states P_1, $P_2 \cdots$, connected by equilibrium constants, K_1, $K_2 \cdots$:

$$P_1 \overset{K_1}{\rightleftharpoons} P_2 + H^+$$

$$P_2 \overset{K_2}{\rightleftharpoons} P_3 + H^+ \quad \text{etc.}$$

Let us further suppose that the state P_6, which has lost five protons, passes over into the unfolded, inactive form, P_D, much more rapidly than any of the preceding states. We may then write the over-all scheme for denaturation in the pH range 6 to 7, as in Eq. (5), competing first- and

$$P_1 \overset{K_1K_2K_3K_4K_5}{\rightleftharpoons} P_6 + 5H^+ \qquad (5)$$

$$k' \downarrow \qquad \downarrow k'' [P_6]$$

$$P_D + \sim 2H^+$$

153. Ref. 121, p. 117.
154. M. A. Golub and M. A. Pickett, *J. Polymer Sci.* **13**, 427 (1954).
155. H. Edelhoch, *JACS* **78**, 2644 (1956).
156. B. Jirgensons, *ABB* **74**, 70 (1958).
157. S. S. Yanari and F. A. Bovey, *Federation Proc.* **17**, 340 (1958).
158. M. Levy and A. E. Benaglia, *JBC* **186**, 829 (1950).

second-order pathways for the inactivation of P_6 being represented to take account of the apparent competing first- and second-order processes observed by Yanari and Bovey (*150*). If the dissociation of protons is assumed to be much faster than the formation of P_D from P_6, this reaction scheme conforms to the observed kinetics. The detailed analysis of such a system of reactions (*143, 158*) shows that, if any of the rate constants k_5, $k_4 \cdots$ for the inactivation of P_5, $P_4 \cdots$ are not negligible in comparison to k_6, *or* if the equilibrium constants K_1, $K_2 \cdots$ are not negligible in comparison to $[H^+]$, then the over-all denaturation rate will exhibit a dependence on $1/[H^+]$, which becomes less than fifth-order as the pH is increased; this we have seen to be experimentally observed (*79, 143–145*). Steinhardt (*143*) proposed that the dissociating groups were the α-ammonium groups of cystine, but this suggestion can no longer be entertained in view of the established composition and structure of pepsin (Section VI). It appears most probable at present that these groups are carboxyl groups whose pK_a' values are considerably greater than the "normal" 4.0 to 4.5 either because they are hydrogen-bonded to each other (*145*) or because they are clustered in one portion of the molecule (*159*). Assuming that they dissociate reversibly, we have (*143*):

$$\Delta H^* = \Delta H_6^* + \Delta H_1 \cdots \Delta H_6$$

$$\Delta S^* = \Delta S_6^* + \Delta S_1 \cdots \Delta H_6$$

where ΔH^* and ΔS^* are the observed over-all enthalpy and entropy of activation for denaturation, found by Steinhardt (*143*) to be 63.5 kcal. and 135.9 e.u., respectively. Assuming the enthalpies of ionization, ΔH_i, to be 1 kcal. for each carboxyl group and the entropies of ionization, ΔS_i, to be -20 e.u. (*28*), we find that, for the transformation of P_6 to P_D, $\Delta H_6^* =$ 58.5 kcal., approximately equal to that observed by Casey and Laidler (*151, 152*) at pH 4.83, where presumably ionization is not a necessary prelude to denaturation. For the entropy, $\Delta S_6^* = 236$ e.u., a large value, but not outside the known range for protein denaturation. When ionization of five carboxyl groups has occurred, the electrostatic repulsion is apparently sufficient to force an unfolding of a portion of the molecule, this portion containing the active site. We may suppose that unfolding results in an increase in the acid strength of these and other neighboring carboxyl groups, as indicated by the more nearly normal pH-titration curve of denatured pepsin, and that this in turn results in the observed time-dependent expulsion of approximately two additional protons (*79*). The complex effect of ionic strength on the denaturation rate may be ex-

159. K. Linderstrøm-Lang, *Compt. rend. trav. lab. Carlsberg* **15**, No. 7 (1924); C. Tanford and J. G. Kirkwood, *JACS* **79**, 5333 (1957); C. Tanford, *JACS* **79**, 5340 (1957).

plained if we imagine that at lower ionic strengths the pK_a' values of the clustered carboxyl groups are increased, encouraging ionization and unfolding, whereas at high ionic strength this effect is overcome and unfolding is retarded because of lessened repulsion between carboxylate groups.

The participation of carboxylate groups in alkali denaturation is supported by the observation (*160, 161*) that metals with high affinity for carboxyl groups (Cu^{++}, Pb^{++}, Cd^{++}, and Zn^{++}) accelerate the process. Urea, although it does not unfold pepsin at lower pH and room temperature (*39, 62, 80*), accelerates denaturation at higher pH (*160, 161*) or temperature (*62*). The slopes of log-log plots of denaturation rate versus urea concentration suggest that the process proceeds by way of complexes containing several denaturant molecules (*160*).

Although the unfolding which accompanies alkali denaturation is usually regarded as irreversible, Northrop (*162*) has shown that a small fraction of the initial enzymic activity can be recovered by allowing alkali-inactivated pepsin to stand for some time at pH 5.4, and that from the renatured preparation crystalline pepsin can be isolated which, except possibly for a higher optical rotation (very inaccurately measured), is identical to the original native protein. This "refoldable" fraction may perhaps be a somewhat modified form occurring in a minor proportion in the original preparation.

2. *Acid Denaturation*

Pepsin gradually loses activity in strongly acid solution (*163, 164*). At pH 1.8 and 50°, loss of activity is paralleled by loss of protein nitrogen; under these conditions, autolysis probably also contributes to the loss of activity (*164*). In very strong acid, there appears to be first a denaturation, followed by hydrolysis of peptide bonds (*164*). The rate of inactivation is directly proportional to $[H^+]$ (*163*).

3. *Inactivation by Radiation*

The quantum yield for the inactivation of pepsin by ultraviolet radiation increases from about 0.002 at pH 4 to 6 to about 0.009 at pH 6.0 to 6.2, where alkali denaturation begins (*165*). This result appears to be intuitively reasonable but is not so easily explained in detail (*166*).

160. H. Edelhoch, *BBA* **22**, 401 (1956).
161. H. Edelhoch, *JACS* **80**, 6648 (1958).
162. J. H. Northrop, *J. Gen. Physiol.* **14**, 713 (1931).
163. R. Ege, *Z. physiol. Chem.* **143**, 159 (1925); *Biochem. Z.* **244**, 243 (1932); *Z. physiol. Chem.* **224**, 129 (1934).
164. J. H. Northrop, *J. Gen. Physiol.* **16**, 33 (1932).
165. A. D. McLaren and S. Pearson, *J. Polymer Sci.* **4**, 45 (1949).
166. A. D. McLaren and C. Lewis, *J. Polymer Sci.* **3**, 379 (1949).

Bellamy and Lawton (*167*) studied the inactivation of pepsin by high-energy electrons. They found the ionic yield for the direct inactivation of dry pepsin to be 0.22 molecule per ion pair. As is commonly observed, the indirect inactivation of pepsin in aqueous solution, where free radicals generated by interaction of the electrons with water are the reactive species, was only about one-tenth as efficient.

X. Pepsin-like Proteases

The occurrence of a gastric cathepsin has been supported by some workers (*123, 168–171*) and opposed by others (*172–174*). Merten *et al.* (*170*) demonstrated the presence of two fractions in pepsin preparations with differing pH optima by means of paper electrophoresis. The isolation, crystallization, and partial characterization by Tang *et al.* (*175, 176*) of a protease from human gastric juice which they named *gastricsin* appears to have partially resolved this controversy. Gastricsin can be distinguished from human and pork pepsins by its higher pH optimum of 3.0, greater heat stability, and lower electrophoretic mobility on starch gel. However, there is some doubt whether gastricsin and gastric cathepsin are identical, since the reported activation of cathepsin by HCN, iodoacetic acid, and H_2S (*168, 169, 171*) does not occur with gastricsin (*175*). Tang *et al.* (*175*) also point out the possibility that gastricsin may be an autolysis product of pepsin; in accord with this possibility is the fact that autolysis of pepsin produces active material less anionic than native pepsin (*49*).

Rennin is the milk-clotting enzyme found in the fourth stomach of the calf; crystalline rennin is prepared from commercial rennet (*177*). Berridge (*177*) has reviewed earlier studies on rennin. Fish (*178*) has shown that rennin is a proteolytic enzyme with a specificity nearly identical to that of pepsin, as evidenced by its action on the B-chain of oxidized insulin; the possibility that the proteolytic activity of purified rennin is due to pepsin contamination was ruled out. Fish also found that purified

167. W. D. Bellamy and E. J. Lawton, *Nucleonics* No. 4, 54 (1954).
168. E. Freudenberg, *Enzymologia* **8**, 385 (1940).
169. S. Buchs, *Enzymologia* **13**, 208 (1949); *ibid.* **16**, 193 (1953); *Z. physiol. Chem.* **301**, 201 (1955).
170. R. Merten, G. Schramm, W. Grassman, and K. Hannig, *Z. physiol. Chem.* **289**, 173 (1952).
171. Z. Ramer, *Z. physiol. Chem.* **296**, 73 (1954).
172. L. K. Christensen, *Scand J. Clin. Lab. Invest.* **7**, 225 (1955) [*C. A.* **50**, 6622i (1956)].
173. L. W. Masch and I. Huchting, *Z. physiol. Chem.* **301**, 49 (1955).
174. W. H. Taylor, *BJ* **71**, 373 (1959).
175. J. Tang, S. Wolf, R. Caputto, and R. E. Trucco, *JBC* **234**, 1174 (1959).
176. V. Richmond, J. Tang, S. Wolf, R. E. Trucco, and R. Caputto, *BBA* **29**, 453 (1958).
177. N. J. Berridge, "Methods in Enzymology" Vol. II, p. 69, 1955.
178. J. C. Fish, *Nature* **180**, 345 (1957).

rennin, unlike commercial rennet, does not have phosphoamidase activity, as claimed by Holter and Li (*179*); i.e., it will not hydrolyze *N*-(*p*-chlorophenyl)amidophosphoric acid.

Schwander *et al.* (*180*) have determined the amino acid composition of rennin by starch column chromatography (*181*). These workers also estimated the isoelectric point of rennin as 4.5 from paper electrophoresis studies and found a molecular weight of 40,000 by sedimentation and diffusion measurements. Jirgensons *et al.* (*182*) have further purified rennin on DEAE-cellulose (*24*). The *N*-terminal residue is glycine and the *C*-terminal residue leucine or isoleucine (*182*). Optical rotatory dispersion studies indicate that rennin belongs to a class of proteins, including pepsin, whose dispersion constants are abnormal (*182*); the dispersion constant of native rennin (210 mμ) is lower than that of denatured rennin (222 mμ).

Gastric gelatinase was isolated and characterized by Northrop (*183*). Gelatinase is removed from pepsin preparations only after repeated recrystallizations. The activity of gelatinase is reduced only slightly under conditions which cause the so-called alkali denaturation of pepsin.

Pepsin-like enzymes from sources other than gastric juice have also been studied. Lundquist and Seedorff (*184*) have isolated a crystalline enzyme from human seminal fluid which resembles pepsin in its pH optimum, specificity, and lack of stability near neutrality; it occurs as a proenzyme which, like pepsinogen, is stable at pH 8.5 and is activated at low pH values. Cathepsin A (*185*) and a crystalline enzyme from *Aspergillis saitoi* (*186*) have certain similarities to pepsin.

Note Added in Proof. Ryle and Porter (*187*) have recently reported detailed studies on two proteases, parapepsins I and II, which they isolated from crude pepsin preparations. Parapepsins I and II differ from pepsin in containing no phosphate group and having different *N*-terminal amino acid residues. In their molecular weights and sedimentation constants the parapepsins are similar to pepsin, although they are slightly less acidic. Studies on synthetic substrates and on the sites of cleavage of the B chain of oxidized insulin indicate that the specificities of pepsin and the parapepsins are similar. Although parapepsin II resembles pepsin in its instability at pH 7, it differs further in having a higher pH optimum, similar to that of the so-called gastric cathepsins (see Section X). Parapepsin I is stable at pH 7 and appears similar to gastric gelatinase (*183*).

179. H. Holter and S. O. Li, *Acta Chem. Scand.* **4**, 1321 (1950).
180. H. Schwander, P. Zahler, and H. Nitschmann, *Helv. Chim. Acta* **35**, 553 (1952).
181. W. H. Stein and S. Moore, *JBC* **176**, 337 (1948).
182. B. Jirgensons, T. Ikenaka, and V. Gorguraki, *Makromol. Chem.* **28**, 96 (1958).
183. J. H. Northrop, *J. Gen. Physiol.* **15**, 29 (1931).
184. F. Lundquist and H. H. Seedorff, *Nature* **170**, 1115 (1952); *Acta Physiol. Scand.* **25**, 178 (1952).
185. J. S. Fruton, "The Enzymes," 2nd ed., Vol. 4, p. 000.
186. F. Yoshida, *in* B. Hagihara, "The Enzymes," 2nd ed., Vol. 4, p. 000.
187. A. A. Ryle and R. R. Porter, *BJ* **73**, 75 (1959).

Chapter 5

Chymotrypsin

P. Desnuelle

Laboratoire de Chimie Biologique, Faculté des Sciences, Marseille, France

Table of Contents

I. Introduction

A. The Endopeptidase Concept

Chymotrypsin and trypsin are two so closely related enzymes that it might appear desirable to study them in a single chapter. Two separate

CHAPTER 7

Metabolism of Triacylglycerols

David N. Brindley

Triacylglycerols play a major role in energy storage in animals, where they are deposited in adipose tissue. When this storage is excessive it is manifested as obesity, and there is considerable medical interest in trying to understand why some people are so prone to this condition, whereas others find it equally difficult to gain weight. In plants, the storage of triacylglycerols is best illustrated by the oil seeds, in which the triacylglycerols provide energy for growth. These seeds constitute very important commercial crops. Triacylglycerols are ideally suited to this storage function because of the highly reduced state of their fatty acids. Thus, they have high energy contents—about 37 kJ/g, compared with 17 kJ/g for protein and 16 kJ/g for carbohydrate, including glycogen, which is also used to store energy. The other advantage of the triacylglycerols is their insolubility in water, which means that they do not alter the osmotic pressure of the cell.

Most of the triacylglycerols in animals are stored in adipose tissue. However, triacylglycerols can be deposited in liver, heart, and skeletal muscle under several conditions of metabolic stress when the supply of fatty acids from adipose tissue exceeds the need or capacity of the cells to oxidize them. This formation of triacylglycerols removes the potentially toxic effects of the fatty acids and of their acyl-CoA esters, which could damage membranes and inhibit enzymes. The process of triacylglycerol synthesis also regenerates CoA. When required, the triacylglycerols are hydrolyzed by intracellular lipases so that the fatty acids can then be oxidized.

The transport of fatty acids is a major function of the triacylglycerols. Since these compounds are insoluble in the aqueous environment of the cells and of the blood and lymph, the triacylglycerol must be stabilized by association with other lipids and proteins. Such aggregates are called *lipoproteins*. There are two major classes of triacylglycerol-rich lipoproteins, namely, *chylomicrons* and *very low density lipoproteins* (VLDL), as shown in Table 7.1. Chylomicrons carry absorbed dietary fat from the small intestine to other organs, whereas VLDL mainly carry triacylglycerol from the liver to other organs.

Lipoprotein metabolism is of considerable clinical relevance. A high rate of VLDL secretion from the liver, if coupled with a relatively low rate of removal from the circulation, leads to hypertriglyceridemia. This condition in combination with a low circulating concentration of high-density lipoprotein (HDL) is considered to be a major risk factor in the development of premature atherosclerosis (see also Chapter 13). VLDL are eventually converted to low-density lipoproteins (LDL) after the triacylglycerol is removed. These LDL particles still contain much of the cholesterol that was originally used to package and stabilize the VLDL. Thus, if VLDL secretion increases, then the flux into LDL is also increased, resulting in hypercholesterolemia if this flux is not balanced by an increased rate of removal of LDL and VLDL from the circulation. An increase in the concentration of LDL in the serum, especially when accompanied by a low concentration of HDL, is thought to provide one of the best indications of an increased risk for atherosclerosis (Chapter 13).

The initial understanding of the pathways of triacylglycerol synthesis began in the early 1950s with the discovery by Kornberg and Pricer that fatty acids are activated to acyl-CoA esters before they are esterified to phosphatidate. This discovery was followed by descriptions by the groups of Kennedy and Shapiro (Brindley and Sturton

Table 7.1. Composition and Properties of Human Triacylglycerol-rich Lipoproteins

Property	*Chylomicrons*	*Very low density lipoprotein*
Density (g/mL)	0.92–0.96	0.95–1.006
Diameter (nm)	75–1000	30–75
Composition (dry wt %)		
Triacylglycerol	80–95	55–65
Phospholipids	3–8	15–20
Free cholesterol	1–3	10
Esterified cholesterol	2–4	5
Protein	1–2	9–10
Major apoproteins	B, C	B, C, E
Minor apoproteins	A	A

BIOSYNTH
PHOSPHA

Figure 7.1. Reaction mechanism of acyl-Co synthetase.

1982) of how phosphatidate is converted to various phospholipids and triacylglycerol. In the early 1960s Hübscher and Clark showed that the small intestine can synthesize triacylglycerols from monoacylglycerols and that this conversion is involved in the transport of dietary fat across the *enterocytes*, which are the absorptive cells of the small intestine (Brindley 1984). Later in that decade Hajra and Agranoff provided evidence that dihydroxyacetone phosphate can act as an alternative acyl-acceptor to glycerolphosphate for initiating the de novo synthesis of glycerolipids (Brindley and Sturton 1982).

Most of the work in this area of biochemistry is now directed to understanding how the metabolism and transport of triacylglycerols is controlled. These investigations have generally not advanced as quickly as those dealing with the metabolism of water-soluble compounds. Lipids are difficult to use as enzyme substrates, and the kinetics are very complicated. Most of the enzymes are bound tightly to membranes, which makes them relatively more difficult to purify and characterize. The rest of this chapter will attempt to provide an overview of the routes of triacylglycerol metabolism in mammals and its control as we understand it at present.

BIOSYNTHESIS OF PHOSPHATIDATE

The synthesis of triacylglycerols begins with the activation of fatty acids to acyl-CoA esters (Figure 7.1), which involves the input of energy from ATP. Acyl-CoA is then used by the various acyltransferases in the synthesis of carboxylic ester bonds. Triacylglycerols generally contain long-chain fatty acids (C_{16} or greater). Consequently, the most important activating enzyme is the long-chain acyl-CoA synthetase, which is found mainly in the endoplasmic reticulum and mitochondria of mammalian cells. However, cells also contain distinct acyl-CoA synthetases that can activate short- and medium-chain acids. These acids can then be esterified in some situations. For example, the milk of many species contains short- and medium-chain acids, and coconut fat also has a high proportion of lauric acid (C_{12}).

The main acceptor for acyl-CoA in most tissues is thought to be the *sn*-glycerol-3-phosphate that is formed by glycolysis or by the phosphorylation of glycerol. Glycerol-3-phosphate is first acylated to 1-monoacylglycerol-3-phosphate (lysophosphatidate) by glycerolphosphate acyltransferase. A different acyltransferase is responsible for esterifying the hydroxyl at the 2-position of lysophosphatidate (Figure 7.2).

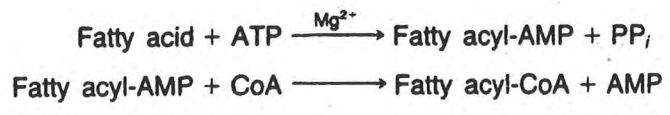


Figure 7.1. Reaction mechanism of acyl-CoA synthetase.

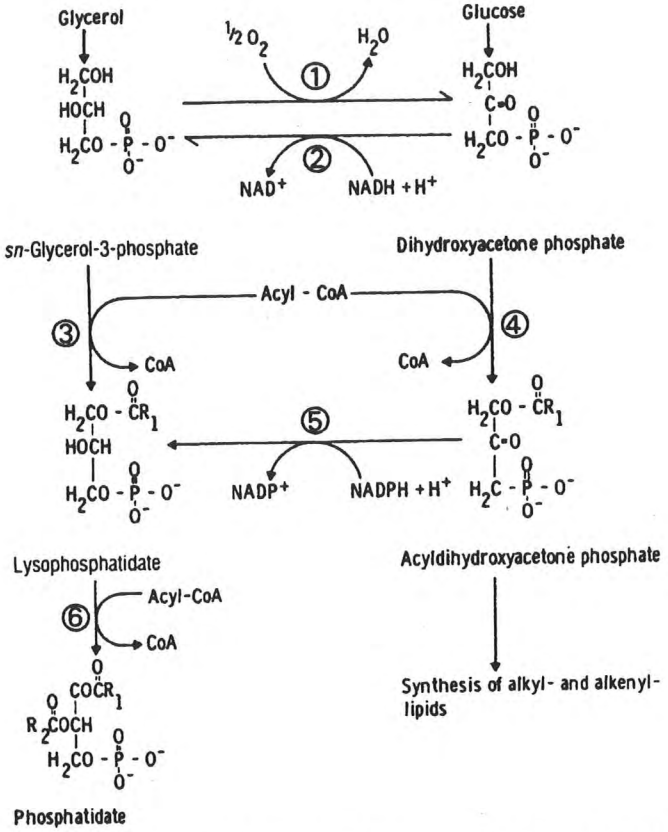


Figure 7.2. Synthesis of phosphatidate. Enzyme activities are indicated by (1) glycerol-3-phosphate dehydrogenase, (2) glycerol-3-phosphate dehydrogenase (NAD^+), (3) glycerolphosphate acyltransferase, (4) dihydroxyacetone phosphate acyltransferase, (5) acyldihydroxyacetone phosphate reductase, (6) monoacylglycerol phosphate (lysophosphatidate) acyltransferase.

In the liver, the glycerolphosphate acyltransferase is divided almost equally between the mitochondrial and microsomal fractions. By contrast, in heart, kidney, adrenal glands, and adipose tissue, the mitochondrial activity is only about 10% of the total activity, and the microsomal activity predominates (Brindley and Sturton 1982). The microsomal activity is found on the cytoplasmic side of the rough and smooth endoplasmic reticulum, and the mitochondrial enzyme is on the inner surface of the outer mitochondrial membrane. These acyltransferases are different enzymes, as is shown by the relative resistance of the mitochondrial acyltransferase to inhibition by heat, proteolytic enzymes, and *N*-ethylmaleimide (Table 7.2). The mitochondrial enzyme also has a lower K_m than the microsomal acyltransferase for acyl-CoA esters and for glycerolphosphate, and it

Table 7.2. Acyltransferases Involved in Phosphatidate Synthesis

Acyltransferase location	*Substrates*	*Effect of N-ethylmaleimide*
Endoplasmic reticulum	Glycerolphosphate or dihydroxyacetone phosphate Saturated and unsaturated acyl-CoA	Inhibition
Mitochondrial outer membrane	Glycerolphosphate Saturated acyl-CoA] low K_m	None
Peroxisomes	Dihydroxyacetone phosphate Saturated acyl-CoA	Slight stimulation

prefers saturated acyl-CoA esters, for example, palmitoyl-CoA. By contrast, the acyltransferase from the endoplasmic reticulum is able to use a variety of saturated and unsaturated acyl-CoA esters.

The esterification at the 2-position by monoacylglycerol phosphate (lysophosphatidate) acyltransferase is relatively selective for unsaturated fatty acids in both mitochondrial and microsomal fractions. However, this specificity is by no means absolute, and the fatty acid composition of the newly synthesized phosphatidate will depend upon the fatty acids available in the cell. The activity of the monoacylglycerol phosphate acyltransferase is relatively low in mitochondrial fractions, and lysophosphatidate is often the major product of esterification from glycerolphosphate. By contrast, phosphatidate is the main product of esterification with microsomal fractions.

The function of the relatively high mitochondrial glycerolphosphate acyltransferase activity, particularly in liver, is unclear. Phosphatidate is an intermediate in the synthesis of diphosphatidylglycerol (cardiolipin) (Chapter 8), which is characteristic of the inner mitochondrial membrane. However, the flux of phosphatidate into diphosphatidylglycerol is relatively low when compared with overall glycerolipid synthesis. The conversion of phosphatidate to triacylglycerol and to phosphatidylcholine is normally considered not to take place in mitochondria or to occur at very low rates.

The lower K_m values of the mitochondrial acyltransferase for glycerolphosphate and acyl-CoA esters (Table 7.2) should mean that the mitochondrial system is favored compared with the enzyme in the endoplasmic reticulum, particularly at low substrate supply. The physiological function of the mitochondrial acyltransferase is therefore uncertain and remains a major unanswered question in the control of glycerolipid synthesis. Phosphatidate can be hydrolyzed back to glycerolphosphate in a substrate cycle (Figure 7.3), which

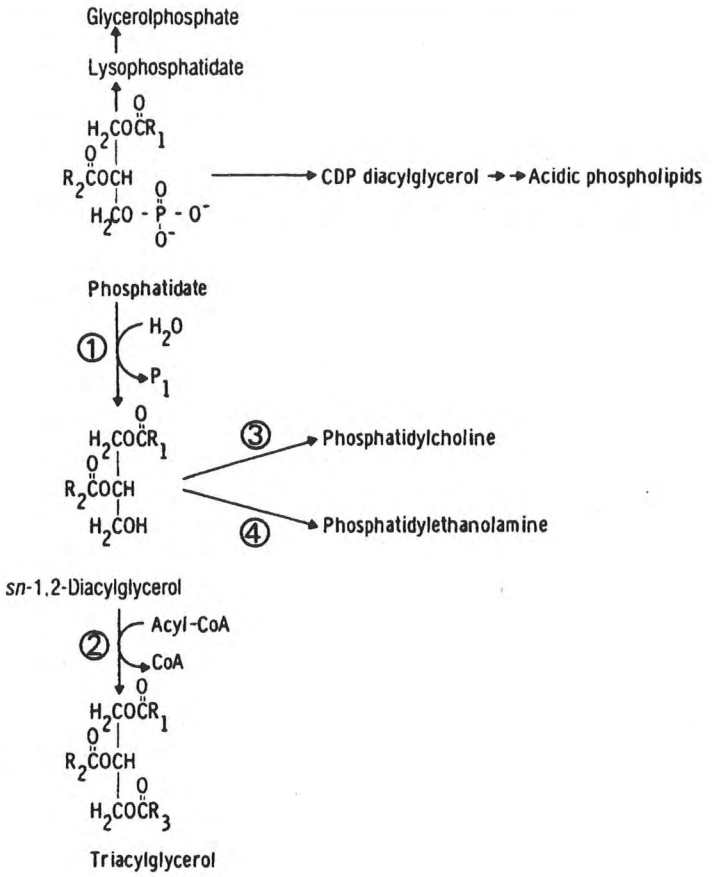


Figure 7.3. Metabolism of phosphatidate. Enzyme activities are indicated by (1) phosphatidate phosphohydrolase, (2) diacylglycerol acyltransferase, (3) choline phosphotransferase, (4) ethanolamine phosphotransferase.

might help in regulating the balance between β-oxidation and esterification (see later discussion). Alternatively, the phosphatidate might be transferred to the endoplasmic reticulum for the synthesis of triacylglycerol and phospholipids.

The other route for the de novo biosynthesis of phosphatidate is by the esterification of dihydroxyacetone phosphate (Figure 7.2). This reaction can be catalyzed in the endoplasmic reticulum by glycerolphosphate acyltransferase, which can use glycerolphosphate and dihydroxyacetone phosphate, which compete for the enzyme (Table 7.2). In addition, there is a dihydroxyacetone phosphate acyltransferase that neither uses glycerolphosphate as a substrate, nor is inhibited by it. The specific enzyme is not inhibited by *N*-ethylmaleimide, which further indicates that it is a different enzyme from the microsomal glycerolphosphate acyltransferase. The specific dihydroxyacetone phosphate acyltransferase is located mainly in peroxisomes. It has a higher reaction rate with the CoA esters of saturated than with unsaturated fatty acids; this specificity

therefore favors the incorporation of saturated fatty acids into the 1-position of glycerolipids.

Acyldihydroxyacetone phosphate can undergo two different biosynthetic modifications (Figure 7.2). First, it can be converted into alkyldihydroxyacetone phosphate, which can then serve as the precursor for the synthesis of the alkyl- and alkenyl-lipids (Chapter 9). Alternatively, the acyldihydroxyacetone phosphate can be converted to 1-monoacylglycerol-3-phosphate (lysophosphatidate) by a reductase that uses NADPH. This reductase is found in both peroxisomes and the endoplasmic reticulum, and the same reductase probably acts on alkyldihydroxyacetone phosphate.

It is generally accepted that acyldihydroxyacetone phosphate is an obligatory intermediate in the synthesis of alkyl- and alkenyl-lipids. However, the quantitative importance of the esterification of dihydroxyacetone phosphate relative to glycerolphosphate for the synthesis of glycerolipids is very much in dispute. On the one hand it has been argued that the esterification of glycerolphosphate must predominate. This conclusion takes into account the relatively high concentration of glycerolphosphate in the cell and the K_m, K_i, and V_{max} values for glycerolphosphate and dihydroxyacetone phosphate for the glycerolphosphate acyltransferase that is located in the endoplasmic reticulum (Brindley and Sturton 1982). These calculations do not consider the specific acyltransferases that exist in the mitochondria and peroxisomes. However, it is not clear whether these organelles can efficiently donate phosphatidate for the synthesis of triacylglycerols and the zwitterionic phospholipids.

Other arguments are based upon the use of glycerol labeled either with ^{3}H at the 2-position or with ^{14}C. Conversion of glycerolphosphate to dihydroxyacetone phosphate before its incorporation causes the loss of ^{3}H, whereas ^{14}C is retained. The relative activities of the two pathways can be estimated by comparing the $^{3}H/^{14}C$ ratio in lipid with that in the glycerolphosphate. When this was done, 50–75% of the glycerolipid synthesized by rat liver slices appeared to be derived by acylation of dihydroxyacetone phosphate. A similar technique showed that about 56% of the phosphatidylglycerol and 64% of the phosphatidylcholine were formed from dihydroxyacetone phosphate in type II cells of lung (Brindley and Sturton 1982).

An alternative approach was to incubate cells with a mixture of D-[U-^{14}C]- and D-[3-^{3}H]-glucose and to measure the labeling of the lipids. This procedure generated [4-^{3}H]NADPH, which was incorporated into the 2-position of glycerolipids by the dihydroxyacetone phosphate pathway (Figure 7.2). It was then possible to calculate the relative contribution of the two pathways to the synthesis of glycerolipids on the assumption that the $^{3}H/^{14}C$ ratio in alkyl- and alkenyl-lipids results entirely from synthesis by the dihydroxyacetone phosphate pathway. A correction factor also had to be applied to

compensate for some labeling of glycerolphosphate at the 2-position. From these results it was estimated that 49–61% of the glycerolipid synthesised by BHK-21 and BHK-Ts-a/1b-2 cells occurs by acylation of dihydroxyacetone phosphate.

It is also possible to postulate that the dihydroxyacetone phosphate pathway should be important in glycerolipid synthesis from theoretical considerations of cofactor requirements. Most anabolic pathways use NADPH as a cofactor in reductive synthesis. The formation of glycerolipids from glucose via the dihydroxyacetone phosphate pathway fulfills this requirement (Figure 7.2). NADPH production increases when there is active fatty acid synthesis. This increase could also favor the incorporation of the fatty acids into triacylglycerols by the dihydroxyacetone phosphate pathway. The conversion of dihydroxyacetone phosphate to glycerolphosphate requires NADH, which is normally involved in reductive degradation. In the liver, glycerolphosphate dehydrogenase, which catalyzes this reaction, is also involved in gluconeogenesis from glycerol and in maintaining the redox state.

We do not know the relative importance of the three acyltransferases that are described in Figure 7.2 in initiating glycerolipid synthesis or whether their relative contributions change in different physiological conditions. However, it is normally assumed that the synthesis of triacylglycerols takes place mainly in the endoplasmic reticulum by the esterification of glycerolphosphate.

CONVERSION OF PHOSPHATIDATE TO TRIACYLGLYCEROL

Phosphatidate lies at a branch point in glycerolipid synthesis which provides three routes for further metabolism (Figure 7.3). First, it can be hydrolyzed to glycerolphosphate in a substrate cycle by phospholipases of the A-type. This hydrolysis probably prevents the excessive accumulation of phosphatidate in the membranes. Second, it can be converted to CDP-diacylglycerol, which provides a precursor for the synthesis of the acidic phospholipids, phosphatidylinositol, phosphatidylglycerol, and diphosphatidylglycerol (Chapter 8). Third, it can be hydrolyzed to diacylglycerol, which in turn can serve as a common precursor in the synthesis of the zwitterionic phospholipids, phosphatidylcholine, and phosphatidylethanolamine (Chapter 8) and also for the synthesis of triacylglycerol. We shall now discuss the last of these routes of metabolism.

The conversion of phosphatidate to diacylglycerol is catalyzed by phosphatidate phosphohydrolase. This enzyme activity has been reported to occur in plasma membranes, lysosomes, mitochondria, the endoplasmic reticulum, and the cytosol. However, it is difficult to be certain about the distribution of the phosphohydrolase, since the assays used in many of these studies measured the release of

inorganic phosphate or water-soluble phosphates from phosphatidate (Brindley and Sturton 1982). These phosphates can also arise by deacylation of the phosphatidate by phospholipase A activities to form glycerolphosphate, which can be further hydrolyzed by acid or alkaline phosphatases. The occurrence of a phosphatidate phosphohydrolase has been confirmed in the cytosol and endoplasmic reticulum in several cell types by specifically measuring the production of diacylglycerol. This activity is the one that is thought to be important for the synthesis of triacylglycerol, phosphatidylcholine, and phosphatidylethanolamine. Recent evidence shows that phosphatidate phosphohydrolase can be translocated between the cytosol and the endoplasmic reticulum (Brindley 1985). This phenomenon will be discussed later in terms of the control of glycerolipid synthesis.

This phosphatidate phosphohydrolase activity can be stimulated by Mg^{2+}. However, this is probably not a cation requirement in the normal sense of being essential for the binding of the phosphate group to the enzyme, since amphiphilic (detergentlike) amines such as chlorpromazine can substitute for Mg^{2+} in the presence of high concentrations of EDTA. The function of the Mg^{2+} is probably to adjust the surface charge and packing arrangement of the membrane that contains the phosphatidate, thus facilitating the interaction of substrate with the phosphohydrolase.

The final stage of triacylglycerol synthesis is completed by the action of diacylglycerol acyltransferase, which forms the final ester bond at the 3-position (Figure 7.3). This enzyme is located in the endoplasmic reticulum, and its activity can be stimulated by Mg^{2+}. As far as we know, this enzyme competes for a common pool of diacylglycerol with the choline- and ethanolamine-phosphotransferases that are responsible for the synthesis of phosphatidylcholine and phosphatidylethanolamine.

The pathways that have been described so far are the possible routes for the de novo synthesis of triacylglycerols that occurs in a wide variety of cell types. However, there is another route for triacylglycerol synthesis that operates in the enterocytes of the small intestine. This pathway will be discussed in relation to its role in lipid absorption and transport.

DIGESTION, ABSORPTION, AND TRANSPORT OF LIPIDS

Lipids constitute an important part of the diet, and in the more developed countries contribute 40–45% of energy intake. The largest proportion of this energy comes from triacylglycerols. This means that a person eats 60–130 g of fat each day and that the body has to digest, absorb, resynthesize and transport this quantity of lipid. Approximately 90–95% of the triacylglycerols from the diet are

absorbed from the intestinal lumen, whereas only about 50% of the cholesterol is absorbed. Since some of the cholesterol in the intestinal lumen comes from the bile, the incomplete absorption of cholesterol enables the body to excrete it.

Digestion of Lipids

Relatively little digestion of triacylglycerols takes place in the stomach, but the action of the proteolytic enzymes releases lipid from food particles, and a coarse emulsion is formed by the churning action of the stomach. A gastric lipase that is distinct from pancreatic lipase has been identified in the stomach contents of several mammals, including man. This lipase is active at pH 3–4, and it preferentially releases short- and medium-chain-length fatty acids from acylglycerols. These fatty acids are found in the triacylglycerols of the milk of several species, and they are preferentially located at the 3-position (Figure 7.4). The partial hydrolysis of triacylglycerols to 1,2-diacylglycerols makes the milk fat micelles more susceptible to the subsequent action of pancreatic lipase in the small intestine. The short- and medium-chain fatty acids can be absorbed directly by the gastric mucosa. They are bound to albumin in the blood and carried to the liver, where they can be readily oxidized to produce energy.

The remaining lipid enters the proximal part of the small intestine, where the pH rises to 5.8–6.5. The coarse droplets of oil become

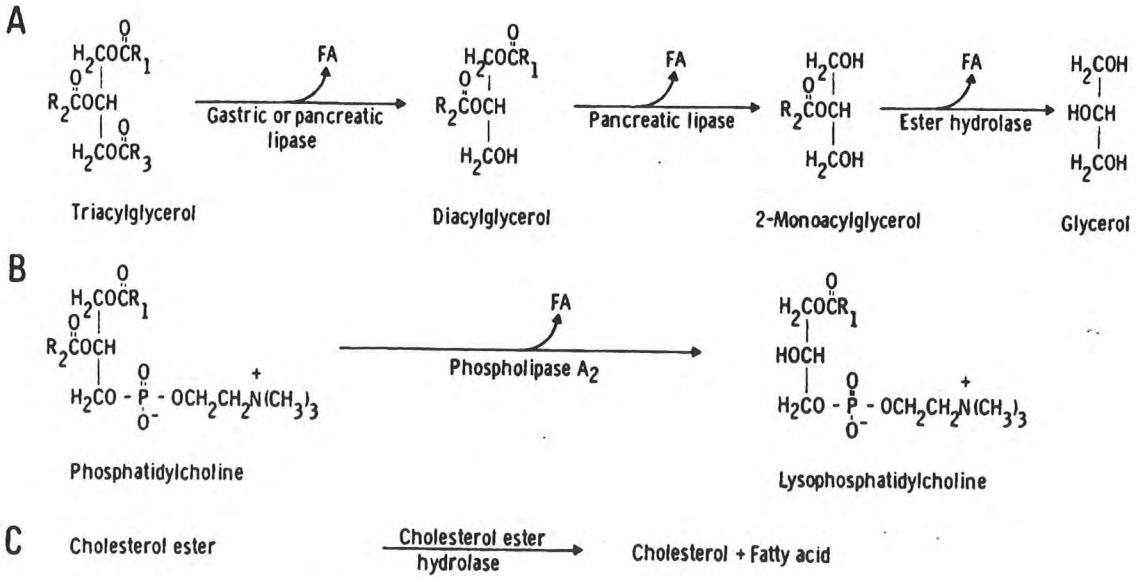


Figure 7.4. Reactions in lipid digestion in the gastrointestinal tract.

Figure 7.5. S
representatic
digestion and
in the small in

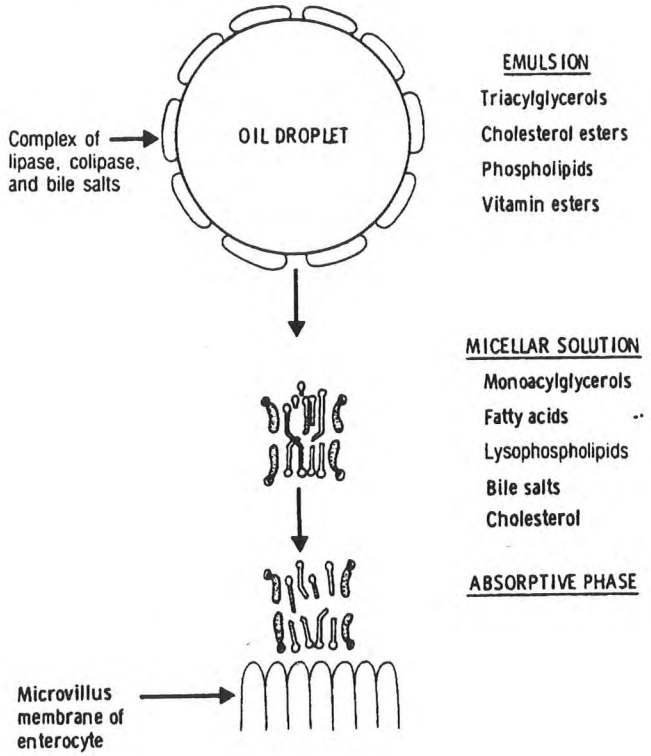


Figure 7.5. Schematic representation of lipid digestion and absorption in the small intestine.

coated with bile salts that are secreted by the liver. These compounds are amphiphilic, and they dissolve at the oil-water interface with their hydrophobic faces pointing into the oil and the hydrophilic surfaces pointing into the aqueous phase of the luminal contents. The bile salts help to disperse the oil droplets, and they donate a negative charge to their surfaces. A protein called colipase (molecular weight of about 10,000), which is secreted by the pancreas, is then adsorbed onto the surface of the oil droplets (Figure 7.5). This protein acts as an anchor for the attachment of pancreatic lipase at the oil-water interface. The bile salts, colipase, and lipase are thought to interact in a ternary complex which in the presence of Ca^{2+} is able to hydrolyze the 1- and 3-ester bonds of triacylglycerol (Figure 7.4A). About 85% of the digestion of dietary triacylglycerol in nonruminant animals ends with the formation of 2-monoacylglycerol, since pancreatic lipase is unable to hydrolyze the ester bond at the 2-position.

Limited hydrolysis of 2-monoacylglycerols does take place in nonruminant animals, but it is catalyzed by a nonspecific ester hydrolase that is secreted by the pancreas. The same enzyme probably also hydrolyzes cholesterol esters (Figure 7.4C) and the esters of fat-soluble vitamins.

In ruminants the complete digestion of triacylglycerols is effected by lipases that are produced by bacteria in the rumen. Bacteria are also responsible for hydrogenating unsaturated fatty acids that are present in the diets of ruminant animals, which accounts for the highly saturated nature of their body fats.

Pancreatic juice also contains phospholipases of the A_1- and A_2-types, which remove fatty acids from the 1- and 2-positions of phospholipids, respectively. These phospholipids can be derived either from the diet or from bile delivered to the small intestine. The digestion of phospholipids need not be complete: 30–40% of the dietary phosphatidylcholine can be absorbed as lysophosphatidylcholine (Figures 7.4B and 7.5).

The process of digestion converts lipids that have limited abilities to interact with water into more polar compounds with amphiphilic properties. Thus, triacylglycerols are transformed into 2-monoacylglycerols; cholesterol esters into cholesterol; and phospholipids into their lyso-derivatives, which are strong detergents. The fatty acids that are released are ionized at the pH that prevails in the intestinal lumen, so these compounds are also surface active, that is, they can interact with both lipid and aqueous environments (Brindley 1984).

As digestion proceeds, the monoacylglycerols, fatty acids, lysophospholipids, cholesterol, and bile salts dissociate from the surface of the oil droplets to form a micellar solution (Figure 7.5). The micelles are aggregates of amphiphilic lipids that orient themselves with the hydrophobic regions on the inside of the micelles and the polar groups exposed to the aqueous environment (Chapter 2). Further digestion of phospholipids and cholesterol esters can take place at the microvillus membrane of the enterocytes, which contains phospholipases and a cholesterol ester hydrolase.

Absorption of Lipids from the Intestines

Most of the lipids are absorbed from the small intestine (in the jejunum) with the exception of the bile salts, which remain in the lumen of the small intestine to facilitate further digestion. The bile salts are finally absorbed in the distal ileum and are transported back to the liver by the portal blood. This cycle constitutes the *enterohepatic circulation*.

Lipid absorption occurs when the micellar solution of lipids comes into contact with the microvillus membrane of the enterocytes (Figure 7.5). The lipids are transported across the membrane by an energy-independent process which relies on the maintenance of an inward diffusion gradient. This gradient can partly be achieved by the attachment of the fatty acids to a binding protein (Z-protein) with a

Figure 7.6. M
for the resyn
triacylglycero
phosphatidyl
the enterocyt
small intestin

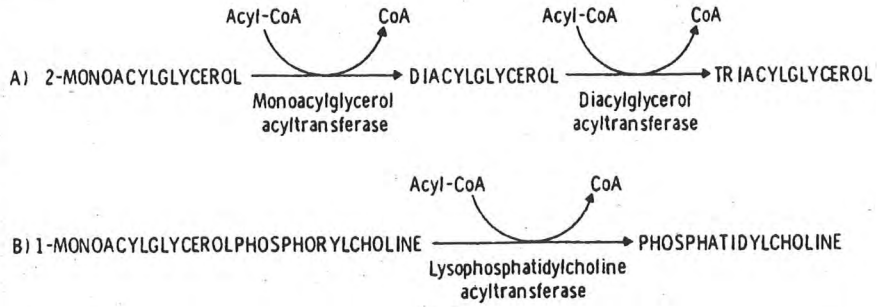


Figure 7.6. Major routes for the resynthesis of triacylglycerol and phosphatidylcholine in the enterocytes of the small intestine.

molecular weight of about 12,000. However, the main driving force for absorption comes from the rapid re-esterification of the lipids, which is an ATP-dependent process due to activation of fatty acids (Figure 7.1).

In the nonruminant animal the major acyl-acceptor is the 2-monoacylglycerol that is formed in the lumen of the intestine during the partial hydrolysis of triacylglycerol. Triacylglycerol is resynthesized by the sequential actions of monoacylglycerol acyltransferase and diacylglycerol acyltransferase (Figure 7.6A). This pathway is estimated to account for 75–85% of the total synthesis of triacylglycerol in the enterocytes of nonruminant animals. The remainder comes from the esterification of glycerolphosphate (and perhaps of dihydroxyacetone phosphate), as described earlier (Figures 7.2 and 7.3). Synthesis of triacylglycerol from glycerolphosphate is necessary, since some of the dietary triacylglycerol is completely hydrolyzed to glycerol. In ruminant animals, bacterial lipases are responsible for the hydrolysis of 2-monoacylglycerols; therefore the pathways for de novo synthesis of the triacylglycerol are required.

The diacylglycerol that is formed as an intermediate in triacylglycerol synthesis can also be used to synthesize phosphatidylcholine. However, the most prominent source of phosphatidylcholine in nonruminant animals is the esterification of the 1-monoacylglycerophosphorylcholine (lysophosphatidylcholine) (Figure 7.6B) that is formed by the partial digestion of phosphatidylcholine. The reactions of glycerolipid synthesis are very rapid—they take place within seconds after the precursors are absorbed.

The transport and metabolism of the absorbed cholesterol is much slower than that of triacylglycerols, and it is estimated that the $t_{\frac{1}{2}}$ for cholesterol in the enterocyte is about 12 h. During absorption the cholesterol becomes incorporated into the membranes of the enterocytes and diluted with existing cholesterol. A large proportion of the cholesterol that is transported from the enterocyte is esterified, mainly with oleic acid.

Formation of Chylomicrons and VLDL

The transport of lipids from the small intestine requires that they be packaged into a physically stable form that can exist in an aqueous environment. Thus, the most hydrophobic of the lipids (cholesterol esters and triacylglycerols) are coated with a layer of amphiphilic compounds including phosphatidylcholine, cholesterol, and various apoproteins. Apart from stabilizing the surface of the lipoproteins, the apoproteins also provide "address labels" that govern which cells in the body receive and metabolize these lipoproteins (Chapter 13). The main lipoproteins responsible for transporting dietary fat from the intestine are the chylomicrons. However, the intestine is also able to secrete VLDL, and it is responsible for about 10% of this lipoprotein that appears in the blood. The chylomicrons are bigger than the VLDL, and they have a higher content of triacylglycerol (Table 7.1). However, the size and composition of the lipoproteins can vary according to the composition of the diet, the rates of lipid absorption, and apoprotein synthesis. Thus, with a high intake of dietary fat and at the peak of absorption, the chylomicrons tend to be larger and contain more triacylglycerol than when the rate of lipid transport is lower.

The assembly of the chylomicrons and VLDL begins with the resynthesis of triacylglycerols, which takes place in the smooth endoplasmic reticulum in the apical region of the enterocyte. The triacylglycerols appear in the form of lipid droplets in the cisternae of the smooth endoplasmic reticulum within minutes of exposure of the cells to micellar lipid in the lumen (Figure 7.7). The droplets are stabilized with a coat of phospholipid and with proteins that are synthesized in the rough endoplasmic reticulum. As time progresses, the droplets increase in number as the endoplasmic reticulum extends and pinches off to form vesicles. It is thought that these lipid-filled vesicles fuse with the Golgi apparatus. The nascent chylomicrons and VLDL are then carried to the lateral surfaces of the enterocyte by the process of *exocytosis*. Fusion of the Golgi and surface membranes takes place, and the chylomicrons and VLDL are secreted into the intercellular spaces which drain into the lymph vessels (Figure 7.7). The lipoproteins in these vessels pass via the thoracic duct and enter the circulation at the level of the jugular vein.

The hydrolysis of the triacylglycerols of chylomicrons and VLDL involves their binding through the apo C-2 on their surfaces to an enzyme called lipoprotein lipase, which is found in the capillary beds of various extrahepatic tissues including skeletal muscle, cardiac muscle, and adipose tissue. The enzyme is anchored to polysaccharide chains on the endothelial wall of the capillaries (Figure 7.8). Lipoprotein lipase is synthesized by the cells that underlie the capillaries. The cells receive the fatty acids that are hydrolyzed from triacylglycerols

Figure 7.8. Schem
representation o
of a chylomicron
lipoprotein lipase.
is reproduced fron
(1984) by the per
Butterworth's Sc

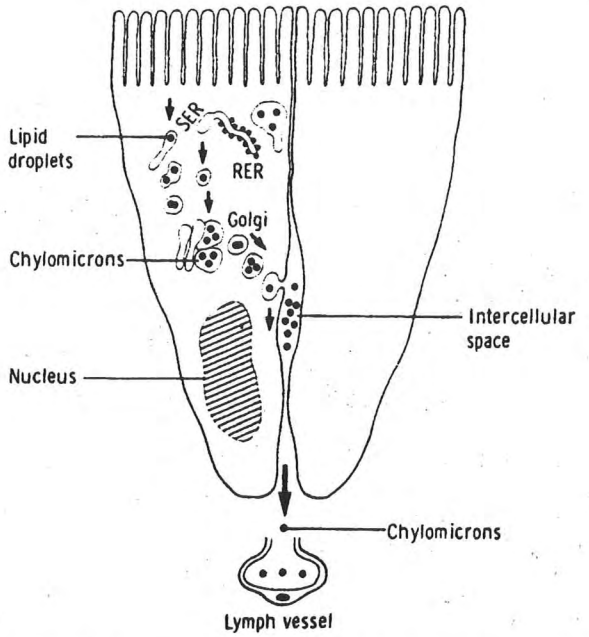


Figure 7.7. Schematic representation of the transport of lipids through the enterocytes and the formation of chylomicrons, where SER represents smooth endoplasmic reticulum and RER represents rough endoplasmic reticulum. The figure is reproduced from Brindley (1984) by the permission of Butterworth's Scientific Ltd.

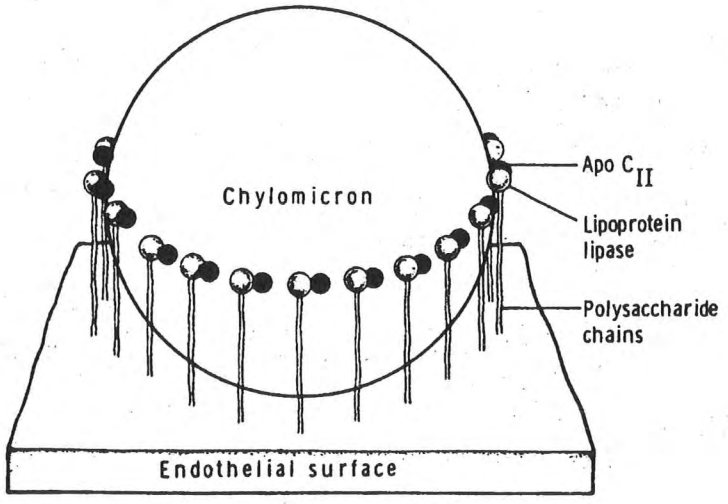


Figure 7.8. Schematic representation of the binding of a chylomicron to lipoprotein lipase. The figure is reproduced from Brindley (1984) by the permission of Butterworth's Scientific Ltd.

by the lipase (Chapter 13). The body is able to direct chylomicrons and VLDL to particular organs through its tissue-specific control of lipoprotein lipase activity. This subject will be considered in more detail later in this chapter.

The metabolic fate of fatty acids depends on the particular cells into which they are delivered. Fatty acids are readily used for energy production by muscle tissues. If the uptake exceeds the immediate requirement for β-oxidation, then the fatty acids are temporarily stored as triacylglycerols within the muscle cells. Storage is, of course, the major fate of fatty acids that enter adipocytes.

Partitioning of Fatty Acids Between the Portal Blood and the Lymphatic System

The description so far has dealt with the metabolism and transport of glycerolipids and cholesterol. Most of the long-chain fatty acids are transported from the intestine by incorporation into the esterified lipids of the chylomicrons or VLDL. These two lipoproteins also transport cholesterol, the fat-soluble vitamins, and other hydrophobic compounds and drugs. However, fatty acids with 10- to 12-carbon chain lengths are found both in the triacylglycerols of chylomicrons and in their unesterified forms bound to albumin in portal blood (Table 7.3). Short-chain-length fatty acids are transported almost exclusively by the latter route to the liver.

There are a number of reasons for the partitioning of fatty acids between portal blood and the fluid in the lymphatics on the basis of their chain lengths (Table 7.3). First, the short- and medium-chain acids which are often located at the 3-position are more readily hydrolyzed from triacylglycerols. This means that they are unlikely to be retained in the monoacylglycerols that are re-esterified in the enterocytes. The enterocytes also contain ester hydrolases that selectively degrade short- and medium-chain acylglycerols. Second, the fatty acids in the intestinal lumen partition between the micellar and aqueous phases. The shorter the chain length, the greater the tendency to partition into the aqueous phase and to become separated from the bulk of the lipid. The uptake of short-chain fatty acids into the enterocytes can take place against a concentration gradient. Finally, the enzymes responsible for re-esterification in the entero-

Table 7.3. Partition of Fatty Acids after Absorption Between the Portal Blood and the Lymphatic System

Chain length	*Portal route (unesterified)*	*Lymphatics (esterified)*
$C_2 - C_8$	Majority	Little
C_{10}	Majority	Significant
C_{12}	Significant	Significant
$>C_{12}$	Little	Majority

cytes discriminate against the short- and medium-chain-length fatty acids.

Use is made of the body's ability to absorb fatty acids by the portal route in conditions where the digestion, absorption, or transport of triacylglycerols in chylomicrons is impaired. The clinical management of these malabsorption syndromes (for example, pancreatic or biliary insufficiency, damage to the intestinal mucosa, abetalipoproteinemia) is facilitated by feeding medium-chain triacylglycerols, from which the fatty acids can be efficiently absorbed by the portal route. This energy-rich diet supplies the liver with a readily oxidizable substrate.

CONTROL OF TRIACYLGLYCEROL SYNTHESIS

The control of triacylglycerol synthesis varies from tissue to tissue. In the enterocytes of the small intestine, it is assumed that the load of dietary fatty acids is the major factor that controls the rate of triacylglycerol synthesis. The availability of 2-monoacylglycerol governs the balance between the monoacylglycerol pathway and the pathways of de novo synthesis. The major function of the enterocyte in this respect is to transport dietary fats and to synthesize chylomicrons.

In the liver the situation is more complex. This organ is required to incorporate into triacylglycerols fatty acids when they are derived primarily from de novo synthesis, from the diet, or from lipolysis in adipose tissue (Figure 7.9). In terms of metabolic and hormonal

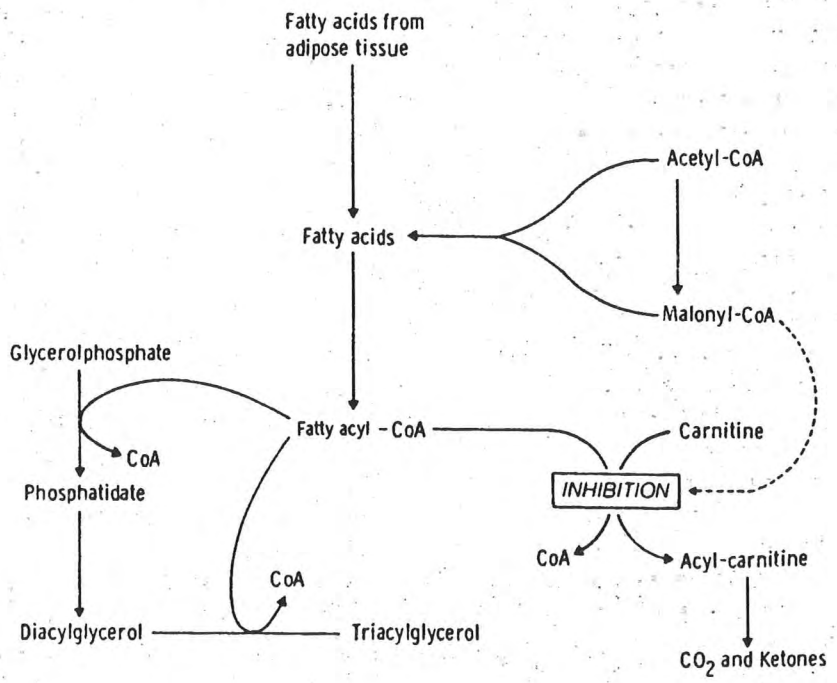


Figure 7.9. The relationship between fatty acid synthesis, oxidation, and esterification in the liver.

control, these represent quite different conditions. In addition to triacylglycerol synthesis, fatty acids can be used for β-oxidation and for the synthesis of phospholipids. These lipids are needed for membrane turnover and for the secretion of bile, VLDL, and HDL components.

The major function of triacylglycerol synthesis in heart and skeletal muscle is to enable fatty acids to be stored temporarily if their supply is greater than the rate of β-oxidation.

Long-term storage of triacylglycerols occurs in adipose tissue. In adipocytes the activities of the enzymes that synthesize triacylglycerols are controlled reciprocally with that of the hormone-sensitive lipase, which is responsible for fatty acid mobilization (Saggerson 1985). However, because of the high fatty acid availability during lipolysis, some of these can be recycled back to triacylglycerol.

Control of Phosphatidate Synthesis in the Liver

One of the obvious sites for the enzymic regulation of fatty acid esterification is at the level of the glycerolphosphate and dihydroxyacetone phosphate acyltransferases, since it is here that fatty acids become committed to glycerolipid synthesis. The maximum velocity of acyl-CoA synthetase is very much higher than those of the acyltransferases, but the flux into acyl-CoA esters is, of course, dependent upon fatty acid availability. At present there is no evidence for separate pools of acyl-CoA esters that selectively supply one pathway of fatty acid metabolism. Although fatty acids can be desaturated and elongated, the major competition to glycerolipid synthesis is provided by β-oxidation. Regulation of the relative activites of glycerolphosphate acyltransferase and carnitine palmitoyltransferase should therefore control this branch point (Figure 7.9). The competition for acyl-CoA esters might explain why there is such an active glycerolphosphate acyltransferase with low K_m values for glycerolphosphate and acyl-CoA esters in the mitochondria. In addition, there might be competition between the peroxisomal β-oxidation system and the specific dihydroxyacetone phosphate acyltransferase that is present in this organelle. However, as explained before, the metabolic fate of the majority of the phosphatidate and lysophosphatidate that can potentially be formed in mitochondria and peroxisomes respectively remains unclear.

We do know that carnitine palmitoyltransferase can be acutely regulated by inhibition with malonyl-CoA (Chapter 4). At present no such regulator of the acyltransferases has been found. Consequently, it is tempting to assume that the esterification system acquires those fatty acids not required by carnitine palmitoyltransferase. It might be wise, however, to reflect that the discovery of the malonyl-CoA inhibition of this enzyme is relatively recent. Previously, the prevail-

ing view was that fatty acid oxidation responded to fatty acid flux without being regulated itself.

The mitochondrial glycerolphosphate acyltransferase is thought to be increased by insulin; it is decreased more in starvation and in diabetes than the acyltransferase of the endoplasmic reticulum. The latter enzyme should provide most of the phosphatidate that is used for the synthesis of triacylglycerols. Theoretically, the acyltransferase could operate at a high rate of fatty acid flux, and it presumably provides the liver with its large capacity for triacylglycerol synthesis. Its activity changes relatively little in different physiological conditions, although small decreases have been observed during starvation. The physiological significance of such changes is not clear, since the capacity of the liver to synthesize triacylglycerols is not decreased in this condition. This fact can be demonstrated by blocking β-oxidation, which diverts an increased flux of fatty acids into triacylglycerol synthesis.

Control of the Conversion of Phosphatidate to Triacylglycerol in Liver

The activity of phosphatidate phosphohydrolase in the liver changes much more dramatically than activity of other enzymes of triacylglycerol synthesis. The synthesis of phosphatidate phosphohydrolase is stimulated over a period of hours by glucocorticoids (for example, cortisol or corticosterone) and high concentrations of cAMP. It is also known that high concentrations of glucocorticoids in vivo increase hepatic triacylglycerol synthesis, produce a fatty liver, and stimulate VLDL secretion (Brindley and Sturton 1982).

The effects of glucocorticoids and cAMP on the synthesis of phosphatidate phosphohydrolase are antagonized by insulin. These hormonal interactions account for the high activities of this enzyme that are observed in the liver during conditions when metabolism is controlled to a greater extent by the stress hormones (adrenalin, glucagon, glucocorticoids) than by insulin (Table 7.4). In addition to these long-term effects (of glucocorticoids, cAMP, and insulin), vasopressin (and probably other hormones that mobilize intracellular Ca^{2+}, including angiotensin II and α_1-agonists) can rapidly stimulate phosphatidate phosphohydrolase activity. The action of vasopressin is also accompanied by an increase in triacylglycerol synthesis (Brindley 1985).

The increased phosphohydrolase activities that are illustrated in Table 7.4 provide the liver with an increased potential to synthesize triacylglycerol, which protects the liver against the large influxes of fatty acids that can be caused by high rates of fatty acid mobilization from adipose tissue. Phosphatidate phosphohydrolase exists both in the cytosol and on the endoplasmic reticulum. It is believed that the cytosolic form of the enzyme is physiologically inactive until it

Table 7.4. Changes in Phosphatidate Phosphohydrolase Activity in the Liver That Can Be Associated with an Increased Effect of Glucocorticoids and Other Stress Hormones Relative to Insulin

Treatment or condition	*Duration*	*Increase in phosphohydrolase activity*
Metabolic stress		
Starvation	6–40 h	1.3- to 2.3-fold
Surgical stress	6 h	3-fold
Subtotal hepatectomy	6 h	5.5-fold
Diabetes (mildly ketotic)	10 weeks	1.4-fold
Diabetes (ketotic)	48 h	2.8-fold but reversed by insulin
Hypoxia	24 h	2.5-fold
Toxic conditions produced by		
Hydrazine	4–24 h	2- to 4-fold
Morphine (mouse)	60–180 min	2-fold
Hormone injections		
Corticotropin	6 h	3.3- to 4.3-fold
Cortisol	5 days	2.4-fold
Genetic obesity (ob/ob mouse)	Long-term	2-fold
Ingestion of some nutrients		
Fructose, sorbitol, or glycerol	6 h	1.9- to 2.3-fold
Ethanol	7 h	6.9-fold

Source: Brindley and Sturton (1982);
Note: The results are from work with rats with the exception of morphine injections and the ob/ob mouse.

translocates onto the membranes on which phosphatidate is being synthesized. The signal for this translocation is the increased availability of fatty acids, which also stimulates the phosphohydrolase activity (Figure 7.10) (Brindley 1985).

The ability of fatty acids to promote the translocation of phosphatidate phosphohydrolase from the cytosol to the membrane-associated compartment appears to be under hormonal control. cAMP displaces the enzyme from the membranes, but this effect can be overcome by higher concentrations of fatty acids (Table 7.5). By contrast, insulin should decrease the intracellular concentration of cAMP and therefore the concentration of fatty acids required to promote the translocation. The mechanisms that cause these changes are not yet established, but they may involve the reversible phosphorylation of the phosphohydrolase. This type of control for phosphatidate phosphohydrolase is reminiscent of the regulation of CTP:phosphocholine cytidylyltransferase, which regulates the synthesis of phosphatidylcholine (Chapter 8). Phosphatidate phosphohydrolase and CTP:phosphocholine cytidylyltransferase can be classified as *ambiquitous*, which means that they exist in different locations of the cell and can regulate metabolism by moving from one location to another

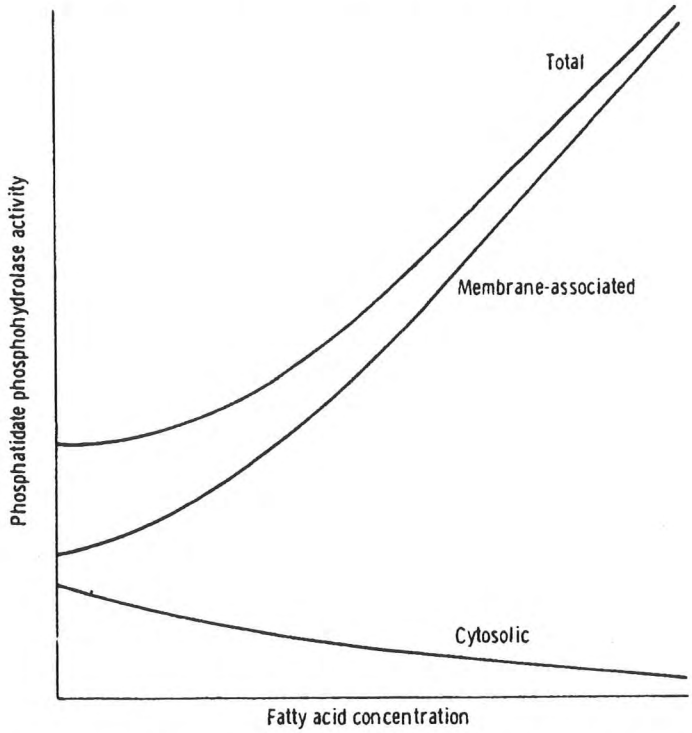


Figure 7.10. Activation and translocation of phosphatidate phosphohydrolase by fatty acids in isolated hepatocytes. The figure shows that a fatty acid (oleate) increases the total activity of phosphatidate phosphohydrolase and also increases its association with membranes rather than with the cytosol. It is believed that the membrane-associated enzyme is physiologically active and that the cytosolic enzyme forms a reservoir of potential activity.

Table 7.5. Effects of a cAMP Analogue and Oleate on the Location of Phosphatidate Phosphohydrolase in Isolated Hepatocytes

	Phosphohydrolase activity (%)	
Additions	*Cytosolic*	*Membrane-associated*
None	68	32
*CPT-cAMP (0.5*mM*)	86	14
Oleate (1*mM*)	48	52
CPT-cAMP (0.5*mM*) + Oleate (1*mM*)	42	60

* 8-(4-Chlorophenylthio)adenosine-3',5'-cyclic monophosphate, a potent analogue of cAMP.

(Wilson 1980). The reason for coordinating the synthesis of phosphatidylcholine and triacylglycerol can readily be appreciated in terms of the synthesis of VLDL, which contain 20% and 60% (by weight), respectively, of these lipids (Table 7.1).

Diacylglycerol is a Precursor of Triacylglycerol, Phosphatidylcholine, and Phosphatidylethanolamine

The synthesis of phosphatidylcholine and phosphatidylethanolamine competes for the diacylglycerol that is produced by phosphatidate phosphohydrolase (Figure 7.3). The choline- and ethanolamine-phosphotransferases have higher affinities for diacylglycerol than does diacylglycerol acyltransferase. Therefore, at low fatty acid availability and a low rate of formation of diacylglycerol, the major flux from diacylglycerol is directed to phospholipid synthesis. This maintains membrane turnover and bile secretion.

It has been suggested that the activity of diacylglycerol acyltransferase could be decreased by a cAMP-dependent phosphorylation that would exaggerate this discrimination against triacylglycerol synthesis and favor phospholipid formation in conditions of stress (Brindley and Sturton 1982). If this were to occur it would be overcome by increased fatty acid availability, which would supply acyl-CoA for the acyltransferase. Thus, as the supply of fatty acid increases, the relative proportion of diacylglycerol converted to triacylglycerol would increase. The synthesis of phosphatidylcholine and phosphatidylethanolamine might eventually be limited by the availability of CDP-choline and CDP-ethanolamine; however, the abundance of fatty acids, would also stimulate the cytidylyltransferase, which would attempt to maintain CDP-choline synthesis (Chapter 8) provided that the phosphocholine was available.

In order to try to relate these changes in hepatic triacylglycerol synthesis to other metabolic pathways and to metabolism in other organs, two very different metabolic situations will be considered, namely, the conditions of high and low availabilities of insulin relative to glucagon, adrenalin, corticotropin, and glucocorticoids.

METABOLISM OF TRIACYLGLYCEROLS WHEN THE AVAILABILITY OF INSULIN IS HIGH

A high concentration of insulin is normally seen after the consumption of meals rich in glucose in the form of starch. Such a diet also produces good insulin control, since the tissues normally respond well to insulin action. If insulin predominates relative to glucagon, the catecholamines, and corticotropin in the control of metabolism, then glycolysis, and the synthesis of glycogen and fatty acid in the liver, are stimulated (Chapter 5). At the same time, β-oxidation and gluconeogenesis are suppressed (Figures 7.9 and 7.11). Glucocorticoids, which are also stress hormones, can have a permissive effect on some of these processes provided that insulin is available. Thus, they augment insulin action in stimulating the synthesis of glycogen and fatty acids.

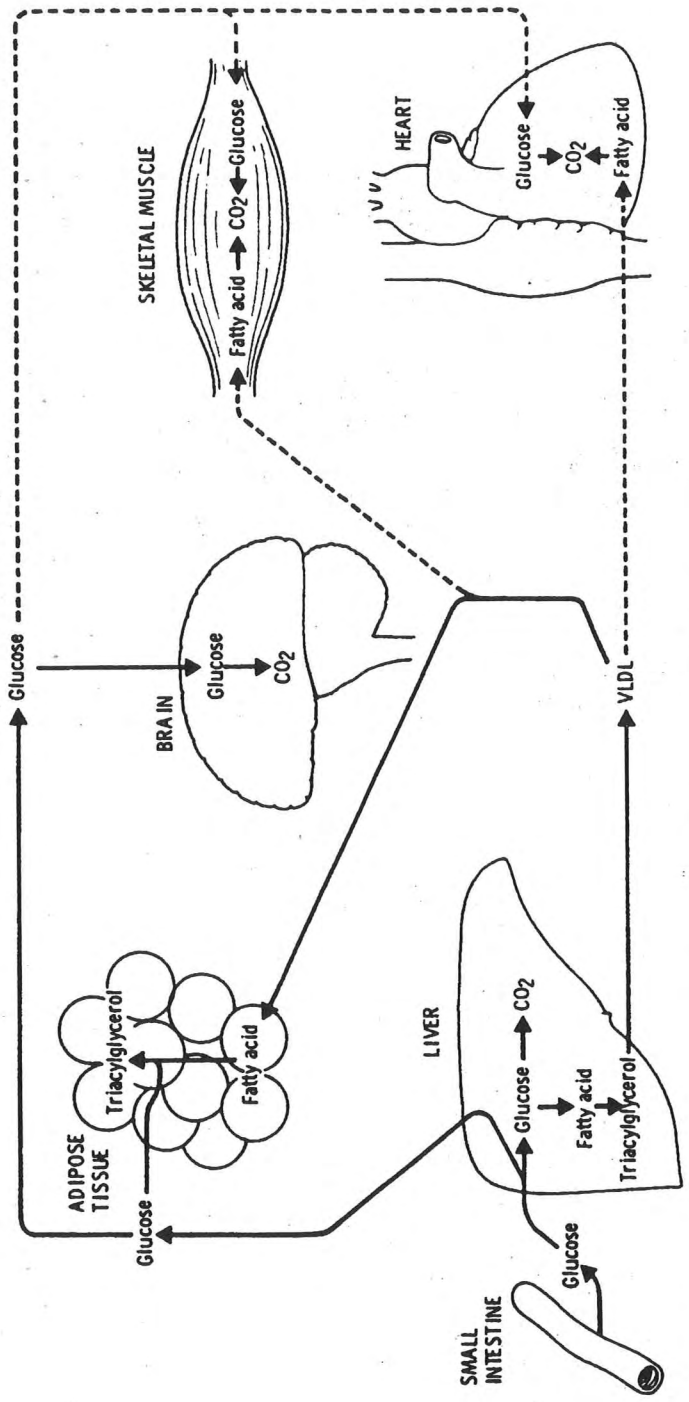


Figure 7.11. Some effects of insulin on metabolism. The solid lines represent major routes of metabolism.

The fatty acids that are produced are readily incorporated into triacylglycerols. Glycerolphosphate and dihydroxyacetone phosphate are provided by glycolysis, and fatty acids are not diverted into β-oxidation because of the inhibition of carnitine palmitoyltransferase by malonyl-CoA (Figure 7.9). Normally, the intracellular concentration of acyl-CoA esters remains relatively low, but because of effects of insulin and the relatively low concentration of cAMP, phosphatidate phosphohydrolase is readily translocated from the cytosol to the membrane-associated compartment, which facilitates glycerolipid synthesis. Furthermore, if glucocorticoids are relatively high, then the phosphohydrolase activity also increases and provides an increased potential for the synthesis of triacylglycerols.

Triacylglycerols do not normally accumulate in the liver but are efficiently packaged into VLDL and secreted into the blood (Figure 7.11). The main site of triacylglycerol hydrolysis and fatty acid uptake is in adipose tissue, since lipoprotein lipase activity in this organ is increased by insulin. This action of insulin can be augmented by the effects of glucocorticoids (Robinson et al. 1983; Cryer 1981). In lactation some of the VLDL (and chylomicrons) are diverted to the mammary gland, where prolactin increases lipoprotein lipase activity. The activity of lipoprotein lipase in skeletal and cardiac muscle is relatively low compared with adipose tissue, since muscle lipase activity is not maintained by insulin, but rather by glucocorticoids, with the possible involvement of glucagon, catecholamines, and thyroid hormones (Cryer 1981).

Insulin not only promotes the transfer of fatty acids (as triacylglycerols) from liver to adipose tissue, but it also stimulates glucose entry and the activities of the enzymes of triacylglycerol synthesis (Saggerson 1985). This leads to the efficient storage of triacylglycerol (Figure 7.9). The release of fatty acids from triacylglycerols in adipose tissue is suppressed because of the inactivation of hormone-sensitive lipase. This condition is maintained by the predominance of insulin over the stress hormones (Figure 7.12). The activation of the hormone-sensitive lipase by a cAMP-dependent phosphorylation is therefore prevented (Belfrage et al. 1984).

TRIACYLGLYCEROL METABOLISM IN CONDITIONS OF METABOLIC STRESS

Stress is characterized by the low activity of insulin relative to glucagon, the catecholamines, corticotropin, and the glucocorticoids in regulating metabolism. This type of hormonal balance occurs in starvation, diabetes, trauma, and toxic conditions. To a lesser extent, nutrients such as fructose, ethanol, and fat present the body with an energy load without releasing insulin. In fact, fructose and ethanol can provoke the release of stress hormones, and high-fat diets can cause tissues to become insulin-insensitive.

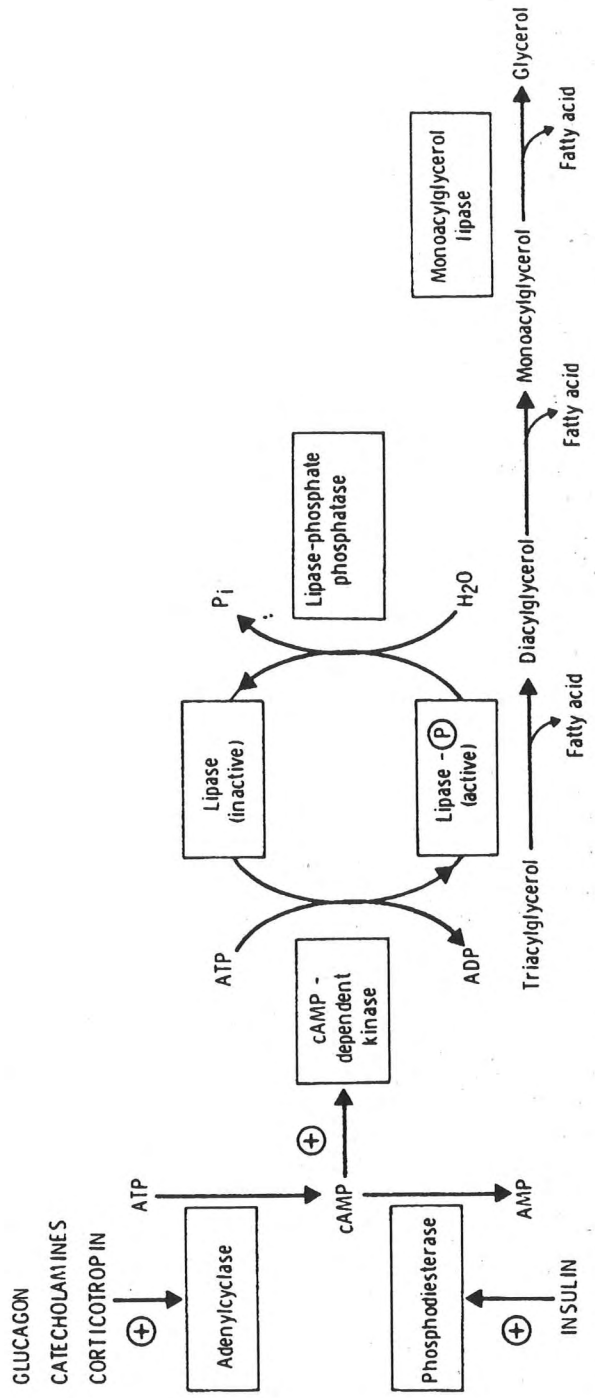


Figure 7.12. The control of the activity of hormone-sensitive lipase in adipose tissue. The plus sign indicates a stimulation.

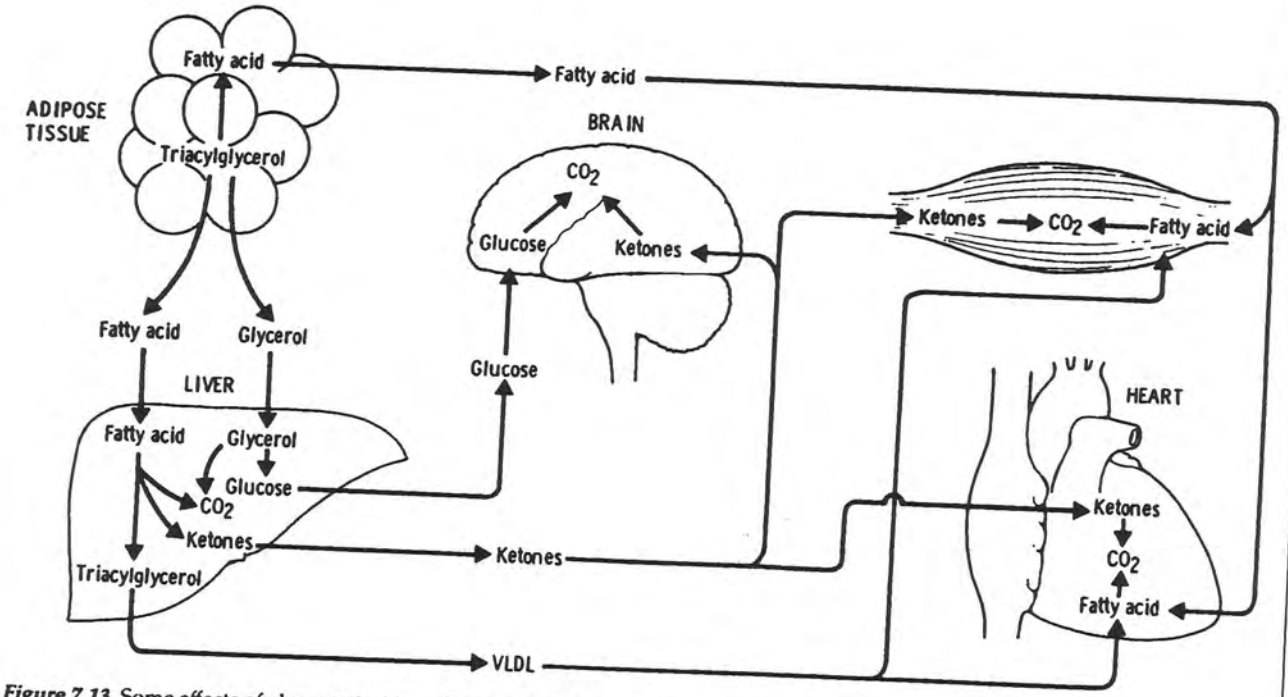


Figure 7.13. Some effects of glucocorticoids and high cAMP concentrations on metabolism.

The acute-acting stress hormones (glucagon, adrenalin) stimulate adenyl cyclase (adenylate cyclase) in a variety of tissues. In adipose tissue this stimulation leads to phosphorylation and activation (Belfrage et al. 1984) of the hormone-sensitive lipase (Figure 7.12). At the same time, the enzymes of triacylglycerol synthesis are inhibited (Saggerson 1985), and fatty acids and glycerol are consequently released into the circulation. Some of the fatty acids enter muscle tissue to become oxidized and provide energy (Figure 7.13). However, a large proportion of fatty acids is taken up by the liver. The high concentration of acyl-CoA esters that results, together with the increase in cAMP concentration, decreases the activity of acetyl-CoA carboxylase (Chapter 5), and the concentration of malonyl-CoA falls (Figure 7.9). Consequently, the inhibition of β-oxidation is relieved, and more of the fatty acids are metabolized to CO_2 and ketones. The oxidation process at low fatty acid supply can also be facilitated by the effect of cAMP in displacing phosphatidate phosphohydrolase from the membranes (Table 7.5), thereby inhibiting triacylglycerol synthesis.

However, in severe stress the supply of fatty acids and glycerol to the liver normally increases dramatically, as do the concentrations of acyl-CoA esters and glycerolphosphate in the liver. The high fatty acid availability overcomes the effects of cAMP in displacing phosphatidate phosphohydrolase (Table 7.5) and CTP:phosphocholine cytidylyltransferase (Chapter 8) from the membranes, and the synthesis of triacylglycerols and phosphatidylcholine is promoted. A fatty liver can often develop in stress conditions, when the capacity to secrete VLDL cannot cope with the increased rate of triacylglycerol synthesis. Fatty liver occurs especially in several toxic conditions where the synthesis of the apoproteins required in VLDL formation is impaired. Normally, the fatty condition is reversed when the fatty acid mobilization from adipose tissue subsides and the triacylglycerols can be secreted, or hydrolyzed so that the fatty acids can be oxidized by the liver.

The VLDL that are released during severe stress are preferentially metabolized by muscle tissue, since its lipoprotein lipase activity is increased by glucocorticoids and other stress hormones. The relative lack of insulin ensures that lipoprotein lipase activity is low in adipose tissue; the fatty acids are therefore prevented from recycling back into the storage depots (Figure 7.13).

One of the main functions of the liver in stress conditions and diabetes is to distribute energy to other organs in the form of glucose, ketones, and triacylglycerols (Table 7.6). It is significant that the control of phosphatidate phosphohydrolase in the liver by glucocorticoids, cAMP, the Ca^{2+}-mobilizing hormones, and insulin resembles

Table 7.6. Export of Energy from the Liver in Stress Conditions

Metabolite	*Taken up by*
Glucose	Brain and erythrocytes
Ketones	Skeletal and cardiac muscle
"	Brain
Triacylglycerols (VLDL)	Skeletal and cardiac muscle

that of enzymes involved in amino acid breakdown and gluconeogenesis (for example, tyrosine aminotransferase, argininosuccinate synthetase, argininosuccinate lyase, and phosphoenolpyruvate carboxykinase). Lipolysis from adipose tissue supplies the liver with fatty acids and glycerol for triacylglycerol synthesis, fatty acids for ketogenesis, and glycerol for gluconeogenesis (Figure 7.13). The hypertriglyceridemia and hyperglycemia that occur in stress and diabetes therefore seem to have a similar physiological basis. They result from increased hepatic secretion of triacylglycerol and glucose and their decreased removal from the circulation. The latter condition occurs because the uptake systems for glucose in muscle and adipose tissue and for triacylglycerols through lipoprotein lipase in adipose tissue are insulin-dependent.

FUTURE DIRECTIONS

Athough we have made progress in understanding the complex control of triacylglycerol metabolism, it is clear that there are many gaps in our knowledge. For instance, what are the major functions of the mitochondrial and peroxisomal systems that synthesize phosphatidate or lysophosphatidate? What are the relative contributions of glycerolphosphate and dihydroxyacetone phosphate to glycerolipid synthesis under different physiological conditions in different organs? Are the initial acyltransferases of glycerolipid synthesis controlled to a greater extent than is known at present? How are the enzymes of triacylglycerol synthesis regulated differentially in different tissues?

In order to answer some of these questions, a number of technical problems have to be overcome. The increased use of tissue- and cell-culture techniques has already meant that the specific effects of hormones and metabolites on triacylglycerol metabolism can be investigated. However, it is often difficult to reproduce or to appreciate the situation that exists in vivo. Care must be taken in interpreting the results and in deciding what is, or is not, an artifact. As always, the conclusions have to be compatible with what occurs in vivo.

One of the biggest problems in lipid metabolism has always been to isolate, characterize, purify, and to raise antibodies against the

enzymes which are often tightly associated with membranes. Until this is done, the mechanisms by which the enzymes act and are controlled will not be fully understood. It is also difficult to work with substrates that are insoluble in water. In addition, a strict kinetic analysis of reaction rates is normally not possible.

If we can overcome some of these technical problems, then the rewards are potentially high. The regulation of triacylglycerol metabolism is an important part of intermediary metabolism. A greater appreciation of this regulation would help to alleviate or prevent a number of clinical conditions, including fatty liver, diabetes, obesity, hyperlipidemias, and atherosclerosis.

PROBLEMS

1. Describe the routes by which phosphatidate and lysophosphatidate can be synthesized in the liver. In which part of the hepatocytes do these reactions take place, and what is known about the significance of these different routes of metabolism?
2. Compare and contrast the digestion, absorption, and transport of the short-, medium-, and long-chain fatty acids that are present in dietary triacylglycerols of nonruminant animals.
3. What differences are there between the control of the synthesis of triacylglycerols in liver and adipose tissue?
4. Describe how the synthesis of triacylglycerols in the liver and their transport in VLDL are controlled after the consumption of a meal rich in starch.
5. Describe the mechanism and control processes by which fatty acids are mobilized from adipose tissue in ketotic diabetes. What are the major metabolic fates of these acids?

BIBLIOGRAPHY

All entries are suggested for further reading.

Belfrage, P.; Fredrikson, G.; Stralfers, P.; and Tornqvist, H. 1984. In *Lipases*, eds. B. Borgström and H. Brockman, 365–416. Amsterdam: Elsevier.

Brindley, D. N. 1984. In *Fats in animal nutrition*, ed. J. Wiseman, 85–103. London: Butterworth's Scientific Ltd.

Brindley, D. N. 1985. Intracellular translocation of phosphatidate phosphohydrolase and its possible role in the control of glycerolipid synthesis. *Prog. Lipid Res.* 23: in press.

Brindley, D. N., and Sturton, R. G. 1982. "Phosphatidate metabolism and its relation to triacylglycerol biosynthesis." In *Phospholipids: New comprehensive biochemistry*, Vol. 4, eds. J. N. Hawthorne and G. B. Ansell. Amsterdam: Elsevier Biomedical Press.

Cryer, A. 1981. Tissue lipoprotein lipase activity and its action in lipoprotein metabolism. *Int. J. Biochem.* 13:525–41.

Robinson, D. S.; Parkin, S. M.; Speake, B. K.; and Little, J. A. 1983. In *The adipocyte and obesity*, eds. A. Angel, C. H. Hollenberg, and D. A. K. Roncari, 127–36. New York: Raven Press.

Saggerson, E. D. 1985. In *New perspectives in adipose tissue*, eds. A. Cryer and R. L. R. Van, Chapter 5. London: Butterworth's Scientific Ltd.

Wilson, J. E. 1980. Brain hexokinase, the prototype ambiquitous enzyme. *Curr. Top. Cell. Regul.* 16:1–44.

Alkyl Isocyanates as Active-Site-Specific Reagents for Serine Proteases. Reaction Properties†

William E. Brown‡ and Finn Wold*

ABSTRACT: Alkyl isocyanates have been found to inhibit serine proteases with a degree of specificity which strongly suggests a derivatization of the active site which depends on specific recognition of the proper alkyl chain by each enzyme. Thus, chymotrypsin is inactivated by stoichiometric reaction with octyl isocyanate and near stoichiometric reaction with butyl isocyanate; elastase does not react with octyl isocyanate, but is inactivated by stoichiometric reaction with butyl isocyanate; and trypsin is not affected by either reagent. Using chymotrypsin as the model, it has been shown that the inactivation reaction requires an intact unprotected active site (chymotrypsinogen and inactive forms of chymotrypsin produced by protonation or by treatment with acetone do not react, and indole protects the active enzyme against inactivation) and also that the carbamoylated enzyme derivative has an altered active site (loss of all-or-none activity in the titration with *N*-trans-cinnamoylimidazole, and altered reactivity of one of the two histidines in chymotrypsin). The selectivity observed in the reaction of the alkyl isocyanates with the three proteases is consistent with their known substrate specificities, and can also be explained on the basis of the structure of the substrate binding pockets observed in the three-dimensional models derived from X-ray diffraction analyses.

The pancreatic serine proteases chymotrypsin, trypsin, and elastase have been extensively studied during the last few decades. The three enzymes have been found to have very similar properties in their roles as catalysts for the hydrolysis of amides (and esters), and differ only in their substrate specificity; chymotrypsin recognizes large hydrophobic side chains, trypsin the positively charged side chains of lysine and arginine, and elastase small, aliphatic side chains in the position immediately to the N-terminal side of the peptide bond which is cleaved. The complete covalent structures of all three enzymes have been determined (Hartley, 1964; Walsh and Neurath, 1964; Hartley *et al.*, 1965; Shotton and Hartley, 1970) and more recently complete three-dimensional models have been constructed from high-resolution crystallographic analyses of derivatives of the three enzymes (Sigler *et al.*, 1968; Stroud *et al.*, 1972; Watson *et al.*, 1970). A comparison of the three structures reveals a remarkable degree of homology, the three catalytic sites are essentially identical, all containing the active-site serine which is acylated during the catalytic process and the participating side chains of histidine, aspartic acid, and another serine in very similar geometrical arrangements. Adjacent to the catalytic site all three enzymes contain a very pronounced cavity or pocket, and the dimensions and chemical makeup of the three individual pockets have been proposed to match perfectly with the substrate specificity of the three enzymes (Shotton and Watson, 1970). Thus, the chymotrypsin binding pocket is hydrophobic and open, the trypsin binding pocket is similar to that of chymotrypsin but uniquely contains a negatively charged aspartic acid residue in the bottom (Stroud *et al.*, 1972), and the elastase binding pocket, also similar to the other two, contains a bulky valine residue toward the front of the cavity, giving this enzyme a much shallower binding pocket (Shotten and Watson, 1970).

This very attractive structure–function model of these three well-characterized, homologous enzymes rests heavily on information derived from the studies of the crystals of inactive derivatives of the enzymes, and it is of obvious interest to be able to test the model by studies on the active enzymes in solution. In an attempt to perform such a test, we have reacted the serine proteases with a homologous series of alkyl isocyanates. The finding that these reagents behaved as highly specific active-site reagents for some of the serine proteases is not very surprising in view of the highly reactive serine residue in these enzymes. However, the finding that the size of the alkyl group determined the specificity of a given reagent toward a certain enzyme was more surprising and suggested to us that these reagents might represent an experimental tool by which the binding pocket model could be tested for the proteases in dilute aqueous solution. Initial experiments were conducted with ethyl, propyl, butyl, and octyl isocyanates and with hexamethylene diisocyanate, but because of the unpleasant properties of the lower homologues in particular, the bulk of the work to be reported was restricted to butyl isocyanate and octyl isocyanate as representatives of a "short" and a "long" side chain, respectively. Most of the detailed studies were done only with chymotrypsin, but the findings are judged to be relevant to the other enzymes as well.

A preliminary report of alkyl isocyanate specificity has appeared (Brown and Wold, 1971). It is the purpose of the present series of papers (1) to document the usefulness of alkyl isocyanates as active-site-specific reagents for chymotrypsin and elastase, (2) to present the evidence that the active-site serine residue is specifically derivatized in the reaction (Brown and Wold, 1973), and (3) to establish the actual location of the alkyl group in relation to the binding pocket by crystallographic analyses (work in progress).

Experimental Section

Materials and Assay Procedures. The materials and the standard assay procedures used for determination of either

† From the Department of Biochemistry, University of Minnesota, Minneapolis, Minnesota 55455. *Received October 20, 1972.* This work was supported by a U. S. Public Health Service research grant from the National Institutes of Health (GM 15053).

‡ U. S. Public Health Service predoctoral trainee (T1-GM-157). Present address: Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Conn.

TABLE I: Reagents, Assays, and Methods Employed in This Study.

Material (Source)	Purity	Assay Substrate (Source)	Ref
α-Chymotrypsin (Pentex)	Twice-crystallized lots 41 and 43	Bz-L-TyrOEt (Sigma)	*a*
		Dimethyl-casein	*b*
		Transcinnamoylimidazole (Aldrich)	*c*
Elastase (Worthington)	Electrophoretically pure	Elastin (Worthington)	*d*
	Twice crystallized	*N*-Ac-L-Ala-L-Ala-L-AlaOMe (Cyclo)	*e*
Trypsin (Pentex)	Twice crystallized	Tos-L-ArgOMe (Sigma)	*a*
Carboxypeptidase A (Worthington)	Three-times-crystallized lot 6139	Hippuryl-L-phenylalanine (Mann)	*f*
Carboxypeptidase B (Worthington)	Dip-F-treated lot 91 A	Hippuryl-L-arginine (Mann)	*g*
Papain (Worthington)	Twice crystallized	Bz-L-ArgOEt (Sigma)	*h*
Pepsin (Worthington)	Twice-crystallized lot 622	Equine hemoglobin (Pentex)	*i*
n-Butyl isocyanate (K & K Labs)	95–99%	Benzylamine (Eastman)	*j*
n-Octyl isocyanate (K & K Labs)	95–99%		
Hexamethylene diisocyanate (Aldrich)	98%		
[^{14}C]Butyl isocyanate (New England)	3.25 Ci/mol 99+% pure lot 544–046 as determined by radioactivity trace of mass trace of gas chromatography data supplied by New England Nuclear.		

[a] Hummel (1959); Worthington Biochemical Corp. (1972). [b] Lin *et al.* (1969). [c] Schonbaum *et al.* (1961). [d] Schneider *et al.* (1962). [e] Gertler and Hofmann (1970). [f] Folk and Schirmer (1963); Worthington Biochemical Corp. (1972). [g] Folk *et al.* (1960); Worthington Biochemical Corp. (1972). [h] Whitaker and Bender (1965); Worthington Biochemical Corp. (1972). [i] Anson (1938); Worthington Biochemical Corp. (1972).

activity or purity of enzymes and reagents are listed in Table I. Assay procedures not found in the literature are described in detail below.

ASSAY OF ISOCYANATE CONCENTRATION. Owing to the rapid hydrolysis rate and low solubility of alkyl isocyanates in aqueous solvents, all stock solutions and dilutions of the isocyanates were made in anhydrous acetone. In order to determine the actual concentration of isocyanate used in each protein reaction, a procedure was developed for the quantitative titration of the isocyanates based on their reaction with benzylamine in chloroform, extraction of the unreacted benzylamine into aqueous acid solution, and spectrophotometric determination of the unreacted benzylamine as its hydrochloride.

To a series of stoppered test tubes containing 0.5 ml of chloroform was added 5 μl of 1.0 M isocyanate (stock solution in acetone) and graduated quantities (0–100 μl) of 8.7×10^{-2} M benzylamine (stock solution in chloroform). The mixtures were reacted at room temperature for 15 min with shaking. Three milliliters of 0.1 M HCl was then added and the tubes were shaken for 15 min to ensure complete extraction of the unreacted benzylamine. The aqueous layers were collected using Whatman silicone-treated 1-PS phase-separation paper, and their absorbance was measured at 266 nm. A plot of the absorbance *vs.* micromoles of benzylamine added gave the equivalents of isocyanate present in the reaction mixture. This procedure was found to be reproducible to ±1%.

TITRATION OF ISOCYANATES IN AQUEOUS SOLUTION. To a constantly stirring solution (20 ml) of 0.08 M Tris-HCl (pH 7.7), containing 0.1 M $CaCl_2$, was added 200 μl of 5.1×10^{-2} M isocyanate (in acetone). At selected time intervals, 2.0-ml aliquots were removed and added immediately to 5 ml of 1.0×10^{-2} M benzylamine in chloroform. The mixture was stirred vigorously for 1 min on a Vortex mixer and transferred to a separatory funnel containing 50 ml of 0.1 M HCl. This mixture was shaken for 1 min and the optical density of the chloroform layer was measured at 258 nm. In this method, the isocyanate available for reaction with benzylamine after different lengths of exposure to aqueous buffer was determined as the chloroform soluble *N*-alkyl-*N'*-benzylurea derivative.

TITRATION OF ENZYMES WITH ISOCYANATES. To a buffered solution (0.1 M Tris-HCl, pH 7.7) containing enzyme (10^{-6}–10^{-4} M) was added an aliquot of isocyanate (0.25–10 mg/ml in anhydrous acetone). After 15 min of constant stirring at room temperature, an aliquot of the reaction mixture was removed for enzyme activity assay and whenever applicable, for determination of radioactivity. A second aliquot of isocyanate solution was then added. After 15 min, another assay aliquot was removed and the preceding procedure was repeated until the desired end point was reached. The final concentration of acetone in the reaction mixture never exceeded 2% (v/v). Radioactivity was determined with a Beckman LS-133 liquid scintillation spectrometer after gel filtration to remove excess reagent and reagent products from the protein derivative. Reagents tested as active-site "protective" agents were mixed in the buffer solution and preincubated with the enzyme for at least 15 min before the titration was begun.

Results

Characterization of the Alkyl Isocyanates. The stability of alkyl isocyanates in aqueous solvents has been studied (Saunders and Slocombe, 1948); however, it was felt important to determine the stability of the isocyanates under the exact conditions used in this work. The decomposition of isocyanates in the presence of water should be the result of the hydrolysis to carbon dioxide and the corresponding amine (eq 1), and the subsequent reaction of the amine with a second

$$RN{=}C{=}O + HOH \longrightarrow RNH_2 + CO_2 \quad (1)$$

mole of isocyanate to form symmetrical ureas (eq 2). The

$$RN{=}C{=}O + RNH_2 \longrightarrow RNHCNHR \quad (2)$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\quad \| \quad O$$

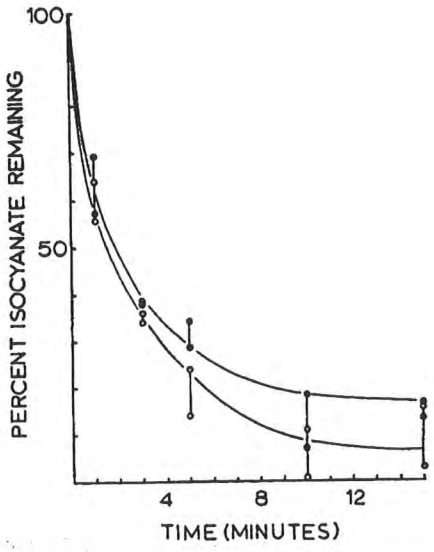


FIGURE 1: The rate of decomposition of octyl isocyanate (O) and butyl isocyanate (●) in aqueous media under the conditions used for the protein reactions (0.08 M Tris-HCl buffer, pH 7.7, containing 0.1 M $CaCl_2$).

results from our studies (Figure 1) show that the half-life of the isocyanates under our reaction conditions is approximately one minute and that there is no significant difference between the rates of decomposition of octyl and butyl isocyanate.

Studies on the Isocyanate–Protease Reaction. SPECIFICITY. The results of the titration of a variety of enzymes with octyl and butyl isocyanate under similar conditions of enzyme concentration, ionic strength, and temperature are given in Table II. Under these conditions octyl isocyanate at a molar ratio of 50:1 (reagent:enzyme) had little or no effect on trypsin, elastase, pepsin, or carboxypeptidases A and B, while chymotrypsin and papain[1] were totally inactivated. Butyl isocyanate under identical conditions, similarly had no effect on trypsin but inactivated both chymotrypsin and elastase with a greater efficiency toward elastase than toward chymotrypsin.

KINETICS. The rate of inactivation of chymotrypsin by octyl isocyanate and elastase by butyl isocyanate is too rapid to measure by normal sampling techniques. A search for special spectral properties of the product which would make it possible to follow the reaction directly in the spectrophotometer was negative, and a detailed study of the reaction kinetics was therefore not attempted. However, knowing the rate of hydrolysis of the isocyanates (Figure 1) it is possible to obtain an indirect measure of the rate of the protein reaction relative to the competing hydrolysis reaction. When chymotrypsin and elastase, at different concentrations, were titrated with either octyl or butyl isocyanate the results in Table III were obtained, showing that as the enzyme concentration increases from 1.9×10^{-6} to 4×10^{-4} M the inactivation reaction becomes increasingly more efficient. This is consistent with the expected second-order reaction which successfully competes with the hydrolysis reaction. The fact that at the higher enzyme concentrations the reaction with the enzyme predominates (50% inactivation with 0.6 mol of isocyanate/mol of enzyme)

TABLE II: Effect of Octyl and Butyl Isocyanate on Proteases.

Enzyme	Enzyme Concn (M × 10^6)	Reaction pH	% Act. Remaining at Reagent to Enzyme Molar Ratio of 50:1	
			Octyl Isocyanate	Butyl Isocyanate
Chymotrypsin	1.9	7.6	0	19
Trypsin	2.0	7.6	85	85
Elastase	1.9	7.6	94	5
Carboxypeptidase A	6.4	7.6	90	
Carboxypeptidase B	3.0	7.6	100	
Pepsin	2.3	2.0	100	
Papain	4.8	3.0	0	

demonstrates that the protease–isocyanate reactions are very fast. The results further emphasize the specificity of octyl isocyanate for chymotrypsin and butyl isocyanate for elastase.

STOICHIOMETRY. All the titration experiments indicated that the protease–isocyanate reaction involved a single reaction site and that loss of activity was associated with the modification of that single site. However, this point is sufficiently important for the argument that the alkyl isocyanates are active-site-specific reagents, that it was deemed essential to determine directly the relationship of reagent incorporation (not merely reagent added, as in the previous experiments) to loss of activity. This was accomplished using [^{14}C]butyl isocyanate (labeled in the carbonyl carbon), and the resulting plot of activity loss *vs.* reagent incorporation (Figure 2) demonstrates the exact stoichiometry of one for both enzymes. The deviation from linearity in the curve for chymotrypsin under conditions where elastase could be titrated directly to the end point is consistent with the lower specificity of butyl isocyanate for chymotrypsin.

Evidence for the Direct Involvement of the Active Site of Chymotrypsin in the Reaction with Alkyl Isocyanates. The conclusion that the alkyl isocyanates indeed act as active-site-specific reagents for the serine proteases, should be based on solid evidence that an intact, operational active site is required for the specific inactivation reaction. There are several approaches through which such evidence can be obtained, and some of these were used to explore the inactivation of chymotrypsin.

REACTION OF ZYMOGEN. It was found early in this work that the inactive zymogen, chymotrypsinogen, could be treated with alkyl isocyanates under condition were chymotrypsin in parallel experiments was completely inactivated, and then be activated with trypsin to give 100% yield of active chymotrypsin. Thus, the specific inactivation reaction does not occur with the zymogen.

pH DEPENDENCE. It is well known that low pH, protonated forms of chymotrypsin are enzymatically inactive. Are they also inactive in the reaction with octyl isocyanate? Chymotrypsin (2×10^{-6} M) was reacted with octyl isocyanate (1.6×10^{-5} M) at variable pH values, and after incubation for 15 min, the chymotrypsin activity was assayed at pH 7.6 using benzoyl-L-tyrosine ethyl ester as substrate. In order to interpret the results in terms of protein reactivity, it was necessary to establish that the competing reaction, the hy-

[1] The reaction of papain will not be discussed further in this paper. The finding that papain was inactivated at any pH below 5.7 has subsequently been followed up in this laboratory and will be reported in a future publication.

TABLE III: Effect of Enzyme Concentration on the Reaction of Chymotrypsin and Elastase with Octyl and Butyl Isocyanate.

Enzyme Concn (M)	Reagent to Enzyme Ratio Required for 50% Inactivation			
	Chymotrypsin		Elastase	
	Butyl	Octyl	Butyl	Octyl
1.9×10^{-6}	16	2.0	3.6	>100
8.0×10^{-5}	1.4	0.6		
4.0×10^{-4}	0.6	0.6	0.6	>100

drolysis of the isocyanate was independent of pH in the range studied. To this end, three samples of octyl isocyanate were incubated in aqueous solution at pH 3, 5, and 7 for exactly 1 min and then added to an excess of chymotrypsin at pH 7.6. The extent of inactivation of chymotrypsin was identical for all three samples, showing that the rate of hydrolysis of octyl isocyanate is identical at the three pH values investigated. (Above pH 8 a very significant increase in the rate of hydrolysis was observed.)

The data for the inactivation of chymotrypsin as a function of pH fit a theoretical titration curve for a functional group with a pK of 5.6 (Figure 3). It is not clear what the nature of this group or function in chymotrypsin is. We had expected to obtain some correlation with the considerably higher pK value (6.4–6.8) characteristic of the acylation step in the catalytic function. Although we have no explanation for this quantitative discrepancy at present, we feel that the general similarity between the pH effects on the catalytic activity and the carbamoylation of chymotrypsin strongly suggest that only the active (deprotonated) form of the enzyme reacts with the isocyanate.

EFFECT OF ACETONE. A second, similar type of experiment was based on the unique reversible inactivation of chymotrypsin by increasing concentrations of acetone (eq 3). The

$$\text{active chymotrypsin (E)} \underset{-\text{acetone}}{\overset{+\text{acetone}}{\rightleftharpoons}} \text{inactive chymotrypsin (E*)} \quad (3)$$

activity assay of chymotrypsin with dimethyl-casein as substrate gives a direct measure of E present at different concentrations of acetone. If an intact active site is required for the isocyanate reaction, octyl isocyanate should react only with E and not with E*. Assuming that the rate of interconversion of E and E* is slow relative to the isocyanate reaction with E, it should be possible to inactivate all E present at any acetone concentration by treatment with octyl isocyanate, and subsequently measure E* present in the original reaction mixture as the amount of active enzyme recovered after removal of the acetone. In order to minimize the conversion of E* to E during the reaction with octyl isocyanate, the reaction was allowed to proceed for only 2 min. The results of this experiment are given in Figure 4, and the best explanation for the complementarity of the two curves seems to be that only active enzyme (E) reacts specifically with the reagent. This can be illustrated using the data at 10% acetone (dotted line in Figure 4). The rate assay (solid circles) shows that 70% of the chymotrypsin is in the active form E in 10% acetone. After

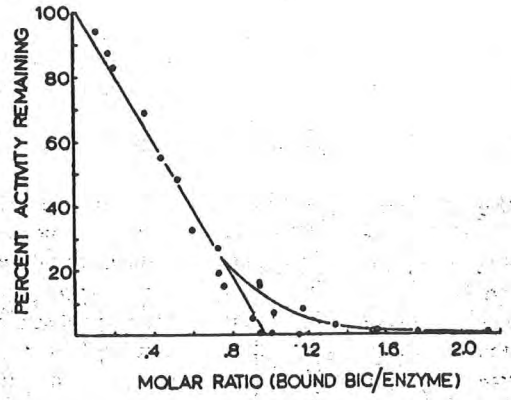


FIGURE 2: Inactivation of chymotrypsin (○) and elastase (●) as a function of incorporation of [^{14}C]butyl isocyanate (BIC). The concentration of both enzymes was 4×10^{-4} M. The reaction conditions and procedures are described in the text.

reaction with octyl isocyanate in 10% acetone, and measurement of activity remaining after removal of acetone (open circles), about 30% of the original activity was recovered, showing that 70% of the enzyme had been inactivated by octyl isocyanate.

KINETIC ANALYSIS OF PARTIALLY INACTIVATED ENZYME. Routinely, the inactivation of chymotrypsin was measured as the decrease in the direct initial rate assay at nonsaturating substrate concentrations. Since the rates measured in this manner could reflect changes in either catalytic turnover or in substrate affinity, it was deemed important to carry out a kinetic analysis of partially inactivated enzyme. When a sample of octylcarbamoyl-enzyme which by the initial rate assay had been 93.3% inactivated was subjected to substrate kinetic analysis, the K_m of the remaining activity was found to be identical to that of the fully active enzyme control, while the V_{max} was decreased to 5.5%. These findings were confirmed by the all-or-none assay (Schonbaum *et al.*, 1961), which simply quantitates the available active sites by acylation with *N*-trans-cinnamoylimidazole. Titration of the above inactivated sample showed that 6.3% of the original active sites remained

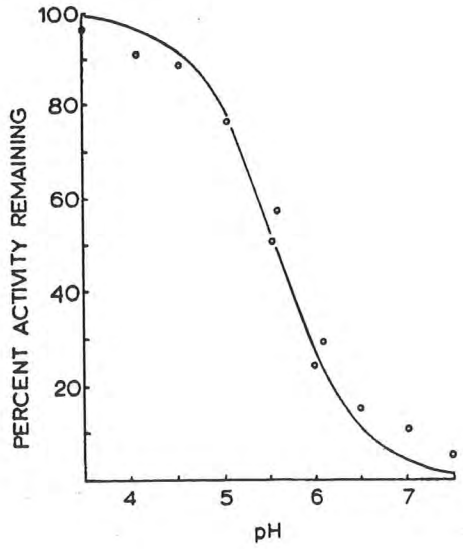


FIGURE 3: The effect of pH on the inactivation of chymotrypsin by an 8:1 molar excess of octyl isocyanate. The solid line represents the theoretical titration curve for a group with an apparent pK of 5.6.

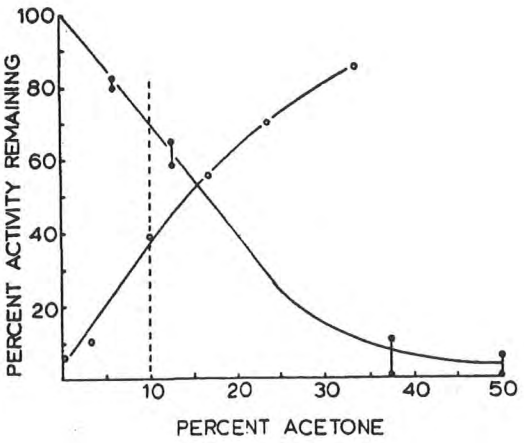


FIGURE 4: The effect of acetone on the reaction of chymotrypsin with octyl isocyanate. Chymotrypsin (3.2×10^{-7} M) in 0.1 M phosphate buffer (pH 7.5) was assayed in the presence of increasing concentrations of acetone using dimethyl-casein as substrate (●). Solutions of $4–8 \times 10^{-5}$ M chymotrypsin in 0.01 M Tris-HCl buffer (pH 7.5), containing increasing concentrations of acetone were reacted with a 1.1:1 molar ratio of octyl isocyanate to chymotrypsin After 2 min, each reaction mixture was diluted by a factor of 30 and the amount of unreacted chymotrypsin was estimated in terms of activity remaining using Bz-L-TyrOEt as substrate (○).

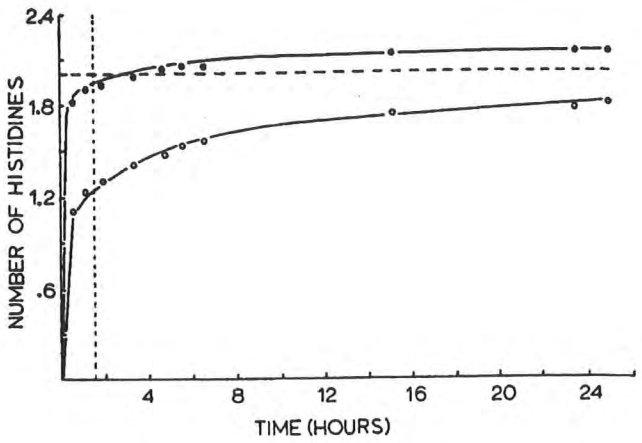


FIGURE 5: The titration of histidine residues in native chymotrypsin (●) and octylcarbamoyl-chymotrypsin (91% inactivated, ○) with diazo-1*H*-tetrazole. The proteins were diluted to 1.25×10^{-5} M in 1 M $KHCO_3$, pH 8.8. Two additions of diazo-1*H*-tetrazole (0.2 M) were made at 10-min intervals and the optical density was read at 480 and 550 nm in the Cary Model 15 spectrophotometer. Using the equation from Sokolovsky and Vallee (1966), the number of histidines were calculated as a function of time. The horizontal dashed line represents the theoretical number of histidines in chymotrypsin and the vertical dashed line represents the incubation time (90 min) recommended in the procedure of Sokolovsky and Vallee (1966).

after reaction with octyl isocyanate. These experiments show that the incorporation of the alkylcarbamoyl group leads to total inactivation of each site involved.

EFFECT OF CARBAMOYLATION WITH OCTYL ISOCYANATE ON THE REACTIVITY OF HISTIDINE IN CHYMOTRYPSIN. Chymotrypsin contains only two histines, one of which (histidine-57) is an essential component of the catalytic site. When a sample of native enzyme and a sample of totally octyl isocyanate inactivated chymotrypsin were titrated in parallel with the histidine-specific reagent diazonium-1*H*-tetrazole according to the method of Sokolovsky and Vallee (1966), the results in Figure 5 were obtained, showing that the reactivity of one of the two histidines had been significantly decreased in the inactivated enzyme. Since, as will be shown in the companion paper (Brown and Wold, 1973), the carbamoylation of the enzyme involves the active-site serine, the most reasonable interpretation of these findings is that the octylcarbamoyl-serine derivative sterically blocks the access of one of the histidines to the tetrazole reagent. We assume that the blocked residue is histidine-57, but we have no direct evidence to back up this assumption. A similar protection of histidine-57 against reaction with tosylamidophenyethyl chloromethyl ketone after derivatization of serine-195 with diisopropyl fluorophosphate has been reported, however (Schoellman and Shaw, 1963).

EFFECT OF INDOLE AS A PROTECTIVE AGENT AGAINST INACTIVATION. Indole is known to be a competitive inhibitor of chymotrypsin with a K_i of 0.8×10^{-3} M (Wallace *et al.*, 1963). If isocyanates bind in the active site, indole should protect chymotrypsin from the inactivation by isocyanates. Chymotrypsin was titrated with octyl isocyanate in the presence and absence of indole, and the data in Figure 6 show that indole gives almost complete protection at a molar ratio of 10^4:1 (indole:chymotrypsin). Finding this protection is especially meaningful since X-ray crystallographic work has provided direct evidence that the actual site of indole binding is the binding pocket (Steitz *et al.*, 1969). In connection with the indole protection it must be mentioned that we also tested several alkyl derivatives as protecting agents. Butyl alcohol at 10^4:1, octyl alcohol at 10^3:1, octylamine at 10^3:1, and hexyl cyanide at 10^4:1 gave no significant protection of chymotrypsin against inactivation by octyl isocyanate.

Discussion

Whereas aromatic isocyanates and isothiocyanates are quite common protein reagents (Edman, 1956; Schick and Singer, 1961; Fasold, 1965), the aliphatic isocyanates have found only limited use, probably because of their lability and lack of specificity. Ozawa (1967) studied the reaction of hexamethylene diisocyanate with the ε-amino group of lysine and with ribonuclease and chymotrypsin. He found in these studies that in 44% acetone, a 50:1 molar excess of hexamethylene diisocyanate gave complete inactivation of chymotrypsin. In view of the effect of acetone found in our work (Figure 4), the requirement for the large excess of reagent is not surprising. We have repeated Ozawa's (1967) experiments with octyl isocyanate and found that the molar ratio (reagent to enzyme) of 50:1 gave 90% inactivation in 44% acetone. In the absence of acetone hexamethylene diisocyanate was found to be as effective as octyl isocyanate, giving complete inactivation at a molar ratio of very nearly 1:1.

The results presented in this work indicate that the alkyl isocyanates represent very useful protein reagents. Because of the identical chemical properties of the reactive group of these readily available members of the homologous series of isocyanates, and because of their rapid reaction with functional groups in proteins under mild conditions, these compounds do indeed offer some unique possibilities as chemical probes in the study of enzyme structure and function. A large number of enzymes use substrates containing simple aliphatic side chains, *e.g.*, alcohols, fatty acids or esters, and amines. In all these cases the alkyl isocyanates should represent potential active-site reagents. As an example of this proposal it can be mentioned that as a result of the work reported here, and

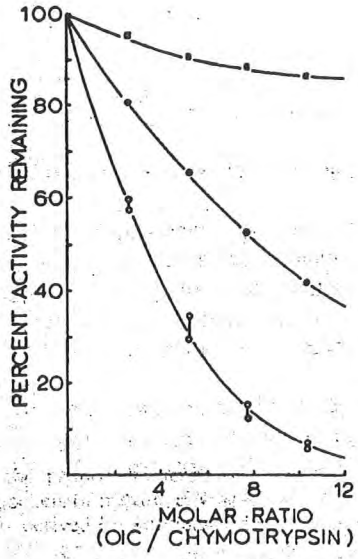


FIGURE 6: The effect of indole on the reaction of chymotrypsin with octyl isocyanate (OIC). Chymotrypsin (1×10^{-6} M) was incubated in 0.08 M Tris-HCl buffer (pH 7.8) plus 0.1 M $CaCl_2$ containing no indole (○), 10^{-3} M indole (●), and 10^{-2} M indole (□). After 15 min, these mixtures were titrated with appropriate aliquots of a solution of 1.29×10^{-2} M octyl isocyanate in acetone. Chymotryptic activity was measured by rate assay using Bz-L-TyrOEt as substrate.

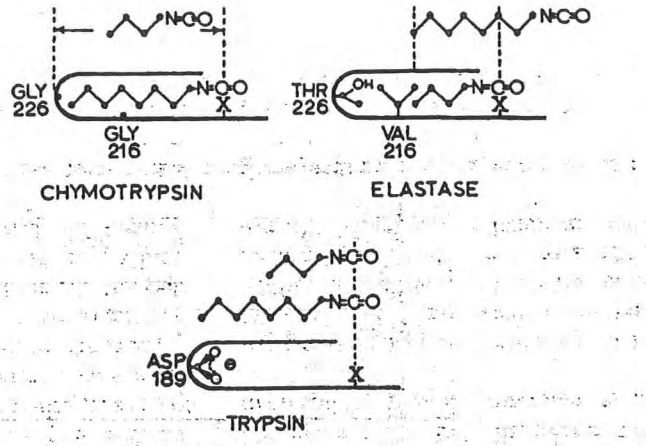


FIGURE 7: Schematic representations of the proposed interaction of the alkyl isocyanates with the binding pockets of trypsin, chymotrypsin, and elastase based on the models derived by Shotten and Watson (1970) and Stroud *et al.* (1972).

after establishing that the isocyanates form moderately stable derivatives with sulfhydryl groups, it was predicted that butyl isocyanate, as an analog of butanol, should be an active-site reagent for alcohol dehydrogenase. This prediction has now been confirmed (Twu and Wold, 1973).

With respect to the proteases, the significant finding in this work is the fact that the alkyl isocyanates show such a high degree of specificity in their reaction with the three proteases. The catalytic apparatus of trypsin, chymotrypsin, and elastase are essentially identical and since the active-site serine residue appears to be equally reactive in all three enzymes, a simple random reaction with alkyl isocyanate should not be expected to distinguish between the three enzymes. The most reasonable explanation for the reagent specificity must therefore be that the alkyl group is specifically recognized by the individual enzyme, and that the inactivation requires the formation of an enzyme–reagent complex that gives proper alignment of the two reactive groups (serine in the enzyme and the isocyanate in the reagent). The requirement for an intact active site documented for chymotrypsin is consistent with this explanation. In Figure 7 we have illustrated in a very schematic form the binding pockets for the three proteases. Assuming that the specificity is determined by reagent binding in these binding pockets, and the fact that indole protected chymotrypsin against inactivation by octyl isocyanate is consistent with that assumption, it is possible to explain most of our data. Octyl isocyanate is a specific reagent for chymotrypsin because this enzyme–reagent complex gives perfect alignment of the reactive groups and therefore, high probability of covalent-bond formation. In the butyl isocyanate–chymotrypsin complex, the alignment is more random. However, since the isocyanate is now in a hydrophobic environment, the hydrolysis rate could be reduced and the probability of covalent-bond formation might still be high. If we accept the argument that specific binding is a prerequisite for inactivation, and the fact that the rates of hydrolysis of octyl and butyl isocyanates are equal, as illustrated in Figure 1, then it should be possible to assess the relative affinity of butyl and octyl isocyanate for chymotrypsin from our data as the ratio of the concentrations of butyl isocyanate and octyl isocyanate required for 50% inactivation. Using the data in Table III at an enzyme concentration of 1.9×10^{-6} M, we determined this ratio to be 8:1. This is in excellent agreement with the results of Antonov *et al.* (1970), who found a K_i ratio of 10:1 for the inhibition of chymotrypsin by boronic acids with *n*-butyl and *n*-octyl chains, respectively.

The binding pocket of elastase is very similar to that of chymotrypsin but, because of the steric interference of the valine and threonine side chains, the binding pocket of elastase is shallower. Thus, only butyl isocyanate can bind with proper alignment for covalent bond formation. The longer octyl side chain probably also binds but because of poor alignment, has virtually zero probability of covalent-bond formation.

The negative results obtained with trypsin are also important for the argument. If the binding to trypsin requires a positive charge, none of the alkyl isocyanates should bind. Unless one also assumes other differences in the trypsin binding pocket such as in the extent of hydration, for example, it is difficult to explain the lack of effect of the shorter chain isocyanates on this enzyme. It is relevant to this model that hexamethylene diisocyanate has been found to inactivate trypsin at very nearly stoichiometric concentration. In this case we would like to propose that the hydrolysis of the diisocyanate leads to the formation of the monoamine–monoisocyanate intermediate, and that this is the real inhibitor of trypsin. It may be possible to check this hypothesis by X-ray crystallography studies of the inactive derivative, but it will also be important to attempt to prepare a compound which contains a protected, unreactive positive charge in addition to the isocyanate function, and test this compound as a specific trypsin inhibitor.

The companion paper (Brown and Wold, 1973) presents the experimental evidence which shows that residue X in Figure 7 is indeed the active-site serine. The evidence discussed above that the alkyl chain is located in the binding pocket, is clearly indirect, however, and it will be essential to prove this point before any direct comparison of solid state and solution properties of the proteases can be carried out in a meaningful manner. The reason for concern about this point is centered

about the question of whether there is more than one binding site which affects substrate binding in chymotrypsin (Jones *et al.*, 1965; Hofstee, 1957). One important test which could resolve this question and prove the binding pocket hypothesis, would be the direct observation by X-ray diffraction analysis of the actual position of the alkyl chain in the enzyme derivatives. The work to obtain this information has been undertaken. Octylcarbamoyl-chymotrypsin has been crystallized, and the crystallographic analysis is in progress.

References

Anson, M. L. (1938), *J. Gen. Physiol. 22*, 79.

Antonov, V. K., Ivanina, T. V., Berezin, I. V., and Martinek, K. (1970), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 7*, 23.

Brown, W. E., and Wold, F. (1971), *Science 174*, 608.

Brown, W. E., and Wold, F. (1973), *Biochemistry 12*, 835.

Edman, P. (1965), *Acta Chem. Scand. 10*, 761.

Fasold, H. (1965), *Biochem. Z. 342*, 295.

Folk, J. E., Piez, K. A., Carroll, W. R., and Gladner, J. (1960), *J. Biol. Chem. 235*, 2272.

Folk, J. E., and Schirmer, E. W. (1963), *J. Biol. Chem. 238*, 3884.

Gertler, A., and Hofmann, T. (1970), *Can. J. Biochem. 48*, 384.

Hartley, B. S. (1964), *Nature (London) 201*, 1284.

Hartley, B. S., Brown, J. R., Kauffman, D. L., and Smillie, L. B. (1965), *Nature (London) 207*, 1157.

Hofstee, B. H. J. (1957), *Biochem. Biophys. Acta 24*, 211.

Hummel, B. C. W. (1959), *Can. J. Biochem. Physiol. 37*, 1393.

Jones, J. B., Kunitake, T., Niemann, C., and Hein, G. E. (1965), *J. Amer. Chem. Soc. 87*, 1777.

Lin, Y., Means, G. E., and Feeney, R. E. (1969), *J. Biol. Chem. 244*, 789.

Ozawa, H. (1967), *J. Biochem. (Tokyo) 62*, 419.

Saunders, J. H., and Slocombe, R. J. (1948), *Chem. Rev. 43*, 203.

Schick, A. F., and Singer, S. J. (1961), *J. Biol. Chem. 236*, 2477.

Schneider, I. J., Tindel, S., Shapiro, D., and State, D. (1962), *J. Lab. Clin. Med. 60*, 514.

Schoellman, G., and Shaw, E. (1963), *Biochemistry 2*, 252.

Schonbaum, G. R., Zerner, B., and Bender, M. L. (1961), *J. Biol. Chem. 236*, 2930.

Shotton, D. M., and Hartley, B. S. (1970), *Nature (London) 225*, 802.

Shotton, D. M., and Watson, H. C. (1970), *Nature (London) 225*, 811.

Sigler, P. B., Blow, D. M., Matthews, B. W., and Henderson, R. (1968), *J. Mol. Biol. 35*, 143.

Sokolovsky, M., and Vallee, B. L. (1966), *Biochemistry 5*, 3579.

Steitz, T. A., Henderson, R., and Blow, D. M. (1969), *J. Mol. Biol. 46*, 337.

Stroud, R. M., Kay, L. M., and Dickerson, R. E. (1972), *Cold Spring Harbor Symp. Quant. Biol. 36*, 125.

Twu, J.-S., and Wold, F. (1973), *Biochemistry* (in press).

Wallace, B. A., Kurtz, A. N., and Niemann, C. (1963), *Biochemistry 2*, 824.

Walsh, K., and Neurath, H. (1964), *Proc. Nat. Acad. Sci. U. S. 52*, 884.

Watson, H. C., Shotton, D. M., Cox, J. M., and Muirhead, H. (1970), *Nature (London) 225*, 806.

Whitaker, J. R., and Bender, M. L. (1965), *J. Amer. Chem. Soc. 87*, 2728.

Worthington Biochemical Corp. (1972), Worthington Enzyme Manual, Freehold, N. J.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 265, No. 3, Issue of January 25, pp. 1369-1375, 1990
Printed in U.S.A.

The *Bacillus thuringiensis* δ-Endotoxin

EVIDENCE FOR A TWO DOMAIN STRUCTURE OF THE MINIMAL TOXIC FRAGMENT*

(Received for publication, June 26, 1989)

Daniel Convents‡§, Claude Houssier¶, Ignace Lasters‡, and Marc Lauwereys‡‖

From the ‡Plant Genetic Systems, N.V., J. Plateaustraat, 22, Gent B-9000, Belgium and the ¶Université Liège, Institut de Chimie (B6), Chimie Macromoleculaire et Chimie Physique, Liège B-4000, Belgium

The conformational characteristics of the minimal toxic fragment of the δ-endotoxin from *Bacillus thuringiensis* berliner 1715 were examined by fluorescence and circular dichroism spectroscopy. This insecticidal protein, specifically toxic to lepidopteran species, was found to consist of two structural domains. Experimental evidence for this conclusion was provided by biphasic guanidine hydrochloride unfolding curves at different pH values and electrophoretic patterns of protease digests. Two stable fragments of comparable molecular weight were obtained using four different broad specificity proteolytic enzymes.

A secondary structure model was constructed using seven *B. thuringiensis* toxin sequences. These toxins were selected on the basis of their limited sequence homology and represent all known insecticidal specificities. Despite this divergence, a consensus secondary structure pattern was obtained, confirming the structural homology among the toxins. The N-terminal halves of all toxins are predicted to be relatively rich in α-helix structure and the C-terminal parts to contain alternating β-strand and coil structures. The latter seems characteristic for a β-sheet conformation. Comparing this model to the unfolding data obtained by circular dichroism, whose far UV signal gives a measure of the α-helix content, allowed us to delineate the structural domains into the primary structure.

During the sporulation phase the Gram-positive bacterium *Bacillus thuringiensis* produces large crystalline inclusions. These parasporal crystals predominantly consist of M_r 130,000–140,000 protoxin molecules called δ-endotoxins (1). Upon ingestion by insect larvae the crystals are dissolved and the protoxin activated by midgut proteases, resulting in toxic fragments of M_r 30,000–70,000 (2). Crystal proteins with a specific toxicity to Lepidoptera, Diptera, and coleopteran insect species have been described (1, 3, 4). A large number of protoxin genes have been cloned. Their toxic fragment has in most cases been located in the N-terminal half of the protoxin primary structure (5–7). The C-terminal part of the protoxin, highly conserved among both Lepidoptera and Diptera specific toxins (8), is probably important for the assembly and stability of the crystal structure.

Limited data have been presented concerning the molecular structure of the protoxin molecules. A dimeric organization, formed by intermolecular disulfide bridges, has been proposed (9, 10). Raman spectroscopy of protoxin crystals revealed nearly equal amounts of α-helical and β-strand structures (about 25% each) (11). At the level of the toxic fragment, even less is known. The involvement of protein domains, each exerting a specific function, has often been discussed. For other protein toxins like diphtheria toxin and *Pseudomonas* exotoxin A, such a spatial separation of a cell recognition site, regions involved in membrane penetration and toxic enzymatic activity are well documented (12, 13).

Recently the high affinity binding of the *B. thuringiensis* toxin to brush border membrane vesicles of target insects has been reported. This membrane interaction is correlated with the protein's insecticidal activities among Lepidoptera (14, 15). The formation of ion channels across the midgut membrane of target insects was postulated to account for the lethal effect (16).

Here we report that the toxic fragment of Bt2, the recombinant Lepidoptera specific toxin of *B. thuringiensis* berliner 1715 (5), consists of at least two structural domains. Evidence for this model was obtained by monitoring protein unfolding using fluorescence and circular dichroism spectroscopy. Protease digests of the toxin analyzed on SDS-PAGE[1] further support this model. The secondary structure of seven *B. thuringiensis* toxins sequences, representing all known insecticidal specificities, was predicted. A consensus prediction was reached with the N-terminal halves of the toxins containing several α-helices, whereas the C-terminal part is rich in β-strand conformation. The helical propensity of the N-terminal part is correlated with the observed physicochemical unfolding behavior of the Bt2 toxic fragment.

EXPERIMENTAL PROCEDURES

Materials

Guanidine hydrochloride (GdnHCl), "Sequanal Grade," was purchased from Pierce Chemical Co. Thermolysin, chymotrypsin, trypsin (diphenyl carbamyl chloride treated) and papain were from Sigma, and Pronase was from Boehringer Mannheim. All other reagents were analytical grade.

Methods

Sequences, Alignments and Secondary Structure Prediction—The *B. thuringiensis* genes used are the following: Bt2, cloned from *B. thuringiensis* berliner 1715 (5) (toxic fragment: amino acid position 1–608); Bt4, cloned from *B. thuringiensis* HD 68[2] (toxic fragment: amino acid position 1 to 592); Bt15, cloned from *B. thuringiensis* HD 110[2] (toxic fragment: amino acid position 1–616); Bt8, cloned from

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Beneficiary of a fellowship from the Institut voor Wetenschappelijk Onderzoek in Nijverheid en Landbouw, Belgium.

‖ To whom correspondence should be addressed.

[1] The abbreviations used are: SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; GdnHCl, guanidine hydrochloride; MES, 2(*N*-morpholino)ethanesulfonic acid; MOPS, 3(*N*-morpholino)propanesulfonic acid; CAPS, 3(cyclohexylamino)-1-propanesulfonic acid; CHES, 2(*N*-cyclohexylamino)ethanesulfonic acid.

[2] H. Höfte, manuscript in preparation.

B. thuringiensis israelensis (7) (toxic fragment: amino acid position 1–634); Bt303, cloned from *B. thuringiensis israelensis* (18) (toxic fragment: amino acid position 1–678); Bt13, cloned from *B. thuringiensis tenebrionis* (19) (toxic fragment: amino acid position 1–644); Bt104, cloned from *B. thuringiensis* HD 263 (20) (P2 toxin, total sequence used; 590 amino acids).

The amino acid sequences of the toxic fragments from these seven toxins were aligned with GENALIGN, using the region program with default parameters (matching weight = deletion weight = 1). GENALIGN is a copyrighted software of IntelliGenetics, Inc.; the program was developed by Dr. Hugo Martinez of the University of California at San Francisco (21).

The aligned sequences were submitted to a combined secondary structure prediction using the algorithm developed by Gibrat *et al.* (22). This method allows the prediction of α-helix, β-strand, and coil structures from the primary amino acid sequence.

Purification of the Bt2 Toxin—The cloning of the Bt2 protoxin has been described previously (5). The gene was put under transcriptional control of the tac (23) promoter.[3] The *Escherichia coli* strain used was WK6 (24). Induction with isopropyl β-D-thiogalactoside resulted in the overproduction of authentic Bt2 protoxin. The protein was produced under the form of insoluble inclusion bodies. After cell lysis the pellet was extensively washed, first with a 1% Triton X-100, 1 M NaCl, pH 7.5, solution, followed by phosphate-buffered saline. The protoxin was solubilized during a 2-h incubation in 50 mM Na_2CO_3, 10 mM dithiothreitol buffer, pH 10, at 37 °C.

The toxic fragment was obtained by trypsin digestion. The protoxin solution was dialyzed at 4 °C against a Tris·HCl buffer, pH 8.6, containing 0.2 M NaCl. Trypsin was added to a final enzyme/substrate ratio of 1:20 (w/w), and digestion allowed for 2 h at 37 °C. The toxic fragment was further purified on a Mono Q 10/10 column connected to a fast protein liquid chromatography system (Pharmacia LKB Biotechnologies Inc.). Purified Bt2 toxin resulted in a single band on Coomassie Brilliant Blue-stained SDS-PAGE with an apparent M_r 60,000. Protein concentration was determined spectrophotometrically at 280 nm using an extinction coefficient of $E^{1\%} = 11.7$ cm^{-1}, calculated from the known amino acid composition.

Spectrofluorimetric Measurements—Fluorimetric measurements were performed on a SFM-25 Kontron fluorimeter equipped with a xenon light source. 10-nm bandwidth slits were used in excitation and emission beams. Cuvettes with a 2-mm excitation and 10-mm emission pathlength were used. The cell holder was maintained at 35 °C using a circulating water bath. Protein concentration was 0.25 μM. The fluorimeter was calibrated at 100% emission intensity at 340 nm with the native protein solution. Prior to measurement the samples were incubated 4 h at 35 °C for unfolding equilibrium to be reached.

Denaturations as a function of pH were done in the following buffer systems: from pH 4 to pH 8, 50 mM acetic acid, 50 mM MES, and 50 mM MOPS adjusted with NaOH; from pH 8 to pH 11, 50 mM CHES and 50 mM CAPS adjusted with NaOH. All buffers contained 0.2 M NaCl. To monitor protein denaturation, the ratio of the emission intensities at 370 and 342 nm was taken as the structure-dependent parameter (P). Since 342 and 370 nm are the emission maxima of the native and denatured protein, respectively, the value of P shifts from <1 to >1 upon protein unfolding.

The data are normalized by calculating the total fractional change (F_{app}) of the observable parameter:

$$F_{app}=(P - P_n)/(P_u - P_n).$$

P represents the parameter value at any point along the transition, while P_n and P_u are the limiting values for the native and unfolded state, respectively.

Circular Dichroism (CD) Measurements—A Jobin-Yvon Mark V dichrograph was used, connected to an Apple II microcomputer for data acquisition. A 1-mm pathlength was used, and the cell holder was kept at 35 °C. Spectra were taken as a mean of nine accumulations. Protein concentration was 7.5 μM. A 5 mM NaH_2PO_4, 0.2 M NaF buffer adjusted at pH 4, 8, and 11 was used for all measurements. The ellipticity at 222 nm was taken as the parameter for protein unfolding. Unfolding data were further analyzed as described for the fluorescence monitored denaturations.

Protease Digests—All proteases were added to 1 mg/ml Bt2 toxin solutions. Native toxin was digested by chymotrypsin, Pronase, and thermolysin, whereas protein equilibrated in 5 M GdnHCl was used for the papain digestion. All proteases were used in an enzyme/substrate ratio of 1:50 (w/w) at 35 °C. Following buffers were used for the digestions: for Pronase and thermolysin 50 mM MOPS, 5 mM $CaCl_2$, 0.2 M NaCl, pH 7; for chymotrypsin 20 mM Tris, 0.2 M NaCl, pH 8, and for papain 50 mM NaH_2PO_4, 0.2 M NaCl, 1% β-mercaptoethanol, pH 6.

Aliquots were taken at time points ranging from 5 min to 4 h incubation. Digestions were stopped by adding trichloroacetic acid to a final concentration of 10%. The precipitated samples were analyzed on SDS-PAGE performed as in Refs. 25.

RESULTS

Structure Prediction—In order to improve the limited confidence level of the secondary structure prediction method (63% overall reliability (22)), a combination of seven *B. thuringiensis* toxin sequences was considered. This set includes three toxins with an insecticidal activity against Lepidoptera, Bt2 (5), Bt4 (17), and Bt15 (17), two which are toxic to dipteran insects, Bt8 (7) and Bt303 (18), one Coleoptera-specific toxin, Bt13 (19), and one sequence of a P2 toxin, Bt104 (20), active to both Lepidoptera and Diptera.

These *B. thuringiensis* toxins were chosen on the basis of their limited sequence homology. The homology parameters are shown in Table I. Clearly, the three Lepidoptera-specific toxins share the highest degree of homology. Bt104 seems to be unrelated to any of the other *B. thuringiensis* toxins at the primary structure level. Fig. 1*a* shows the result of the global alignment on the seven *B. thuringiensis* toxins. Highly conserved stretches among all toxins except Bt104, or only among the lepidopteran specific toxins Bt2, Bt4 and Bt15, are indicated.

This set of aligned sequences was used for a secondary structure prediction study using the algorithm of Gibrat *et al.* (22). The result of this three-state prediction is shown in Fig. 1*b*. *Arrows* indicate local peaks in the combined prediction.

Four α-helical regions are predicted in the N-terminal half of the toxins. The second halves of the sequences almost completely lack consensus α-helix prediction.

Comparing Fig. 1, *a* and *b*, reveals that all helices are situated in rather conserved stretches, especially the second and the third helix. No reliability gain, which can be anticipated when integrating several *B. thuringiensis* toxin primary sequences into one secondary structure model, is expected for these predictions. In other words, the prediction confidence level of these regions is equivalent to that of a single polypeptide chain prediction.

β-Strand conformation is primarily predicted in the C-terminal halves of the toxins. Sharp β-strand peaks appear between residues 440 and 700. Careful analysis of this region with respect to the strand and coil conformations reveals a highly concerted alternation of the two structures in all toxins. Such a conformation may be indicative of a β-sheet conformation.

TABLE I

Primary sequence homology among seven B. thuringiensis toxins values expressed as the ratio of the alignment score (alignment score = matching weight − deletion weight; with the deletion weight = matching weight = 1) to the mean lengths of the two considered sequences.

	Bt2	Bt4	Bt8	Bt13	Bt104	Bt15	Bt303
Bt2		48.0	12.4	16.4	0	36.8	7.3
Bt4			9.0	14.7	0.2	41.4	20.5
Bt8				8.9	4.9	13.1	10.3
Bt13					0.3	16.0	11.8
Bt104						0.2	0
Bt15							3.6

[3] D. Conventa, C. Houssier, I. Lasters, and M. Lauwereys, unpublished results.

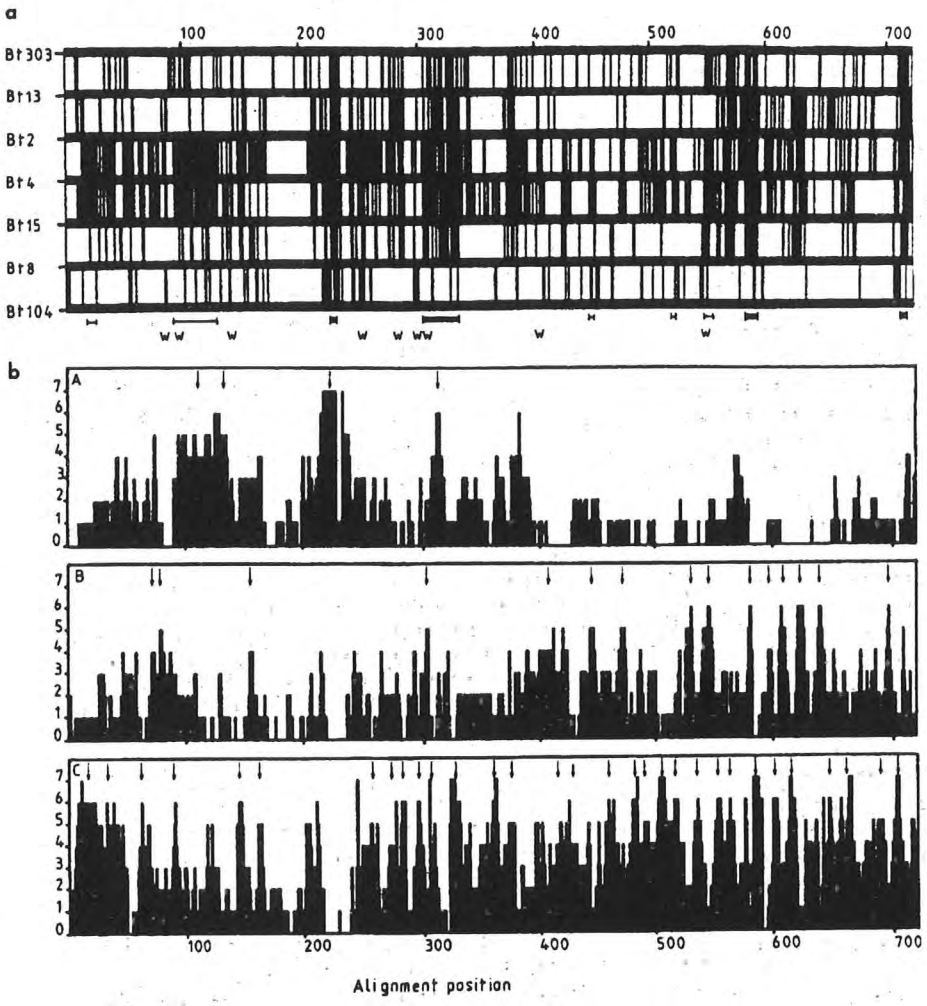


FIG. 1. *a*, alignments of the seven *B. thuringiensis* toxins. The sequences were aligned as described (21) and ordered so that the most related sequences are next to each other. Matching amino acids are represented by *vertical lines connecting the seven horizontal lines* which represent the sequences. Gaps included for optimal alignments are not indicated. Beneath the alignment, regions conserved among the Lepidoptera specific toxins (*thin lines*) and all toxins except Bt104 (*thick lines*) are indicated. The positions of the tryptophan residues in the Bt2 toxin are indicated by a *W*. *b*, combined secondary structure predictions of the seven *B. thuringiensis* toxins shown on the same scale as the alignment. α-Helix, β-strand, and coil predictions are displayed on separate histograms: A–C, respectively. At each position in the aligned set of toxins, the total number of residues predicted in one state was added and assigned to that position. Gaps in the alignment were not taken into account. *Arrows* indicate local peaks for a particular prediction. These were discriminated using the following criteria: at least four out of seven sequences should be predicted in the assigned conformational state; a minimum length of 4 residues for an α-helix; a minimum length of 3 residues for a β-strand; a minimum length of 4 residues for a coil; a stretch can only be interrupted at at most one position with a lower number of predictions in the considered state.

Only limited sequence homology among all toxins can be detected in the region between residue 525 and 700, except for a short stretch at position 580–590 which is predicted as a coil. Nevertheless, the overall structure pattern is conserved, with β-strand and coil conformations predicted in at least five out of the seven sequences. It is thus demonstrated that combining several sequences of related proteins, for which a strong conformational similarity can be assumed, results in a structural model likely to be more reliable. The most obvious result is the observation that the *B. thuringiensis* toxins seem to consist of two domains of secondary structure, the N-terminal part rich in helix, followed by a sheet conformation.

Denaturation Studies with Guanidine Hydrochloride—To investigate the overall structure of the *B. thuringiensis* toxins, the Bt2 toxic fragment (denoted further as toxin) was chosen as our model system. This toxin has a calculated molecular mass of 66 kDa and contains 9 tryptophan residues (see Fig. 1*a*).

Tracing protein unfolding using fluorescence spectroscopy scopes upon changes in tertiary structure after adding a chaotropic agent (26). Complementary, circular dichroism can be used when particular interest exists in observing a protein's secondary structure content (27, 28).

Fluorescence—Preliminary experiments indicated that 8 M urea is insufficient to completely unfold the Bt2 toxin. As a consequence GdnHCl was used as a denaturant. Fig. 2 shows the fluorescence spectra of native and denatured Bt2 toxin at pH 8.

The fluorescence characteristics of the indol side chains are altered both with respect to the emission intensity and the emission wavelength maximum (λ_{max}) (26). Both phenomena reflect a change in the microenvironment of the tryptophan side chains. The drop in the quantum yield may be correlated to the water exposure of tryptophan, although such a correlation is not straightforward (26). On the other hand the shift

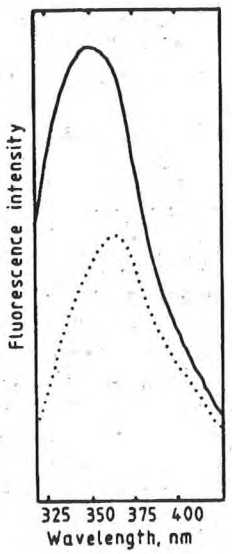


FIG. 2. **Fluorescence emission spectra of the Bt2 toxin in the native state (——) and denatured by 8 M GdnHCl (····).** The spectra were recorded in 20 mM Tris, 0.2 M NaCl, pH 8.

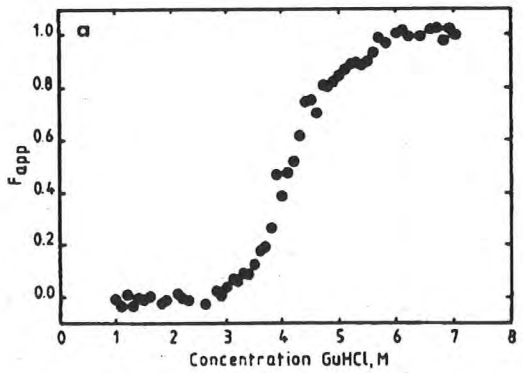
of λ_{max} to higher wavelengths is a general observation when studying protein unfolding by fluorescence (26). Therefore the red shift of λ_{max} was taken as the parameter of unfolding.

Fig. 3*a* shows the normalized denaturation curve (see "Experimental Procedures") of the Bt2 toxin at pH 8.0. Below 3 M GdnHCl no conformational change can be detected. Above 3 M GdnHCl the Bt2 toxin unfolds in a transition that continues to 6 M GdnHCl, where the protein is completely denatured. In a first approximation this curve can be considered to be the result of a single transition from the native to the unfolded state. As relatively high concentrations of denaturant are needed for complete unfolding, the Bt2 toxin can be considered as a protein of high conformational stability.

Denaturation curves at pH 4, 8, and 11 are shown in Fig. 3*b*. Clearly the pH has a considerable impact on the protein's conformational stability. Both the transition midpoint and the shape of the denaturation curve are pH-dependent.

Whereas no multiple transitions could be discriminated at pH 8, both the pH 4 and 11 unfolding curves are obviously biphasic. At pH 4 the first transition coincides with the initial unfolding at pH 8. Around 4 M GdnHCl a stable intermediate appears, which then further denatures between 5 and 6 M GdnHCl. In contrast, at pH 11 part of the Bt2 toxin becomes less stable, denaturing between 2 and 4.5 M GdnHCl, whereas another region has the same stability as at pH 8. The shapes of the unfolding curves at intermediate pH values between pH 4 and 11 fall between the two extreme curves depicted (experiments done at pH unit increments, data not shown).

Control experiments exclude the possibility that buffer components are responsible for the altered unfolding equilibrium as a function of GdnHCl concentration (data not shown).



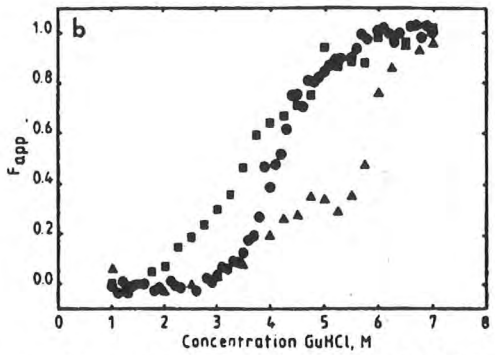


FIG. 3. *a*, denaturation of the Bt2 toxin by GdnHCl (*GuHCl*) at pH 8.0 monitored by fluorescence. Data were normalized as described. *b*, normalized denaturation curves of the Bt2 toxin at pH 11 (■), pH 8 (●), and pH 4 (▲) measured by fluorescence. Buffers are described under "Experimental Procedures."

Also no red shift of the emission spectrum was observed for the native Bt2 at the different pH values (data not shown).

It should be emphasized that, since at alkaline and acid pH a biphasic denaturation is observed, a similar behavior can be expected at pH 8. However, the stabilities of both protein regions do not differ sufficiently at this pH to permit discrimination of their independent unfolding.

Circular Dichroism—Since the CD signal is highly sensitive to changes in secondary structure, especially the α-helix content, this technique permits us to verify whether a correlation can be established between the unfolding experiments and the prediction of the protein's secondary structure.

At high protein concentrations the Bt2 toxin precipitated at pH 4 making CD measurements impossible. Only at alkaline pH were unfolding experiments monitored by CD feasible.

The far UV spectra of the native Bt2 toxin at pH 8 and both native and denatured protein at pH 11 are shown in Fig. 4. Within experimental error the CD spectra of the native protein solutions can be considered as identical, so no gross conformational change is monitored upon shifting the pH.

The unfolding curves at pH 8 and 11 are displayed in Fig. 5. Contrary to the fluorescence denaturation curves, the CD data show a single transition unfolding, both at pH 8 and 11, with a decreased stability at pH 11. Since the α-helix conformation gives rise to a CD band at 222 nm, circular dichroism at this wavelength is a good measure of the helix content of a

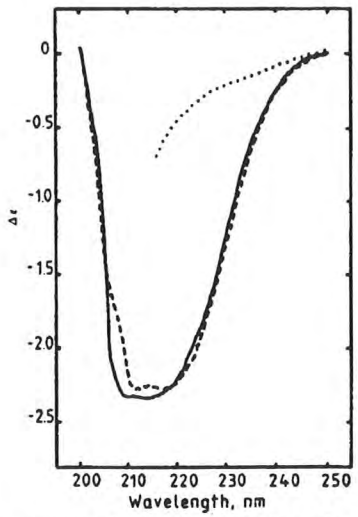


FIG. 4. **CD spectra of the native Bt2 toxin at pH 8 (——), and at pH 11, both in the native state (– – –) and denatured by 7 M GdnHCl (····).** All spectra were recorded in 5 mM NaH_2PO_4, 0.2 M NaF adjusted to the indicated pH.

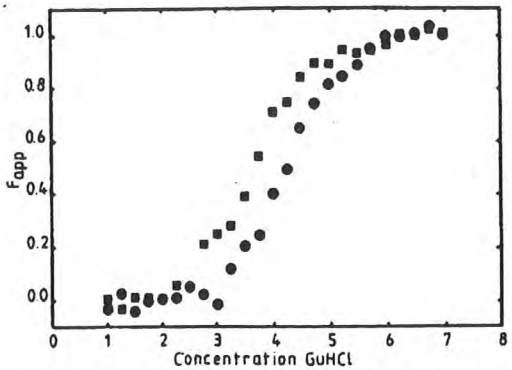


FIG. 5. **Normalized unfolding curves of the Bt2 toxin at pH 8 (●) and pH 11 (■), measured by the decrease in optical activity at 222 nm.** *GuHCl*, guanidine hydrochloride.

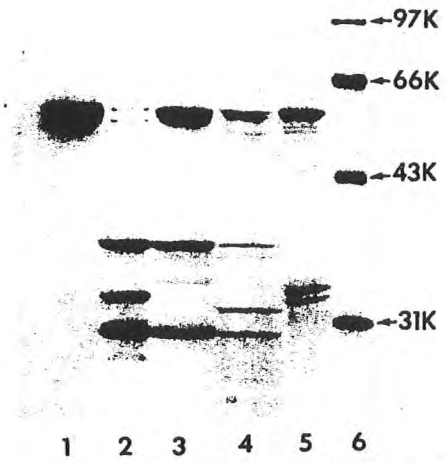


FIG. 6. **Protease digests of the Bt2 toxin on SDS-PAGE.** *Lane 1*, intact Bt2 toxin; *lane 2*, chymotrypsin digestion of native Bt2 after 8 h of incubation; *lane 3*, thermolysin digestion of native Bt2 after 7 h of incubation; *lane 4*, Pronase digestion of native Bt2 after 2 h of incubation; *lane 5*, papain digestion of Bt2 in the presence of 5 M GdnHCl after 8 h of incubation; *lane 6*, molecular mass markers: rabbit muscle phosphorylase *b* 97.4 kDa, bovine serum albumin 66.2 kDa, hen egg white ovalbumin 42.7 kDa, bovine carbonic anhydrase 31 kDa.

protein (29). These curves are therefore interpreted as an unfolding of a protein region rich in α-helix structure. This part of the protein is less stable at pH 11 than at pH 8.

Protease Digests—Whether the independently unfolding protein regions of the Bt2 toxin are organized as two spatially distinct parts, or as one globular unit, can be investigated by protease digestions. Three proteases were able to cleave the native protein. Common fragments of M_r 30,000 and 36,000 were generated by chymotrypsin, thermolysin, and Pronase even upon prolonged incubation, as shown in Fig. 6, *lanes 2–4*. Papain was ineffective and left the protein intact (data not shown).

As thermolysin and papain retain their catalytic properties at high concentrations of denaturant, the Bt2 toxin was treated with these enzymes after equilibration with GdnHCl, up to 5 M. Addition of thermolysin in the presence of denaturant (more than 4 M GdnHCl), resulted in random cleavage and complete degradation of the Bt2 protein (data not shown). On the other hand when papain was used discrete bands of M_r 33,000 and 32,000 appeared on SDS-PAGE as shown in Fig. 6, *lane 5*. These findings support a model in which the Bt2 toxin is composed of two protease resistant structural entities, linked by a flexible region.

Additional evidence for this model is provided by fluorescence unfolding data. Denaturation curves of intact Bt2 and the unfractionated chymotrypsin digest are completely superimposable (data not shown).

DISCUSSION

Using spectrophotometrical techniques we demonstrate that the toxic fragment of Bt2, a Lepidoptera specific *B. thuringiensis* toxin, consists of two regions which fold independently. Evidence therefore is obtained from biphasic denaturation curves at alkaline and acid pH. Such an unfolding behavior is characteristic of multidomain proteins (30) (31).

A secondary structure prediction model supports this hypothesis. The N-terminal half of the Bt2 toxin is predicted to contain several helical stretches, whereas the C-terminal part is devoid of this conformational state. Such a segregation of secondary structure along the polypeptide chain is indicative of the presence of two globular units (32). A combined secondary structure prediction carried out on a number of homologous *B. thuringiensis* toxins differing in insecticidal activity spectrum reveals that this pattern is highly conserved. This is not unexpected for such a set of proteins, as their folding mechanism is most likely under evolutionary pressure. Two stretches in the N-terminal half, predicted to be helices, are highly conserved in primary structure. This feature immediately prompts the question whether these regions are essential for the protein fold, it's stability, or activity.

Taking into account the result of the structure predictions, a careful analysis of the Bt2 toxin denaturation data permits an educated guess on the localization of the structural domains within the primary structure of this protein.

At pH 11, we observe a biphasic denaturation profile using fluorescence spectroscopy. The first transition accounts for the largest part of the total unfolding signal. Under these conditions, CD monitored unfolding reveals an α-helix rich region of decreased stability. Thus, the less stable domain at pH 11 should be rich in α-helix conformation and contain most of the tryptophan residues. As the N-terminal half is predicted to be an α-helix-rich region and meets the requirement of being tryptophan rich, we are convinced that the first transition corresponds to the unfolding of the N-terminal part of the toxin and that the biphasic unfolding behavior is due to the independent denaturation of two structural domains.

As stable fragments of comparable molecular weight are generated upon digestion of the Bt2 toxin with several proteases, additional evidence for a multidomain organization is provided. Referring to studies on thermolysin (33, 34) and the α-subunit of tryptophan synthase (35, 36), it is generally accepted that protease-sensitive sites are located in flexible regions connecting structural domains.

Our model of the domain structure and the unfolding pathway of the Bt2 toxin is presented in Fig. 7.

The possibility that the biphasic denaturation behavior is due to a pH-induced conformational change of the native state, resulting in an alternative unfolding pathway, or to the formation of stable intermediates different from those proposed in Fig. 7 can not *a priori* be excluded. Such an unfolding intermediate would have a conformation which is unrelated to both the native and the completely denatured states (as all

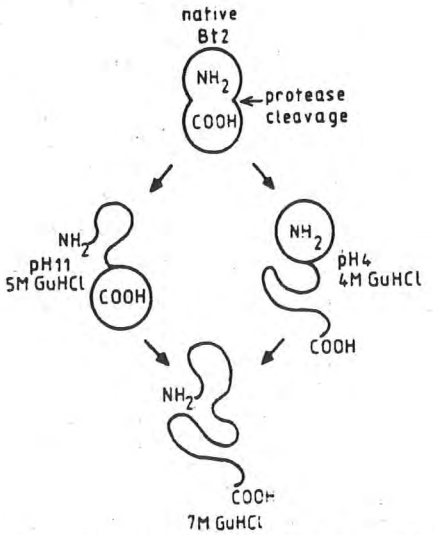


FIG. 7. **Model of the Bt2 toxin domain structure and unfolding pathway.** *GuHCl*, guanidine hydrochloride.

other models are equivalent to the presented unfolding pathway).

Since a biphasic transition is observed at both acid and alkaline conditions, the latter assumption becomes very unlikely. Electrostatic interactions stabilizing an unfolding intermediate at a particular pH are not expected to become exactly replaced at the other extreme of the pH scale, again resulting in a stabilized conformation.

Through comparison of the fluorescence and CD spectra of the native protein recorded at different pH conditions we can also reject a pH-induced conformational change as the underlying mechanism for the altered stability. No red shift of the fluorescence emission spectra is observed between pH 4 and 11. This means that the microenvironment of the tryptophan residues is hardly changed. Moreover the CD spectra of the native protein at pH 8 and 11 are near to identical, indicating a conserved secondary structure.

Several hypotheses can be formulated to explain the pH-dependent stability of the Bt2 toxin and more especially the denaturation of the N-terminal half. As we postulate that the N-terminal half of the protein contains several α-helices, their conformational characteristics are most probably responsible for this phenomenon. The pH-mediated stability of an α-helix can be rationalized by considering its permanent dipole formed by the aligned peptide bonds (37). The helix dipole results in an additional positive and negative electrostatic charge at the N- and C-terminal ends of the helix (38), respectively. Depending on the nature and the ionization state of residues flanking these helices, favorable or unfavorable interactions occur which lead to an increased or decreased stability (39). The gradual protonation of residues at the C-terminal end of the helices in the first part of the toxin could well be responsible for the stabilizing effect at acid pH.

An alternative mechanism by which an α-helix can be stabilized in a pH-dependent manner is by ionic interaction among residues 4 positions apart in the primary sequence (40, 41). Deprotonation of one amino acid side chain would again lead to helix destabilization. Our rationalization of the Bt2 toxin unfolding behavior being correlated to the protein's secondary structure is to our knowledge, quite unique. Recently the pH-dependent stability of barnase (42) was found to be exactly correlated with the helix dipole stabilization effect.

So, apart from the obvious importance of a structural investigation of the toxin as an insecticide, the Bt2 protein could serve as model for protein stability studies.

Isolation of the Bt2 structural domains can be attempted by means of the fragments obtained from protease digests. Their full characterization will provide further evidence for the structural model of the Bt2 toxin described in this paper. The partition of the toxin into functional domains can be examined through the purified structural domains. In this perspective, apart from *in vivo* toxicity tests, the high affinity binding of the Bt2 toxin to membrane components of target insects has been described (14, 15). Recently a phospholipid vesicle perturbation by *B. thuringiensis* toxins has also been reported (43).

Various other bacterial toxins like diphtheria toxin or *Pseudomonas* exotoxin A have been extensively studied with respect to their domain structure (12, 13). For the latter the three-dimensional structure reveals that this toxin is built up of three spatially distinct domains (44) which are related to the protein's functionalities, like receptor binding, membrane translocation, and enzymatic activities (13).

The final confirmation of a similar molecular architecture for the *B. thuringiensis* toxins will have to await the elucidation of its three-dimensional structure through x-ray crystallography. Several groups are involved in the structure determination of a coleopteran-specific toxin (analogous or identical to the Bt13 sequence used in this study) (17, 45). Further information on the features that determine the high insect specificity and the mode of action of the *B. thuringiensis* toxins will also prove indispensable to a rational application of the structural data presented here.

Acknowledgments—We thank Annemie Van Houtven, Adri Van Vliet, and Jan De Meutter for excellent assistance throughout this work, Patrick Stanssens and Chris Opsomer for DNA constructions on the Bt2 clone. Enno Krebbers, Gaston Matthyssens and Herman Van Mellaert are acknowledged for critically reading and improving the manuscript.

REFERENCES

1. Whiteley, H. R., and Schnepf, H. E. (1986) *Annu. Rev. Microbiol.* **40,** 549–576
2. Lilley, M., Ruffell, R. N., and Somerville, H. J. (1980) *J. Gen. Microbiol.* **118,** 1–11
3. Goldberg, L. J., and Margalit, J. (1977) *Mosq. News* **37,** 355–358
4. Krieg, A., Huger, A. M., Langenbruch, G. A., and Schnetter, W. (1983) *Z. Angew. Entomol.* **96,** 500–508
5. Höfte, H., De Greve, H., Seurinck, J., Jansens, S., Mahillon, J., Ampe, C., Vandekerckhove, J., Vanderbruggen, H., Van Montagu, M., Zabeau, M., and Vaeck, M. (1986) *Eur. J. Biochem.* **161,** 273–280
6. Schnepf, H. E., and Whiteley, H. R. (1985) *J. Biol. Chem.* **260,** 6273–6280
7. Chungjatupornchai, W., Höfte, H., Seurinck, J., Angsuthanasombat, C., and Vaeck, M. (1988) *Eur. J. Biochem.* **173,** 9–16
8. Höfte, H. (1988) *Insecticidal Crystal Proteins of B. thuringiensis: Molecular Organization and Expression in Tragenic Plants,* Ph.D. thesis. Rijksuniversiteit Gent, Belgium
9. Visser, B., van Workum, M., Dullemans, A., and Waalwijk, C. (1986) *FEMS Microbiol. Lett.* **30,** 211–214
10. Huber, H. E., Lüthy, P., Ebersold, H. R., and Cordier, J. L. (1981) *Arch. Microbiol.* **129,** 14–18
11. Carey, P. R., Fast, P., Kaplan, H., and Pozsgay, M. (1986) *Biochim. Biophys. Acta* **872,** 169–179
12. Olsnes, S., Moskaug, L. O., Stenmark, H., and Sandvig, K. (1988) *Trends Biochem. Sci.* **13,** 348–351
13. Hwang, J., Fitzgerald, D. J., Adhya, S., and Pastan, I. (1987) *Cell* **48,** 129–136
14. Hofmann, C., Lüthy, P., Hütter, R., and Pliska, V. (1988) *Eur. J. Biochem.* **173,** 85–91
15. Hofmann, C., Vanderbruggen, H., Höfte, H., Van Rie, J., Jansens, S., and Van Mellaert, H. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 7844–7848
16. Sacchi, V. F., Parenti, P., Hanozet, G. M., Giordana, B., Lüthy, P., and Wolfersberger, M. G. (1986) *FEBS Lett.* **204,** 213–218
17. Garfield, J. L., and Stout, C. D. (1988) *J. Biol. Chem.* **263,** 11800–11801
18. Ward, E. S., and Ellar, D. J. (1987) *Nucleic Acids Res.* **15,** 7195
19. Höfte, H., Seurinck, J., Van Houtven, A., and Vaeck, M. (1987) *Nucleic Acids Res.* **15,** 7183
20. Donovan, W. P., Dankocsik, C. C., Gilbert, M. P., Gawron-Burke, M. C., Groat, R. G., and Carlton, B. C. (1988) *J. Biol. Chem.* **263,** 561–567
21. Sobel, E., and Martinez, H. M. (1985) *Nucleic Acids Res.* **14,** 363–374
22. Gibrat, J.-F., Garnier, J., and Robson, B. (1987) *J. Mol. Biol.* **198,** 425–443
23. de Boer, H. A., Comstock, L. J., and Vasser, M. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 21–25
24. Zell, R., and Fritz, H.-J. (1987) *EMBO J.* **6,** 1809–1815
25. Laemmli, U. K. (1970) *Nature* **227,** 680–685
26. Kronman, M. J., and Holmes, L. G. (1971) *Photochem. Photobiol.* **14,** 113–134
27. Greenfield, N., and Fasman, G. D. (1969) *Biochemistry* **8,** 4108–4116
28. Chen, Y.-H., Yang, J. T., and Martinez, H. M. (1972) *Biochemistry* **11,** 4120–4131
29. Marqusee, S., and Baldwin, R. L. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 8898–8902

30. Jeanicke, R. (1987) *Prog. Biophys. Mol. Biol.* **49,** 117–237
31. Privalov, P. L. (1982) *Adv. Protein Chem.* **35,** 1–104
32. Chothia, C. (1984) *Annu. Rev. Biochem.* **53,** 537–572
33. Colman, P. M., Jansonius, J. N., and Matthews, B. W. (1972) *J. Mol. Biol.* **70,** 701–724
34. Fontana, A. (1988) *Biophys. Chem.* **29,** 181–193
35. Higgins, W., Fairwell, T., and Miles, E. W. (1979) *Biochemistry* **18,** 4827–4835
36. Miles, E. W., Yutani, K., and Ogasahara, K. (1982) *Biochemistry* **21,** 2586–2592
37. Hol, W. G. J., van Duijnen, P. T., and Berendsen, H. J. C. (1978) *Nature* **273,** 443–446
38. Hol, W. G. J., Halie, L. M., and Sander, C. (1981) *Nature* **294,** 532–536
39. Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M., and Baldwin, R. L. (1987) *Nature* **326,** 563–567
40. Bierzynski, A., Kim, P. S., and Baldwin, R. L. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 2470–2474
41. Sundaralingam, M., Drendel, W., and Greaser, M. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 7944–7947
42. Sali, D., Bycroft, M., and Fersht, A. R. (1988) *Nature* **335,** 740–743
43. Haider, M. Z., and Ellar, D. J. (1989) *Biochim. Biophys. Acta* **978,** 216–222
44. Allured, V. S., Collier, R. J., Carroll, S. F., and McKay, D. B. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 1320–1324
45. Li, J., Henderson, R., Carroll, J., and Ellar, D. (1988) *J. Mol. Biol.* **199,** 543–544

The influence of the predaceous mite *Typhlodromus (A.) potentillae* (Garman) on the development of *Panonychus ulmi* (Koch) on apple grown under various nitrogen conditions. *Entomophaga* 15: 291–304
154. Van Driesche, R. G. 1983. Meaning of "percent parasitism" in studies of insect parasitoids. *Environ. Entomol.* 12: 1611–22
155. Van Driesche, R. G. 1988. Field measurement of population recruitment of *Cotesia glomerata* (Hymenoptera: Braconidae), a parasitoid of *Pieris rapae* (Lepidoptera: Pieridae), and factors influencing adult parasitoid foraging success in collards. *Bull. Entomol. Res.* 78:199–208
156. Van Driesche, R. G. 1988. Survivorship patterns of *Pieris rapae* (Lep.: Pieridae) larvae in Massachusetts kale with special reference to mortality due to *Cotesia glomerata* (Hymen.: Braconidae). *Bull. Entomol. Res.* 78:397–405
157. Van Driesche, R. G., Bellows, T. S. Jr. 1988. Use of host and parasitoid recruitment in quantifying losses from parasitism in insect populations. *Ecol. Entomol.* 13:215–22
158. Van Driesche, R. G., Bellows, T. S. Jr., Elkinton, J. S., Gould, J., Ferro, D. N. 1991. The meaning of percentage parasitism revisited: solutions to the problem of accurately estimating total losses from parasitism in a host generation. *Environ. Entomol.* 20:1–7
159. Van Driesche, R. G., Bellows, T. S. Jr., Ferro, D. N., Hazzard, R., Maher, M. 1989. A simplified way of using Richards and Waloff's second method to estimate stage survival, with reference to the Colorado potato beetle, *Leptinotarsa decemlineata* (Say). *Can. Entomol.* 121: 291–300
160. Van Driesche, R. G., Gyrisco, G. G. 1979. Field studies of *Microctonus aethiopoides*, a parasite of the adult alfalfa weevil *Hypera postica*, in New York. *Environ. Entomol.* 8:238–44
161. Van Driesche, R. G., Taub, G. 1983. Impact of parasitoids on *Phyllonorycter* leafminers infesting apple in Massachusetts, *USA Prot. Ecol.* 5:303–17
162. Varley, G. C., Gradwell, G. R. 1960. Key factors in insect population studies. *J. Anim. Ecol.* 29:399–401
163. Varley, G. C., Gradwell, G. R. 1968. Population models for the winter moth. In *Insect Abundance*, ed. T. R. E. Southwood, 4:132–42. London: Symp. R. Entomol. Soc. London
164. Varley, G. C., Gradwell, G. R. 1970. Recent advances in insect population dynamics. *Annu. Rev. Entomol.* 15:1–24
165. Varley, G. C., Gradwell, G. R. 1971. The use of models and life tables in assessing the role of natural enemies. In *Biological Control*, ed. C. B. Huffaker, pp. 93–110. New York: Plenum
166. Varley, G. C., Gradwell, G. R., Hassell, M. P. 1973. *Insect Population Ecology*. Oxford: Blackwell
167. Vickery, W. L., Nudds, T. D. 1984. Detection of density-dependent effects in annual duck censuses. *Ecology* 65: 96–104
168. Wylie, H. G. 1981. Effects of collection method on estimates of parasitism and sex ratio of flea beetles (Coleoptera: Chrysomelidae) that infest rape crops in Manitoba. *Can. Entomol.* 113:665–71
169. Wylie, H. G. 1982. An effect of parasitism by *Microctonus vittatae* (Hymenoptera: Braconidae) on emergence of *Phyllotreta cruciferae* and *Phyllotreta striolata* (Coleoptera: Chrysomelidae) from overwintering sites. *Can. Entomol.* 114:727–32
170. Yano, K., Morakote, R., Satoh, M., Asai, I. 1985. An evidence for behavioral change in *Nephotettix cincticeps* Uhler (Hemiptera: Delphacidae) parasitized by pipunculid flies (Diptera: Pipunculidae). *Appl. Entomol. Zool.* 20:94–96

Annu. Rev. Entomol. 1992. 37:615-36

THE MODE OF ACTION OF *BACILLUS THURINGIENSIS* ENDOTOXINS

Sarjeet S. Gill, Elizabeth A. Cowles, and Patricia V. Pietrantonio

Department of Entomology, University of California, Riverside, California 92521

KEY WORDS: toxin structure, midgut, insecticidal activity, insect resistance

INTRODUCTION

The use of synthetic organic insecticides developed during the last half of this century may pose risks to human health and can cause environmental problems. Consequently, interest has developed in using alternative strategies for insect pest management. One contemporary approach that has received attention is the development of *Bacillus thuringiensis* (Bt) toxins as insecticides. *B. thuringiensis*, a gram-positive bacteria, produces a proteinaceous parasporal crystalline inclusion during sporulation. Upon ingestion by insects, this crystalline inclusion is solubilized in the midgut, releasing proteins called δ-endotoxins. These proteins (protoxins) are activated by midgut proteases, and the activated toxins interact with the larval midgut epithelium causing a disruption in membrane integrity and ultimately leading to insect death.

In spite of significant efforts directed towards the study of Bt, its use in pest control is restricted, in part because of the selectivity of Bt and in part because of its moderate efficacy. Attempts to improve these two qualities have met with only moderate success because inadequate effort has been devoted to understanding the molecular basis of selectivity and insecticidal properties of these toxins.

0066-4170/92/0101-0615$02.00

The feasibility of using Bt for insect control has been increased by advances in recombinant DNA technology. Use of this technology has facilitated the cloning of the toxin genes and their expression in plants, in seeds, and in plant- and soil-inhabiting bacteria, thereby providing novel means of toxin delivery to insects. These new developments in Bt insect control, although safe for humans and the environment and compatible with other agricultural practices, will likely encounter problems with insect resistance in the near future. Recent progress in the introduction of Bt toxin genes into plants and other organisms will be of little use if the potential problems of insect resistance are not understood. As a consequence, Bt products must be engineered within the framework of managing the development of insect resistance prior to the widespread use of Bt.

This review addresses several characteristics of Bt toxins that contribute to their modes of action. We present a brief overview of Bt toxin structures, the interaction of these toxins with insect cell membranes, the histopathology of Bt action, and a discussion of insect resistance to Bt.

TOXIN STRUCTURE

Primary Structure

The primary structure of Bt δ-endotoxins varies with the gene that encodes the protein. Numerous Bt toxin genes have now been identified and are classified according to a designation proposed by Höfte & Whiteley (48). For a more detailed discussion, the reader should consult excellent reviews published recently (4, 48). Briefly, the toxins identified to date mostly exist in three size designations, namely 125–138, 65–75, and 25–28 kilodaltons (kd). These toxins are encoded by the *cryI* and *cryIV*; the *cryII*, *cryIII*, and *cryIV*; and the *cyt* genes, respectively (Table 1). This classification of toxin genes is also related to the biological activity of their gene products. Thus, the toxins encoded by the *cryI*, *cryII*, *cryIII*, and *cryIV* genes are toxic to lepidopterans, dipterans and lepidopterans, coleopterans, and dipterans, respectively. All of these *cry* toxin genes are thought to have a common evolutionary origin because of the observed high amino acid homology (48). The *cryII*, *cryIII*, *cryIVC*, and *cryIVD* genes appear to be natural C-terminal truncations of the larger *cryI* and *cryIV* genes. The *cyt* genes and Cyt proteins are significantly different both in their structure and their biological activities from the *cry* genes and Cry proteins.

Tertiary Structure

The δ-endotoxins produced in Bt are packaged into parasporal inclusions visible with light microscopy. The structures of parasporal inclusions vary among subspecies and also, to some extent, according to the conditions used for bacterial culture. Many, if not most, inclusions contain more than one protein; for example, the inclusion body of *B. thuringiensis* subsp. *kurstaki* (Btk) HD-1 comprises five different polypeptides, three CryI and two CryII protoxins (48). Complex interactions, including hydrophobicity, hydrogen bonding, and disulfide bonds, presumably hold these proteins together in the parasporal inclusion (9).

The disulfide bonds are important for the maintenance of the parasporal inclusion structure. For example, in CryI toxins, the C-terminal half of the protoxin contains cysteines that are usually found as cystines. The cysteine content of each of these toxins, usually between 1.1–1.9%, is sufficient for 7–12 disulfide bonds (4, 9). The toxins contain no detectable free sulfhydryl groups (6), illustrating the importance of disulfide linkages in the stabilization of the parasporal inclusion.

B. thuringiensis subsp. *israelensis* (Bti) parasporal inclusions, like those of Btk, have structural disulfide bonds. These parasporal inclusions, which contain at least four different proteins, have a cysteine content of 2.1%. Alkali solubilization alone releases 50% of the parasporal inclusion protein; the larger 125- to 135-kd CryIVA and IVB proteins require disulfide reduc-

Table 1 Classification of Bt crystal protein genes

Gene type[a]	Host range	Number of amino acids	Predicted mol. mass (kd)	Reference
cryIA(a)[1]	L	1176	133.5	75
cryIA(b)	L	1155	131.0	91
cryIA(c)	L	1178	133.3	1
cryIB	L	1207	138.0	8
cryIC(a)	L	1189	134.8	49
cryIC(b)	L	1177	134.0	7
cryID	L	1165	132.5	46
cryIE	L	1171	132.0	89
cryIF	L	1174	133.6	b
cryIIA	L/D	633	70.9	24
cryIIB	L	633	70.8	94
cryIIIA	C	644	73.1	76
cryIIIB	C	659	74.2	77
cryIVA	D	1180	134.4	92
cryIVB	D	1136	127.8	18
cryIVC	D	675	77.8	85
cryIVD	D	643	72.4	23
cytA	D/Cytol	248	27.4	90

[a] Several other gene types have been reported but not published. These include *cryIA(d)*, *cryIG*, *cryIIIC*, and *cytB*.

[b] The nucleotide sequence for the *cryIF* has been submitted for publication, and the putative *cryIIIC* gene has 652 amino acids with a predicted molecular mass of 74.4 (M. C. Gawron-Burke, personal communication). Modified after Höfte & Whiteley (48).

tion for solubilization (20). The 27-kd CytA protein aggregates as a trimer or tetramer, apparently because of disulfide bonds (20).

Differential solubility of distinct Bt parasporal inclusions can be used for partial separation of the toxins. For example, in Btk HD-1, the CryI toxins are soluble at pH 9.5, while the CryII toxins require a pH of about 12 for complete solubilization. Similarly, in Bti the CryIVA, CryIVB, and CytA toxins are soluble at pH 9.5, while the CryIVD toxin requires a pH of 12. The CryIA(c) toxin is soluble at pH values less than 3.5 and above 9.5; between pH 5.0 and 8.5, the solubility is low (5).

These solubilities can be related, in part, to the three-dimensional structure of the CryIA(c) toxin. Raman spectroscopy revealed that 20% of this toxin is α-helix, 35% β-pleated sheet, and 45% unordered (15, 19). Structurally, the Btk HD-73 protoxin consists of two multidomain regions, giving the protoxin unique chemical and biological characteristics (15, 16, 19). The C-terminal region is more sensitive to 8 M urea, while the entire protoxin unfolds in 6 M guanidine hydrochloride (19). Both the toxin and protoxin can refold to their native, active forms upon dialysis. The toxic moiety of the CryIA(c) protoxin and the N-terminal and cysteine-minus half does not appear to undergo a major conformational change upon activation (14). Disulfide bonds, therefore, are apparently not necessary for the proper refolding of the toxin. Although these disulfide bonds may be important for parasporal inclusion stability, they do not appear to be necessary for toxicity of the lepidopteran-specific toxins.

Many Bt toxins are probably glycoproteins (71), although this point is controversial in the literature. The primary structure of many toxins shows that N- and O-glycosylation can occur at multiple sites. Furthermore, these glycosylated moieties may play a role in insect toxicity (67). Whether this is true remains to be proven, although toxin glycosylation sites apparently contribute little, if anything, to insect toxicity because toxins cloned in *Escherichia coli*, which is not known to glycosylate proteins, are generally equitoxic to toxins derived from Bt.

Three-dimensional structures and domains of the toxins have been explored immunologically. These studies show that antisera raised against a particular CryI protein will cross-react with several other CryI proteins, but will not recognize CryII proteins (47, 57), indicating that these two classes of proteins are antigenically distinct. Yet the proteins are structurally related within each group. Cross-reactivity, however, does not necessarily imply a similar toxicity spectrum, whether in vitro or in vivo, so these results must be viewed with caution.

The toxin's mode of action can be attributed, in part, to its three-dimensional structure. Some of the regions of the C-terminal cell-binding domain of the activated Bt toxins (Figure 1) may be needed for the toxin to

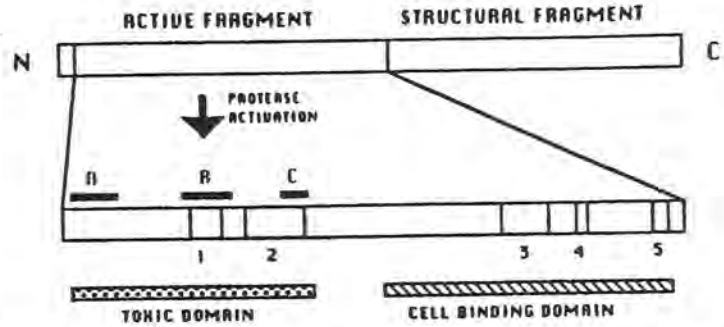


Figure 1 Illustration of the domains of a typical Cry toxin, e.g. CryIA(c), derived from Bt. The 130-kd protoxin is cleaved by proteases, resulting in the loss of the C-terminal structural fragment. The active fragment is formed from the remaining N-terminal half. This active fragment is thought to consist of at least two domains, the toxic and cell-binding domains. Five conserved regions (*shaded*) are found within the active fragment. Deletion of region 5 usually results in a loss of insecticidal activity. The toxic domain contains hydrophobic regions, of which three are identified (*solid dashes*). The putative ion pore formed by the Cry toxins (see Figure 2) is apparently derived from α-helices found within the toxic domain. This toxic domain is thought to contain at least six α-helices (44a). The cell-binding domain consists of the variable and conserved C-terminal regions of the activated toxin. The variable region probably plays a role in receptor binding. The C-terminal conserved region of the active fragment could also be involved in receptor interaction or alternatively be important for structural conformation.

bind specifically to cell receptors. The heterogeneity in the amino acid sequence, however, may be of little importance as long as the correct three-dimensional conformation is maintained (99).

ACTIVATION OF TOXINS

The larger protoxins, of approximately 130–140 kd, all require processing to active toxins (48). This action, mediated by the alkaline pH and proteases of the insect midgut, yields 60- to 70-kd proteinase-resistant toxin fragments that are derived from the N-terminal half of the protoxin (5, 16, 45). The degradation of CryIA(c) protein occurs in an ordered sequence of seven specific cleavages starting at the C terminus of the protoxin and proceeding toward the N-terminal region. C-terminal fragments of approximately 10 kd are produced and rapidly proteolyzed further to small peptides (16). Whether similar precise degradation occurs with other protoxins is not known, but that the 130- to 140-kd toxins give rise to a protease-resistant core is generally recognized. The stability of this core varies with the Bt isolate. The protease-resistant core of the 125- and 135-kd proteins of Bti is less stable than the lepidopteran toxins (18). Nevertheless, the C-terminal truncated toxins are all less stable than the protoxins (4, 48), probably because of the role of the

C-terminal fragment in the formation of parasporal inclusions, which are more stable.

The size of the smallest active fragment, the protease-resistant core, varies with the Bt strain. In general for CryI protoxins, only limited proteolysis occurs at the N-terminal while approximately 500 amino acids are removed from the C-terminal. For example, the CryIA(b) toxic fragment resides between amino acid residues 29–35 and 599–607 (45), and the CryIA(c) toxic fragment lies between residues 29 and 623 (4). Similarly, the toxic fragment of CryIVA protein incorporates residues 30 to 695 (102), while the CryIVB toxic fragment resides between amino acid residues 39 and 677 (69). In contrast, the CryII, CryIII, and CryIVD proteins do not undergo protease-mediated C-terminal cleavage because these proteins appear to be naturally truncated. Even a small C-terminal truncation of 11 amino acids of the CryIIA toxin results in a loss of activity (94). Similarly, C-terminal protease cleavage of both the CryIIIA and CryIVD toxins will likely result in the loss of insecticidal activity because these toxins have C-terminal sequences (conserved domain 5, Figure 1) that are required for toxicity in several other CryI-activated toxins (10, 94). However, these naturally truncated proteins may undergo protease cleavage in the insect midgut. Removal of up to 159 residues from the N terminus of the CryIIIA toxin does not decrease biological activity (10, 66). In the mosquito midgut, the CryIVD toxin undergoes proteolysis resulting in products of 32–40 kd, which are toxic to mosquito larvae (12, 13). The N-terminal protease cleavage site for the CryII toxin, if any, is not known. The 27-kd CytA cytolytic toxin is also cleaved to an active product of 24 kd resulting from both N- and C-terminal proteolysis (3, 34, 84).

The midgut environment can also play a crucial role in specificity, as shown with the activation of *B. thuringiensis* subsp. *aizawai* (Bta) strain IC1. When activated with lepidopteran *Pieris brassicae* midgut extract, the toxin kills both *P. brassicae* and dipteran *Aedes aegypti* larvae; however, when activated by *A. aegypti* midgut extract, the isolate is toxic only to these mosquito larvae (39).

DOMAINS AND TOXICITY

The activated toxin can be divided into three structural regions: a N-terminal region, the toxic domain (amino acids 1–279), consisting of several conserved hydrophobic regions; a conserved C-terminal region (amino acids 461–695); and a variable region between these two regions that contains most of the amino acid differences (4, 14, 48). The N-terminal hydrophobic toxic domain has a high degree of α-helicity with several predicted transmembrane regions (Figure 1). Among several known toxins, one observes conservation of hydrophobicity rather than conservation of amino acids, suggesting that this hydrophobic domain probably has a significant role in toxicity. Computational analyses of this hydrophobic toxic domain show that at least six major α-helices can be identified for most of the Cry toxins (14, 19, 28, 44a). Modification of one of the α-helices by site-directed mutagenesis that resulted in a change in the net charge caused a decrease in toxicity (99). Thus the hydrophobic toxic domain, the α-helices in particular, is probably crucial for the creation of the amphipathic transmembrane spanning channels that are proposed to be critical components in the manifestation of toxicity.

In contrast, the variable and conserved C-terminal regions appear to be more critical in the selectivity of the toxin, which is determined by the binding of the toxin to its receptor in the midgut of the target insect. These variable and conserved C-terminal regions of the activated toxins consist primarily of β-sheets (14, 19, 28). The C-terminal fragment of the activated CryIA(c) toxin from amino acid 322 contributes to *Helicoverpa virescens* selective toxicity. Substitution of a few amino acids between residues 429 and 447 of CryIA(c) toxin with the corresponding amino acids from the CryIA(a) toxin markedly decreased the toxicity of the recombinant CryIA(c) toxin to *Manduca sexta*, while the toxicity towards *H. virescens* diminished slightly (74). For *Bombyx mori*, however, residues 332–450 contribute to the increased toxicity of the CryIA(a) toxin (32). The specificity of CryIA(b) toxin, from Bta strain IC1, depends on the midgut protease processing of the toxin within this variable domain. Changes in the identity of residues adjacent to Arg544 and Arg567 can convert a monospecific lepidopteran toxin to a dual specific lepidopteran/dipteran toxin (39). The regions modified above are probably involved more in selectivity rather than in exerting the toxic effect, although this was not determined in these studies.

The toxic and binding domains appear to be folded in their active conformation even in the protoxin because no apparent change in the conformational spectra of these domains is observed when the protoxin is activated (14, 19). Moreover, the activated toxin appears to consist of at least two domains, as determined by protease cleavage, one for toxicity and the other for receptor binding and toxin structure (14, 16, 19). Structurally the Cry toxins, therefore, appear to be very similar to the *Pseudomonas aeruginosa* exotoxin (2), which has three discrete domains, each with a specific function.

The CytA toxin has several hydrophobic regions packaged in either an antiparallel or parallel helix bundle, which facilitates the insertion of this toxin into cell membranes (65). Potentially, the predicted coil-turn loop region adjacent to helix 3 and 4 and adjacent β-sheets (93) fit these structural requirements, and this region could be important for binding and toxicity. Indeed monoclonal antibodies directed against this region inhibit the binding of the CytA toxin to cell membranes (S. S. Gill, unpublished data).

MODE OF ACTION

Toxin Binding and Membrane Interaction

CRY TOXINS Although the primary action of Bt toxins is in the midgut epithelium of sensitive insects, investigators have only recently demonstrated that the toxin binds specifically to brush border membrane vesicles (BBMV) prepared from the midgut. The BBMV consist primarily of the apical brush border membrane of the midgut columnar cells. The BBMV studies (43, 44, 86–88) show a positive correlation between the biological activity of the Cry toxins and their ability to bind to BBMV of susceptible larvae. These Cry toxins require a specific plasma membrane receptor on the midgut epithelial cells. The receptor affinity of these Cry toxins can vary from relatively low to high, with K_d values between 10^{-7} and 10^{-10} M (Table 2). Furthermore, CF1 cells, although not derived from midgut cells, appear to have a low-affinity receptor (E. Chow & S. S. Gill, unpublished data). These studies show that insect cells probably have multiple toxin receptors. Whether these K_d values reflect distinct receptor subunits or different receptors remains to be ascertained. Further, the toxicity of the Cry toxins appears to be correlated with receptor number rather than receptor affinity (86, 87), that is, the higher the receptor concentration the greater the toxicity. A more recent study, however, shows that at least with the gypsy moth, *Lymantria dispar*, there is no correlation between receptor concentration and toxicity but rather a negative correlation between receptor affinity and toxicity (97). Therefore, a straight correlation between toxicity and toxin binding, according to either receptor affinity or receptor concentration, may not always be a generally applicable concept.

One active area of investigation involves the characterization of the receptor. Researchers believe that the receptor is a glycosylated protein (28, 43, 54), and it has been proposed that receptor glycosylation plays an essential role in toxin binding and toxicity (28). However, evidence that the glycosylated moieties of the receptor are actually involved in toxin binding is equivocal. Several putative receptors have been identified. The CryIA toxin receptor for the CF1 cell line was identified as a 146-kd protein (54), while the receptors for the lepidopteran and dipteran forms of Bta toxin were putatively identified as proteins of 120 or 68 kd and 90 kd from lepidopteran and dipteran cells, respectively (38). Whether these cell-line receptors are in fact the same receptors as in the insect midgut is uncertain because the cell lines are not derived from midgut epithelial cells, and thus, these data must be evaluated with caution. Therefore, the elucidation of the toxin receptor and characterization of its function in the insect midgut and in other cells must wait until the molecular cloning and/or isolation of the receptor is achieved.

Table 2 Binding characteristics of Cry toxins to the BBMV from the midgut of various insect species[a]

Parameters[b] / Species	Toxin type: CryIA(a) LC_{50}	K_d	R_t	CryIA(b) LC_{50}	K_d	R_t	CryIA(c) LC_{50}	K_d	R_t	CryIB LC_{50}	K_{d1} (K_{d2})	R_{t1} (R_{t2})[c]	CryIC LC_{50}	K_{d1} (K_{d2})	R_{t1} (R_{t2})	CryIE LC_{50}	K_{d1} (K_{d2})	R_{t1} (R_{t2})	References
Manduca sexta	20	1.1	9.9	20	0.2	7.9	9	0.2	6.3	>625[d]	Nonsaturable binding		111	0.4	8.95	73	0.09	2.97	44, 86
														(7.05	9.44)		(216	283)	87
		1.5[e]	9.8[e]		0.4[e]	3.4[e]		0.6[e]	4[e]										
		1.48	19.9		0.43	4.3													
Helicoverpa virescens	157	0.8	3.7	7	0.4	21	2	0.4	62.3				>2700	22.4	37.9	>2700	NB[f]	NB	86, 87
		1.2[e]	3.7[e]		0.6[e]	9.7[e]		0.8[e]	19.5[e]										
		1.16	3.5																
Spodoptera littoralis	>1,350	NB	NB										93	0.18	2.04	88	1.18	4.73	87
														(13.9	38.9)		(352	230)	
Lymantria dispar				1.08	19.8	2.7	425	2.03	3.69										97
Pieris brassicae				0.9[g]	3.2	13.6				4[h]	46	0.2							43, 44
											(490	30)							
											47.6	18.8							
Plodia interpunctella																			
susceptible				0.03[i]	0.72	1.44							0.11[i]	0.31	0.38				
														(154	6.17)				
resistant to CryIA(b)				26.3[i]	36.3	1.77							0.03[i]	0.18	1.15				88

[a] Data estimated from competition experiments, except where indicated.
[b] LC_{50}, ng activated toxin/cm² diet unless otherwise indicated; K_d, nM of activated toxin; R_t, pmol/mg BBMV protein; K_{d1}, R_{t1}, parameters of a high affinity binding site.
[c] K_{d2}, R_{t2}, Parameters of a low affinity binding site.
[d] Protoxin LC_{50}.
[e] Data estimated from equilibrium binding studies.
[f] No binding.
[g] μg Protoxin/ml, leaf disc assay (5 μl/disc); 1st instar larvae.
[h] LD: 43% mortality w/4 ng per larva; 93% mortality w/6 ng per larva.
[i] LD_{50} μg toxin/larva.

CYT TOXINS The Cyt toxins of mosquitocidal Bt display a mechanism of cell-membrane interaction different from that of the Cry toxins (17, 34, 53, 84). Initial binding of this Cyt toxin occurs with unsaturated phospholipids (34, 84); the toxin has no effect on bacterial protoplasts, which are devoid of phosphatidylcholine, sphigomyelin, and cholesterol (84). Acyl unsaturation of the *syn*-2 position of the phospholipid is essential for this initial binding; however, the head group is relatively unimportant (34). The toxin binding is nonsaturable at toxin concentrations at which the cells do not lyse. Initially, the toxin binds as a monomer; upon continued exposure, aggregates of the toxin, which occur after a time lag, form in the cell membrane (17, 60). The aggregates are approximately 300–400 kd, which is equivalent to 16 toxin molecules if the aggregates consists solely of the CytA toxin (17). These aggregates then apparently lead to the formation of pores resulting in cytolysis (17, 33, 53, 60, 61). Membrane-bound toxin associates with free toxin, so the pores are not formed by lateral diffusion. At low temperatures, the toxin binds but does not aggregate to form a pore (60, 61). The pores that are formed by the CytA toxin are monovalent cation-selective channels as demonstrated in lipid bilayers (53). The Cyt toxins of mosquitocidal Bt appear to act in a manner similar to that observed with other membrane-disrupting toxins, some of which aggregate on cell membranes (68, 83).

SYNERGISM The mixture of toxins derived from Bti have a higher mosquitocidal activity than the sum of toxicities of the individual toxins. This increased mosquitocidal activity results from the synergistic action of some of the toxins. The CytA toxin displays greater capacity to synergize the action of the CryIV toxins of Bti (13, 50, 100). In addition, the CryIVB toxin requires the presence of the CryIVC toxin for mosquitocidal activity against *Culex quinquefasciatus* (22). Although the mechanism of synergism has not been elucidated, the synergistic action observed probably results from the promotion of hydrophobic interactions between the toxins and between the toxins and the midgut epithelial cell membrane.

Ion Regulation

To better elucidate the mode of action of Bt, one must understand ion regulation in the insect midgut. The epithelial apical membrane of lepidopteran midgut actively transports K^+ ions, thereby creating relatively large membrane potentials; the gut lumen is electrically positive (>180 mV) to both the cell and hemolymph sides (40). This K^+ transport and potential gradient is sustained by an electrogenic V-type ATPase pump that actively pumps H^+ ions into the gut lumen; K^+/H^+ exchange occurs through an electroneutral antiport (95). This proton pump, which is located in the apical membrane of the goblet cell cavity (25), also appears to be involved in nutrient absorption and pH regulation (for a review of ion transport, see 25).

DISRUPTION OF AMINO ACID TRANSPORT The uptake of amino acids across the midgut epithelium of lepidopteran larvae depends on the existence of these transepithelial electrical potential differences. The uptake is carried out through a voltage-dependent symport mechanism, located in the apical border of the columnar epithelial cells, which allows the transport of K^+ ions and amino acids into the cytoplasm (35). Other monovalent cations can effectively replace K^+ for amino acid transport in vitro (96). Six independent amino acid transport systems have been identified in the lepidopteran midgut epithelium, and two different transporters, or at least a transporter with different affinities for Na^+ and K^+ (35, 96), might function in phenylalanine uptake.

Activated Bt toxins apparently insert in the midgut membrane and increase the K^+ conductance of the columnar cell apical membrane. This action leads to disruption of the electrical, K^+, and pH gradients (40, 73). Consequently, the CryIA(c) toxin inhibits, in a dose-dependent fashion, the Na^+ and K^+ gradient-dependent phenylalanine uptake in BBMV of *M. sexta* larvae (96). Similar results were reported for the effect of the CryIA(b) toxin on alanine uptake by *M. sexta* BBMV (42). CryIB protein, which is not toxic to *M. sexta*, had no effect on the amino acid uptake (42). Preincubation of *M. sexta* BBMV with CryIA(c) or CryIA(b) toxins increased both Na^+ and K^+ permeability, while a mixture of CryIA toxins increased only the K^+ permeability of *P. brassicae* BBMV (42, 73, 96).

Neither Ba^{2+} nor Ca^{2+}, two K^+ channel blockers, have any effect on the toxin-induced inhibition of amino acid uptake in *M. sexta* BBMV (42, 96). However, in the isolated *M. sexta* midgut, these two cations prevent the fall in toxin-induced short-circuit current, probably by blocking the pore in the membrane produced by the toxin (21). Differences in these results have been explained by the fact that the BBMV consist only of the apical microvillar borders of the columnar cells, which may limit the structural and biochemical capacity of this tissue to respond to pore formation (96). Alternatively, the protection of Ba^{2+} and Ca^{2+} might be nonspecific and occur through the closure of gap junctions between goblet and columnar cells (42).

ION-PORE FORMATION Three hypotheses on the mode of action of Bt toxins have been proposed in recent years. First, these Bt toxins directly inhibit plasma membrane ATPases (30); second, Bt toxins increase K^+ permeability selectively (40); or third, the toxins form a nonspecific pore (55, 59).

Btk toxin reportedly inhibits a purified dog kidney Na^+/K^+ ATPase in vitro (30), but the primary site of Bt action is probably not on the midgut ATPase because these ATPases are not located on the outside of the apical membrane of midgut columnar cells (25). In *H. virescens*, midgut BBMV possess a vanadate and CryIA(c) toxin-sensitive phosphatase. The phosphorylation of a 72-kd protein is inhibited in vitro both by vanadate and the toxin, suggesting

that the inhibition of the 72-kd protein phosphorylation is responsible for the decrease in phosphatase activity. This putative 72-kd phosphatase is not an ATPase (31).

Several researchers have investigated the hypothesis that the CryIA toxins act either to open up K^+ channels or to form new channels. Ba^{2+} and Ca^{2+}, which are inhibitors of neuronal K^+ channels, blocked the action of the toxin on isolated *M. sexta* midguts, while other ions did not (21). However, as indicated above, these ions did not block the amino acid transport action of the CryIA(c) toxin.

Finally, some researchers have proposed that Bt toxins act by forming nonspecific cation pores (55, 59). Confirmation of pore formation has also been demonstrated in several model membranes. The CytA toxin forms channels in a phosphatidylethanolamine planar bilayer (53); the channels seemed to open and close in synchrony and were selective for K^+ and Na^+, but were not Cl^- permeable. The presence of Ca^{2+} reduced the number and frequency of the channel openings. Liposomes of the target phospholipids for CytA and CytB toxin became leaky to $^{86}Rb^+$ when exposed to the toxins (26). CryIA(c) and CryIIIA toxins made gated channels that lacked a strong voltage dependence and were highly cation selective (79). Moreover, some Cry toxins derived from *B. thuringiensis* subsp. *entomocidus* (103) could also cause the release of markers from loaded liposomes. However, the toxin apparently does not enter the cell nor does it exert its toxicity via second messengers because no changes in the levels of cAMP or other nucleotides were observed upon Bt intoxication (56).

Analysis of the Cry toxins reveals that the putative membrane-spanning amphipathic helices located in the N-terminal hydrophobic toxic domain might be involved in pore formation (28). Any introduction of charged amino acid residues into this conserved hydrophobic domain, for example, amino acid residues 162–171 (Figure 1) of the CryIA(c) toxin, results in lower insecticidal activity (99). This observation has led to the speculation that Cry toxins also form or interact with K^+ channels (99) since this region has some homology to the *Drosophila melanogaster Shaker* protein, which is a K^+ channel–forming protein (52). More recent evidence on K^+ channels does not support this hypothesis because the homologous region in the *Shaker* protein is apparently not involved in pore formation (101).

More likely, Bt toxins form a nonselective pore. The radius of this Bt pore, 0.6–1.0 nm (55), is larger than the crystal radii of 0.13, 0.10, and 0.14 nm for K^+, Na^+, and NH_4^+, respectively (80). However, although the Bt toxin pore appears to be a nonselective ion pore, some ion selectivity could be achieved by a gating mechanism.

Thus the working model of Bt toxin action (Figure 2) is consistent with the formation of a nonselective pore (28, 53, 55, 59). This model consists of

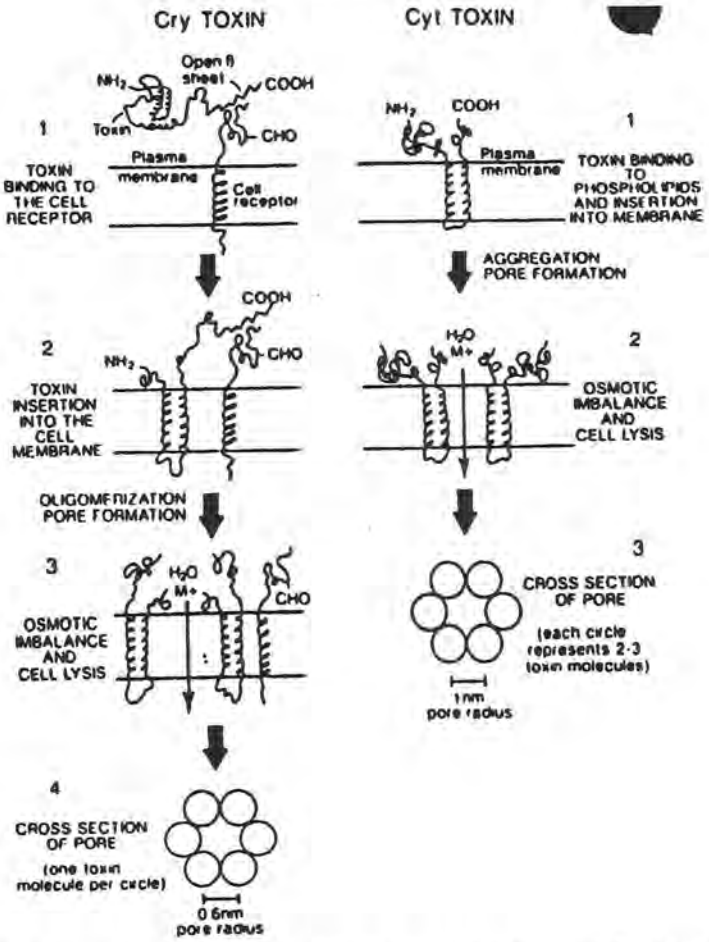


Figure 2 Proposed model for the formation of the Bt toxin-induced pores. (*A*) Putative mechanism by which the Cry toxins form a pore. 1. The cell membrane–binding domain of the toxin binds to a high affinity receptor on the apical membrane of the insect midgut columnar cells. The carbohydrate moiety of the receptor may not be involved in binding toxin, as e.g. in *H. virescens*. 2. The toxic domain of the toxin inserts into the cell membrane after a change in toxin conformation (19). 3. Oligomerization of toxin molecules in the cell membrane, resulting in the formation of a pore, which leads to osmotic imbalance resulting from the influx of water and cations (M^+) or other small molecules. 4. Cross-section of the resulting pore; each circle represents one toxin molecule. (*B*) Putative mechanism for pore formation by Cyt toxins. 1. The C-terminal hydrophobic region of the Cyt toxin inserts into the cell membrane following hydrophobic interactions between the toxin and the membrane phospholipids. 2. Following this initial binding, aggregation of the Cyt toxin molecules occurs in the cell membrane, resulting in the formation of a pore. Cell lysis occurs because of osmotic imbalance resulting from the influx of water and cations (M^+) and other small molecules. 3. Cross-section of the Cyt toxin pore; each circle represents 2–3 molecules of toxin, with 12–18 toxin molecules per pore. Whether the Cry and Cyt toxin pores are formed from the α-helices or β-sheets is not known, although α-helices are depicted in the formation of pores in both *A* and *B*. Modified after Ellar (28) and Wolfersberger (98).

toxin molecules interacting to form a membrane spanning channel. Although the α-helices of the N-terminal hydrophobic toxic domain appear to be the most likely regions to form this channel, at this stage we cannot exclude the possibility that β-sheets form membrane-spanning channels as observed with the K^+ channel (101).

Both the Cry and Cyt toxins may ultimately act by similar mechanisms, but different cell-membrane components are required for the initial stages of interaction between these toxins and the cell membrane. The Cry toxins bind to a midgut glycoprotein receptor, while unsaturated phospholipids act as "binding sites" for the Cyt toxins. Subsequently, the toxin molecules insert into the plasma membrane to form a pore that is permeable to small ions and molecules. These pores will enlarge from osmotic swelling, and eventually cell lysis will occur.

Histopathological Effects on the Midgut

Heimpel & Angus (41) provided the first concise account of the mode of action of Bt. They noted that the insect midgut membranes were disrupted by the δ-endotoxin, allowing an ionic flow into the hemolymph, and classified the susceptible insects into three types. In type I insects, which display extreme sensitivity to the toxin, mortality appears to result from equilibration of the basic hypertonic contents of the midgut into the lightly buffered hemolymph. General body and gut paralysis is observed and death occurs in 1–7 h (41). In type II insects, death occurs in 2–7 days, but pH changes do not occur in the blood; only gut paralysis is observed (41). Type III insects are not killed by the toxin but death results from septicemia after spore germination in the midgut.

The overall midgut pathology of Bt toxicity results in a loss of basal involutions in the columnar cells; swelling of the apical microvilli; vesiculation of the endoplasmic reticulum; loss of ribosomes; swelling of mitochondria; swelling of the cell and nucleus and subsequent rupture of nuclear, organelle, and plasma membranes; and finally release of the cell content into the lumen with sloughing of the cells (11, 27, 29, 36, 78). Other signs include increase in the number and size of nuclear pores (70), separation of the cells from each other and from the basement membrane (11, 78), and nearly complete destruction of the goblet cell (36).

Using electron probe X-ray analysis, investigators observed that the toxin produces cellular changes in ionic concentrations, primarily in the goblet cell cavity. These changes and the increase in total cell Ca^{2+} were hypothesized to be consequences of a direct inhibition of the K^+ pump in the goblet cell apical membrane (37).

The presence of isolated damaged cells in the midgut suggests that direct contact with the toxin is required, that the symptoms do not diffuse laterally, and that the damage appears to progress from the apical to the basal [illegible] (70). These observations are supported by immunofluorescence studies on the binding of CryIA(c) toxin in the midgut and Malpighian tubules of *H. virescens* (72). Because the CryIA(c) toxin binds to several epithelial cells, these investigators also speculated that the toxin receptor may be a common protein in the apical membrane of epithelial cells.

Btk induces a cytoplasmic response in the midgut columnar cells of *B. mori* within one minute after ingestion of crystals. The damage is first observed in the mid to apical region of the columnar cell with loss of microfibrillar cores in microvilli. After 5 min, columnar cells exhibited gross structural damage. With time, bulbous membrane eversions appeared, as microvilli disappeared from the apical border. The goblet cells, however, did not show any structural damage up to 5 min after the ingestion of the toxin (70).

The midgut pathology of *M. sexta* larvae fed with alkali-solubilized Btk endotoxin is similar to Btk-treated *B. mori*. In addition, Golgi-associated membranes appear to be instantaneously susceptible, showing accumulation or fusion of vacuoles. Evidence does not indicate endocytosis, but large amounts of lysosomes accumulate with time, arising by autophagia. The intercellular septate and gap junctions on the lateral borders are disrupted and often internalized, but the basal membranes are not affected (58).

Upon ingestion of Bti, mosquito larvae exhibit the first symptoms within 30 min—the enlargement of intracellular and intercellular spaces on the basal side of the midgut epithelium (11) and the disappearance of granular material between the peritrophic membrane and the epithelial cells (78). This initial appearance of intercellular spaces could result from a loss of gap junctions in the epithelial layer. The mosquito larvae cease to feed at this juncture. Coincidentally, with the initial swelling comes a decrease in the number of microvilli. Upon continued exposure, the ultrastructure of cellular organelles, including the endoplasmic reticulum and mitochondria, change considerably, ultimately resulting in cell lysis (11, 78).

RESISTANCE

Although early attempts to select resistance to Bt in the laboratory were unsuccessful, recent studies show that insects have the capacity to develop resistance to the toxins. McGaughey (62) showed that further laboratory selection of the lepidopteran *Plodia interpunctella* that had been exposed to Btk (Dipel) in the field resulted in development of a 27-fold resistance in two generations and a 97-fold resistance after 15 generations of selection. Furthermore, the resistance developed was stable and inherited as a recessive trait. Additional selection pressure on this resistance colony resulted in an even higher level of resistance, 250-fold, after which the resistance level reached a

plateau (63). This resistant colony also showed resistance to different isolates of Btk, *B. thuringiensis* subsp. *thuringiensis*, and *B. thuringiensis* subsp. *galleriae* (64). The resistance developed in *P. interpunctella* did not result from differential processing of the Bt toxin in resistant and susceptible insects (51), but rather from a decrease in toxin affinity K_d (88); however, the receptor number was unchanged in both insect populations (Table 2). Surprisingly, this resistant colony was more susceptible to another Bt toxin, CryIC. This enhanced susceptibility apparently resulted from an increase in receptor-number concentration (88).

This rapid development of resistance, after only two generations of selection, apparently resulted from prior exposure to Bt in the field. Using a similar approach, S. M. Dai & S. S. Gill (unpublished studies), showed that *Culex quinquefasciatus* rapidly developed resistance. Selection of a *C. quinquefasciatus* colony that had been previously exposed for 98 generations to crude Bti powder with just one toxin, CryIVD, resulted in rapid development of 70-fold resistance after just three generations (S. M. Dai & S. S. Gill, unpublished data). In contrast, when a previously unselected field-collected population was exposed to a single toxin, the development of resistance in *C. quinquefasciatus* was much slower—15-fold resistance after 22 generations (S. M. Dai & S. S. Gill, unpublished). Similarly, the resistance level developed slower in a previously unselected population of *H. virescens*, in which a resistance level of 24-fold was achieved after 7 generations with no further increase, even after 14 generations (81).

Field populations with differing levels of Bt resistance were recently observed. The diamondback moth, *Plutella xylostella*, which in Hawaii was treated with Btk 50–100 times during 1978–1982, had up to 36-fold resistance (82).

CONCLUSION

Future research into the mode of action of Bt toxins must focus on the characterization of the insect receptors, the toxin structures that are involved in selectivity and toxicity, and the potential development of insect resistance to Bt.

The elucidation of toxin receptors is crucial to the overall understanding of the mode of action of Bt toxins. Such a characterization will also facilitate studies to determine the role of the putative Bt receptors in the insect midgut. Moreover, once the Bt receptors are characterized, an evaluation of the receptor diversity in insects and an analysis of their distribution will provide researchers with a unique opportunity to engineer Bt toxins.

Also, the domains or epitopes involved in insect selectivity, i.e. receptor binding, must be fully characterized. While recombinant toxins provide a means to characterize these domains, complete mapping of the domains will probably require the use of X-ray crystallization techniques and/or monoclonal antibodies. Utilization of these techniques should yield clues as to the active sites of toxins and their role in toxicity. Similarly, the elucidation of the domains involved in causing toxicity, in conjunction with a recognition of the domains involved in Bt selectivity, will expedite attempts to engineer more potent or chimeric toxins with a predetermined host range.

An appreciation of structure-function relationships between the toxin and the insect will also prevent the widespread occurrence of resistance in insects. Exposure to complex mixtures of toxins, as that found in Btk HD-1 and Bti, likely results in the selection of a gene pool that exhibits low-level resistance to individual toxins, most probably the more insecticidal toxins. However, the presence of such resistance would be difficult to detect unless toxicity assays were performed with single toxins. For example, exposure of *C. quinquefasciatus* to Bti over several generations will likely result in the selection of populations with lower sensitivity to the most mosquitocidal toxin, CryIVD.

Exposure of insect populations to single toxins, after prior exposure to complex mixtures of toxins (which comprise the bulk of current Bt formulations) can result in the rapid increase of resistance. Therefore, the use of single toxins will lead to more rapid development of resistance and the likely failure of Bt as an insect-control agent in the long run. To delay the onset of resistance, Bt formulations must contain toxins that interact either with multiple receptor sites or that have different modes of action. The latter approach is likely to be more reliable because resistance, unlike that observed with *P. interpunctella*, can develop at numerous sites within the cell, such that receptor recognition by toxins may become irrelevant. An additional alternative would be to have more precise targeting, via transgenic organisms, of Bt toxins such that only a small section of the insect population, i.e. that causing an economic loss, is exposed to the toxin.

Finally, although Bt provides an attractive alternative to chemical insecticides, totally substituting Bt for the use of chemicals would be a mistake. Use of the chemical arsenal in combination with Bt would probably enable a more judicious use of both, and would also delay the onset of insect resistance.

ACKNOWLEDGMENTS

While we have tried to review all the relevant literature, we had to limit our citations, and thus may have not cited some papers out of necessity. Original work reported herein was supported in part by grants from the NIH (ES03298) and the University of California Mosquito Research and Biotechnology Programs. The assistance of S. M. Dai and Y. M. Yu is gratefully acknowledged.

Literature Cited

1. Adang, M. J., Staver, M. J., Rocheleau, T. A., Leighton, J., Barker, R. F., Thompson, D. V. 1985. Characterized full-length and truncated plasmid clones of the crystal protein of *Bacillus thuringiensis* subsp. *kurstaki* HD-73 and their toxicity to *Manduca sexta*. *Gene* 36: 289–300
2. Allured, V. S., Collier, R. J., Carroll, S. F., McKay, D. B. 1986. Structure of exotoxin A of *Pseudomonas aeruginosa* at 3.0 Angstrom resolution. *Proc. Natl. Acad. Sci. USA* 83:1320–24
3. Armstrong, J. L., Rohrmann G. F., Beaudreau, G. S. 1985. Delta-endotoxin of *B. thuringiensis* subsp. *israelensis*. *J. Bacteriol.* 161:39–46
4. Aronson, A. I., Beckman, W., Dunn, P. 1986. *Bacillus thuringiensis* and related insect pathogens. *Microbiol. Rev.* 50:1–24
5. Bietlot, H., Carey, P. R., Choma, C., Kaplan, H., Lessard, T., Pozsgay, M. 1989. Facile preparation and characterization of the toxin from *Bacillus thuringiensis* var. *kurstaki*. *Biochem. J.* 260:87–91
6. Bietlot, H. P. L., Vishnubhatla, I., Carey, P. R., Pozsgay, M., Kaplan, H. 1990. Characterization of the cysteine residues and disulphide linkages in the protein crystal of *Bacillus thuringiensis*. *Biochem. J.* 267:309–15
7. Bosse, M., Masson, L., Brousseau, R. 1990. Nucleotide sequence of a novel crystal protein gene isolated from *Bacillus thuringiensis* subspecies *kenyae*. *Nucleic Acids Res.* 18:7443
8. Brizzard, B. L., Whiteley, H. R. 1988. Nucleotide sequence of an additional crystal protein gene cloned from *Bacillus thuringiensis* subsp. *thuringiensis*. *Nucleic Acids Res.* 16:4168–69
9. Bulla, L. A. Jr., Bechtel, D. B., Kramer, K. J., Shethna, Y. I., Aronson, A. I., et al. 1980. Ultrastructure, physiology, and biochemistry of *Bacillus thuringiensis*. *CRC Crit. Rev. Microbiol.* 8:147–204
10. Carroll, J., Li, J., Ellar, D. J. 1989. Proteolytic processing of a coleopteran-specific δ-endotoxin produced by *Bacillus thuringiensis* var. *tenebrionis*. *Biochem. J.* 261:99–105
11. Charles, J.-F., de Barjac, H. 1983. Action des cristaux de *Bacillus thuringiensis* var. *israelensis* sur l'intestin moyen des larves de *Aedes aegypti* L., en microscopie electronique. *Ann. Microbiol. Inst. Pasteur* 134A:197–218
12. Cheung, P. Y. K., Hammock, B. D. 1986. Separation of three distinct activities from the parasporal crystal of *Bacillus thuringiensis* var. *israelensis*. *Curr. Microbiol.* 12:121–26
13. Chilcott, C. N., Ellar, D. J. 1988. Comparative toxicity of *Bacillus thuringiensis* var. *israelensis* crystal proteins in vivo and in vitro. *J. Gen. Microbiol.* 134:2552–58
14. Choma, C. T., Kaplan, H. 1990. Folding and unfolding of the protoxin from *Bacillus thuringiensis*: evidence that the toxic moiety is present in an active conformation. *Biochemistry* 29:10971–77
15. Choma, C. T., Surewicz, W. K., Carey, P. R., Pozsgay, M., Kaplan, H. 1990. Secondary structure of the entomocidal toxin from *Bacillus thuringiensis* subsp. *kurstaki* HD-73. *J. Protein Chem.* 9:87–94
16. Choma, C. T., Surewicz, W. K., Carey, P. R., Pozsgay, M., Raynor, T., et al. 1990. Unusual proteolysis of the protoxin and toxin from *Bacillus thuringiensis*: structural implications. *Eur. J. Biochem.* 189:523–27
17. Chow, E., Singh, G. J. P., Gill, S. S. 1989. Binding and aggregation of the 25-kilodalton toxin of *Bacillus thuringiensis* subsp. *israelensis* to cell membranes and alteration by monoclonal antibodies and amino acid modifiers. *Appl. Environ. Microbiol.* 55:2779–88
18. Chungjatupornchai, W., Höfte, H., Seurinck, J., Angsuthanasombat, C., Vaeck, M. 1988. Common features of *Bacillus thuringiensis* toxins specific for diptera and lepidoptera. *Eur. J. Biochem.* 173:9–16
19. Convents, D., Houssier, C., Lasters, I., Lauwereys, M. 1990. The *Bacillus thuringiensis* δ-endotoxin. Evidence for a two domain structure of the minimal toxic fragment. *J. Biol. Chem.* 265:1369–75
20. Couche, G. A., Pfannenstiel, M. A., Nickerson, K. W. 1987. Structural disulfide bonds in the *Bacillus thuringiensis* subsp. *israelensis* crystal. *J. Bacteriol.* 169:3281–88
21. Crawford, D. N., Harvey, W. R. 1988. Barium and calcium block *Bacillus thuringiensis* subspecies *kurstaki* δ-endotoxin inhibition of potassium current across isolated midgut of larval *Manduca sexta*. *J. Exp. Biol.* 137:277–86
22. Delecluse, A., Bourgouin, C., Klier, A., Rapoport, G. 1988. Specificity of action on mosquito larvae of *Bacillus thuringiensis israelensis* toxins encoded by two different genes. *Mol. Gen. Genet.* 214:42–47
23. Donovan, W. P., Dankocsik, C. C., Gilbert, M. P. 1988. Molecular characterization of a gene encoding a 72-kilodalton mosquito-toxic crystal protein from *Bacillus thuringiensis* subsp. *israelensis*. *J. Bacteriol.* 170:4732–38
24. Donovan, W. P., Dankocsik, C. C., Gilbert, M. P., Gawron-Burke, M. C., Groat, R. G., et al. 1988. Amino acid sequence and entomocidal activity of the P2 crystal protein. *J. Biol. Chem.* 263:561–67; 1988. Erratum. *J. Biol. Chem.* 264:4740
25. Dow, J. A. T. 1986. Insect midgut function. *Adv. Insect Physiol.* 19:187–328
26. Drobniewski, F. A., Ellar, D. J. 1988. Toxin-membrane interactions of *Bacillus thuringiensis* delta-endotoxin. *Biochem. Soc. Trans.* 16:39–40
27. Ebersold, H. R., Lüthy, P., Muller, M. 1977. Changes in the fine structure of the gut epithelium of *Pieris brassicae* induced by the δ-endotoxin of *Bacillus thuringiensis*. *Bull. Soc. Entomol. Suisse* 50:269–76
28. Ellar, D. J. 1990. Pathogenicity determinants of entompathogenic bacteria. *Proc. 5th Int. Colloquim on Invertebrate Pathology and Microbial Control, Adelaide, Australia, 20 24 August, 1990*, pp. 298–302. Adelaide, Australia: Soc. Invertebrate Pathology
29. Endo, Y., Nishiitsutsuji-Uwo, J. 1980. Mode of action of *Bacillus thuringiensis* δ-endotoxin: histopathological changes in the silkworm midgut. *J. Invertebr. Pathol.* 36:90–103
30. English, L. H., Cantley, L. C. 1986. Delta endotoxin is a potent inhibitor of the (Na, K)-ATPase. *J. Biol. Chem.* 261:1170–73
31. English, L. H., Readdy, T. L. 1989. Delta endotoxin inhibits a phosphatase in midgut epithelial membranes of *Heliothis virescens*. *Insect Biochem.* 19:145–52
32. Ge, A. Z., Nedka, I., Shivarova, N. I., Dean, D. H. 1989. Location of the *Bombyx mori* specificity domain on a *Bacillus thuringiensis* δ-endotoxin protein. *Proc. Natl. Acad. Sci. USA* 86:4037–41
33. Gill, S. S., Hornung, J. M. 1987. Cytolytic activity of *Bacillus thuringiensis* proteins to insect and mammalian cell lines. *J. Invertebr. Pathol.* 50:16–25
34. Gill, S. S., Singh, G. J. P., Hornung, J. M. 1987. Cell membrane interaction of *Bacillus thuringiensis* subsp. *israelensis* cytolytic ...ns. *J. ... Invert ...* 1:300–8
35. Giordana, B., Sacchi, F. V., Parenti, P., Hanozet, G. M. 1989. Amino acid transport systems in intestinal brush-border membranes from lepidopteran larvae. *Am. J. Phys.* 257:R494–R500
36. Griego, V. M., Fancher, L. J., Spence, K. D. 1980. Scanning electron microscopy of the disruption of the tobacco hornworm, *Manduca sexta*, midgut by *Bacillus thuringiensis* endotoxin. *J. Invertebr. Pathol.* 35:186–89
37. Gupta, B. L., Dow, J. A. T., Hall, T. A., Harvey, W. R. 1985. Electron probe X-ray microanalysis of the effects of *Bacillus thuringiensis* var. *kurstaki* crystal protein insecticide on ions in an electrogenic K^+-transporting epithelium of the larval midgut in the lepidopteran *Manduca sexta*, in vitro. *J. Cell Sci.* 74:137–52
38. Haider, M. Z., Ellar, D. J. 1987. Analysis of the molecular basis of insecticidal specificity of *Bacillus thuringiensis* crystal δ-endotoxin. *Biochem. J.* 248:197–201
39. Haider, M. Z., Ellar, D. J. 1989. Functional mapping of an entomocidal δ-endotoxin. Single amino acid changes produced by site-directed mutagenesis influence toxicity and specificity of the protein. *J. Mol. Biol.* 208:183–94
40. Harvey, W. R., Crawford, D. N., Eisen, N. S., Fernandes, V. F., Spaeth, D. D., et al. 1987. The potassium impermeable apical membrane of insect epithelia: a target for the development of safe pesticides. *Mem. Inst. Oswaldo Cruz.* 82(Suppl. III):29–34
41. Heimpel, A. M., Angus, T. A. 1959. The site of action of crystalliferous bacteria in Lepidoptera larvae. *J. Insect Pathol.* 1:152–70
42. Hendrickx, K., De Loof, A., van Mellaert, H. 1989. Effects of *Bacillus thuringiensis* delta-endotoxin on the permeability of brush border membrane vesicles from tobacco hornworm (*Manduca sexta*) midgut. *Comp. Biochem. Physiol.* 95C:241–45
43. Hofmann, C., Lüthy, P., Hutter, R., Pliska, V. 1988. Binding of the delta endotoxin from *Bacillus thuringiensis* to brush-border membrane vesicles of the cabbage butterfly (*Pieris brassicae*). *Eur. J. Biochem.* 173:85–91
44. Hofmann, C., Vanderbruggen, H., Höfte, H., van Rie, J., Jansens, S., et al. 1988. Specificity of *Bacillus thuringiensis* δ-endotoxins is correlated with the presence of high-affinity binding

sites in the brush border membrane of target insect midguts. *Proc. Natl. Acad. Sci. USA* 85:7844–48

44a. Hodgman, T. C., Ellar, D. J. 1990. Models for the structure and function of the *Bacillus thuringiensis* δ-endotoxins determined by computational analysis. *DNA Seq.—J. DNA Seq. Mapp.* 1:97–106

45. Höfte, H., Greve, H., Seurinck, J., Jansens, S., Mahillon, J., et al. 1986. Structural and functional analysis of a cloned delta endotoxin of *Bacillus thuringiensis* Berliner 1715. *Eur. J. Biochem.* 161:273–80

46. Höfte, H., Soetaert, P., Jansens, S., Peferoen, M. 1990. Nucleotide sequence and deduced amino acid sequence of a new Lepidoptera-specific crystal protein gene from *Bacillus thuringiensis*. *Nucleic Acids. Res.* 18: 5545

47. Höfte, H., van Rie, J., Jansens, S., van Houtven, A., Vanderbruggen, H., et al. 1988. Monoclonal antibody analysis and insecticidal spectrum of three types of Lepidoptera-specific crystal proteins of *Bacillus thuringiensis*. *Appl. Environ. Microbiol.* 54:2010–17

48. Höfte, H., Whiteley, H. R. 1989. Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiol. Rev.* 53:242–55

49. Honée, G., Van der Salm, T., Visser, B. 1988. Nucelotide sequence of crystal protein gene isolated from *Bacillus thuringiensis* subspecies *entomocidus* 60.5 coding for a toxin highly active against *Spodoptera* species. *Nucleic Acids Res.* 16:6240

50. Ibarra, J. E., Federici, B. A. 1986. Isolation of a relatively non-toxic 65-kilodalton protein inclusion from parasporal body of *Bacillus thuringiensis* subsp. *israelensis*. *J. Bacteriol.* 165: 527–33

51. Johnson, D. E., Brookhart, G. L., Kramer, K. J., Barnett, B. D., McGaughey, W. H. 1990. Resistance to *Bacillus thuringiensis* by the Indian meal moth *Plodia interpunctella*: comparision of midgut proteinases from susceptible and resistant larvae. *J. Invertebr. Pathol.* 55:235–44

52. Kamb, A., Iverson, L. E., Tanouye, M. A. 1987. Molecular characterization of Shaker, a Drosophila gene that encodes a potassium channel. *Cell* 50:404–13

53. Knowles, B. H., Blatt, M. R., Tester, M., Horsnell, J. M., Carroll, J., et al. 1989. A cytolytic δ-entotoxin from *Bacillus thuringiensis* var. *israelensis* forms cation-selective channels in planar lipid bilayers. *FEBS Lett.* 244:259–62

54. Knowles, B. H., Ellar, D. J. 1986. Characterization and partial purification of a plasma membrane receptor for *Bacillus thuringiensis* var. *kurstaki* lepidopteran-specific δ-endotoxin. *J. Cell Sci.* 83:89–101

55. Knowles, B. H., Ellar, D. J. 1987. Colloid-osmotic lysis is a general feature of the mechanism of action of *Bacillus thuringiensis* δ-endotoxins with different insect specificity. *Biochim. Biophys. Acta* 924:509–18

56. Knowles, B. H., Farndale, R. W. 1988. Activation of insect cell adenylate cyclase by *Bacillus thuringiensis* δ-endotoxins and melittin. *Biochem. J.* 253: 235–41

57. Knowles, B. H., Francis, P. H., Ellar, D. J. 1986. Structurally related *Bacillus thuringiensis* δ-endotoxins display major differences in insecticidal activity in vivo and in vitro. *J. Cell Sci.* 84:221–36

58. Lane, N. J., Harrison, J. B., Lee, W. M. 1989. Changes in microvilli and Golgi-associated membranes of lepidopteran cells induced by an insecticidally active bacterial δ-endotoxin. *J. Cell Sci.* 93: 337–47

59. Lüthy, P., Jaquet, F., Hofmann, C., Huber-Lakac, M., Wolfersberger, M. G. 1986. Pathogenic actions of *Bacillus thuringiensis* toxin. *Zentralbl. Bakteriol. Microbiol. Hyg. Abt. I* 15:161–66 (Suppl.)

60. Maddrell, S. H. P., Lane, N. J., Harrison, J. B., Overton, J. A., Moreton, R. B. 1988. The initial stages in the action of an insecticidal δ-endotoxin of *Bacillus thuringiensis* var. *israelensis* on the epithelial cells of the Malpighian tubules of the insect, *Rhodnius prolixus*. *J. Cell Sci.* 90:131–44

61. Maddrell, S. H. P., Overton, J. A., Ellar, D. J., Knowles, B. H. 1989. Action of activated 27,000 *M*, toxin from *Bacillus thuringiensis* var. *israelensis* on Malpighian tubules of the insect *Rhodnius prolixus*. *J. Cell Sci.* 94:601–8

62. McGaughey, W. H. 1985. Insect resistance to the biological insecticide *Bacillus thuringiensis*. *Science* 229:193–95

63. McGaughey, W. H., Beeman, R. W. 1988. Resistance to *Bacillus thuringiensis* in colonies of Indianmeal moth and almond moth (Lepidoptera: Pyralidae). *J. Econ. Entomol.* 81:28–33

64. McGaughey, W. H., Johnson, D. E. 1987. Toxicity of different serotypes and toxins of *Bacillus thuringiensis* to resistant and susceptible Indianmeal moth (Lepidoptera: Pyralidae). *J. Econ. Entomol.* 80:1122–26

65. McPherson, A., Jurnak, F., Singh, G. J. P., Gill, S. S. 1987. Preliminary X-ray diffraction analysis of crystals of *Bacillus thuringiensis* toxin, a cell membrane disrupting protein. *J. Mol. Biol.* 195: 755–57

66. McPherson, S. A., Perlak, F. J., Fuchs, R. L., Marrone, P. G., Lavrik, P. B., et al. 1988. Characterization of the coleopteran-specific protein gene of *Bacillus thuringiensis* var. *tenebrionis*. *Biol Technology* 6:61–65

67. Muthukumar, G., Nickerson, K. W. 1987. The glycoprotein toxin of *Bacillus thuringiensis* subsp. *israelensis* indicates a lectinlike receptor in the larval mosquito gut. *Appl. Environ. Microbiol.* 53:2650–55

68. Parker, M. W., Pattus, F., Tucker, A. D., Tsernoglou, D. 1989. Structure of the membrane-pore-forming fragment of colicin A. *Nature* 337:93–96

69. Pao-intara, M., Angsuthanasombat, C., Panyim, S. 1988. The mosquito larvicidal activity of 130-kDa delta-endotoxin of *Bacillus thuringiensis* var. *israelensis* resides in the 72 kDa amino-terminal fragment. *Biochem. Biophys. Res. Commun.* 153:249–300

70. Percy, J., Fast, P. G. 1983. *Bacillus thuringiensis* crystal toxin: ultrastructural studies of its effect on silkworm midgut cells. *J. Invertebr. Pathol.* 41:86–98

71. Pfannenstiel, M. A., Muthukumar, G., Couche, G. A., Nickerson, K. W. 1987. Amino sugars in the glycoprotein toxin from *Bacillus thuringiensis* subsp. *israelensis*. *J. Bacteriol.* 169:796–801

72. Ryerse, J. S., Beck, J. R. Jr., Lavrik, P. B. 1990. Light microscope immunolocation of *Bacillus thuringiensis kurstaki* δ-endotoxin in the midgut and Malpighian tubules of the tobacco budworm, *Heliothis virescens*. *J. Invertebr. Pathol.* 56:86–90

73. Sacchi, V. F., Parenti, P., Giordana, B., Hanozet, G. M., Lüthy, P., et al. 1986. *Bacillus thuringiensis* toxin inhibits K^+-gradient-dependent amino acid transport across the brush border membrane of *Pieris brassicae* midgut cells. *FEBS Lett.* 204:213–18

74. Schnepf, H. E., Tomczak, K., Ortega, J. P., Whiteley, H. R. 1990. Specificity-determining regions of a lepidopteran-specific insecticidal protein produced by *Bacillus thuringiensis*. *J. Biol. Chem.* 265:20923–30

75. Schnepf, H. E., Wong, H. C., Whiteley, H. R. 1985. The amino acid sequence of a crystal protein from *Bacillus thuringiensis* deduced from the DNA base sequence. *J. Biol. Chem.* 260: 6264–72

76. Sekar, V., Thompson, D. V., Maroney, M. J., Bookland, R. G., Adang, M. J. 1987. Molecular cloning and characterization of the insecticidal crystal protein gene of *Bacillus thuringiensis* var. *tenebrionis*. *Proc. Natl. Acad. Sci. USA* 84:7036–40

77. Sick, A., Gaertner, F., Wong, A. 1990. Nucleotide sequence of a coleopteran-active toxin gene from a new isolate of *Bacillus thuringeinsis* subsp. *tolworthi*. *Nucleic Acids Res.* 18:1305

78. Singh, G. J. P., Schwest, L. P. Jr., Gill, S. S. 1986. The toxin action of *Bacillus thuringiensis* var. *israelensis* in *Aedes aegypti* in vivo. I. Relevance of midgut lesions to its poisoning syndrome. *Pestic. Biochem. Phys.* 26:36–46

79. Slatin, S. L., Abrams, C. K., English, L. 1990. Delta-entotoxins form cation-selective channels in planar lipid bilayers. *Biochem. Biophys. Res. Comm.* 169:765–72

80. Stevens, C. F. 1991. Making a submicroscopic hole in one. *Nature* 349: 657–58

81. Stone, T. B., Sims, S. R., Marrone, P. G. 1989. Selection of tobacco budworm for resistance to a genetically engineered *Pseudomonas fluorescens* containing the δ-endotoxin of *Bacillus thuringiensis* subsp. *kurstaki*. *J. Invertebr. Pathol.* 53:228–34

82. Tabashnik, B. E., Cushing, N. L., Finson, A., Johnson, M. W. 1990. Field development of resistance to *Bacillus thuringiensis* in diamondback moth (Lepidoptera: Plutellidae). *J. Econ. Entomol.* 85:1671–76

83. Thelestam, M., Blomqvist, L. 1988. Staphlococcal alpha-toxin: recent advances. *Toxicon* 26:55–65

84. Thomas, W. E., Ellar, D. J. 1983. Mechanism of action of *Bacillus thuringiensis* var. *israelensis* insecticidal δ-endotoxin. *FEBS Lett.* 154:362–68

85. Thorne, L., Garduno, F., Thompson, T., Decker, D., Zounes, M., et al. 1986. Structural similarity between the Lepidoptera- and Diptera-specific insecticidal endotoxin genes of *Bacillus thuringiensis* subsp. *kurstaki* and *israelensis*. *J. Bacteriol.* 166:801–11

86. van Rie, J., Jansens, S., Höfte, H., Degheele, D., van Mellaert, H. 1989. Specificity of *Bacillus thuringiensis* δ-

278–282 Nucleic Acids Research, 1997, Vol. 25, No. 1

O-GLYCBASE version 2.0: a revised database of O-glycosylated proteins

Jan E. Hansen*, Ole Lund, Kristoffer Rapacki and Søren Brunak

Center for Biological Sequence Analysis, The Technical University of Denmark, Building 206, DK-2800 Lyngby, Denmark

Received October 4, 1996; Accepted October 8, 1996

ABSTRACT

O-GLYCBASE is an updated database of information on glycoproteins and their O-linked glycosylation sites. Entries are compiled and revised from the literature, and from the SWISS-PROT database. Entries include information about species, sequence, glycosylation sites and glycan type. O-GLYCBASE is now fully cross-referenced to the SWISS-PROT, PIR, PROSITE, PDB, EMBL, HSSP, LISTA and MIM databases. Compared with version 1.0 the number of entries have increased by 34%. Revision of the O-glycan assignment was performed on 20% of the entries. Sequence logos displaying the acceptor specificity patterns for the GalNAc, mannose and GlcNAc transferases are shown. The O-GLYCBASE database is available through WWW or by anonymous FTP.

INTRODUCTION

Although it is not yet fully acknowledged, it is likely that most secreted proteins are glycosylated. One type is O-glycosylation, which is a post-translational event, where a carbohydrate is covalently linked to the hydroxyl group of serine or threonine. O-linked glycosylation serves a variety of functions in biology: They serve as ligands for selectins thereby mediating cell adhesion and leukocyte migration through endothelia. O-linked glycosylations confers resistance to proteolysis of stem regions of membrane proteins (1,2), and are involved in recognition phenomena (3) as sperm–egg binding (4). Abberant short chained O-linked glycosylations, found in a number of cancers, have been been used as tumor markers and may also be involved in the metastatic spread of tumor cells (5). As a glycoprotein can change glycosylation pattern if expressed in a non-native cell system and that this may introduce changes in function, expression rates and oligomerization, glycosylation has increasing implications for the biotechnological industry (6). It is thus of basic biological as well as of functional interest to know the sites of O-glycosylation in a given glycoprotein. As experimental methods to determine O-glycosylation are cumbersome as they require solid phase Edman degradation, tools for predicting the sites of O-glycosylation from primary sequence are needed (7–9). A compilation of known and verified O-glycosylation sites facilitates such research.

NEW FEATURES OF VERSION 2.0

Several cases of unprecedented O-linked glycosylations of prokaryotic glycoproteins have been included. One example is Pilin from *Neisseria gonorrhoeae*, where the site and sugar type of the glycoprotein was resolved by X-ray analysis (10). Another example of prokaryotic glycosylation is the glycoprotein MPT-32 from *Mycobacterium tuberculosis*, where the site of glycosylation recently has been resolved (11). Secondly, O-glycosylation sites from insects as the antibacterial glyco-peptide Lebocin from the silkworm *B.mori* (12) have been included. In order to facilitate easy access to other relevant information about the glycoprotein in question, entries have been extensively cross-referenced, not only to the sequence databases as SWISS-PROT, PIR and EMBL, but also to structural, homology and pattern specific databases such as PDB, PROSITE, HSSP, LISTA and MIM.

O-GLYCOSYLATION TYPES AND SEQUENCE MOTIFS

As only a minor fraction (~5–10%) of serine and threonine residues are modified by glycosylation, sequence context or conformational rules which determine whether a given serine or threonine is O-glycosylated, must exist. Such an acceptor rule has been deduced for N-linked glycosylation of asparagine as the consensus tripeptide Asn-X-Ser/Thr, where X can be any residue except proline (13). Motifs with Asn-X-Cys have also been found. No such clear consensus motif exists for the most abundant mucin-type of O-glycosylation (7,14–16) and neither for the mannose type of glycosylation. Mucin type O-glycosylation is initiated by transfer of *N*-acetylgalactosamine to the hydroxy amino acids serine and threonine by a family of differentially expressed UDP-GalNAc polypeptide *N*-acetylgalactosaminyl transferases (GalNAc transferases) (17–19) in the *cis*-Golgi compartment (20). Elongation of the O-linked oligosaccharide chain proceed in the subsequent Golgi compartments in a stepwise process specific for the expressing cell type.

Heterogeneity in glycosylation is profound, since at some sites only a fraction of the glycoprotein molecules are glycosylated even if expressed in the same cell type. Secondly the carbohydrate composition in the proceeding glycan tree is also found to vary at a given site. Whether this heterogeneity is truly stochastic or relates to the amino acid sequence or structure at the given site remains to be clarified.

* To whom correspondence should be addressed. Tel: +45 4525 2485; Fax: +45 4593 4808; Email: janhan@cbs.dtu.dk

Table 1. Sugars linked to serine and threonine residues

Protein example	Organism	Linked sugar	Reference
Mucins and others	mammals, fish and birds	GalNAc	37,20
Proteoglycans	mammals	Xyl	38,39
Nuclear and cytoplasmic proteins	mammals	GlcNAc	23
Factor IX, VII and protein Z	mammals	Glc	40
Factor IX and XII and other EGF proteins	mammals	Fuc	41,25,42,43
Glycoamylase	fungi	Man	44
Lectins	plants	Gal	45,46

At least six other sugar types (Table 1) can be linked to serine or threonine by different glycosyltransferases. Different peptide acceptor motifs have been proposed for these transferases (21–26). It is therefore important to discriminate between the type of O-linked oligosaccharide before deducing the sequence acceptor motifs for the transferase in question (see Fig. 1). However, for the common mucin type O-glycosylation, the proposed motifs (21,22) account only for a fraction of the known mucin type O-glycosylation sites (7). For the xylosyltransferase has deviations from the proposed motif Ser-Gly-X-Gly, where X is any amino acid (26), also been found (27). The changes from the proposed consensus acceptor motifs of O-glycosyl transferases motivates a compilation of all known O-linked glycosylation sites and a reinterpretation of the sequence acceptor patterns. Solving the problem of acceptor specificity of O-glycan transferases may have to take into account that O-glycosylation is a post-translational process proceeding after N-glycosylation and folding (20). Thus the accessibility of the transferase is restricted to exposed surface residues of the glycoprotein. O-glycosylation of a given site may therefore be dependent of the overall structure of the protein (28). Sequence acceptor motifs which are O-glycosylated at the surface of glycoproteins, may therefore not be O-glycosylated if buried in the hydrophobic core of the protein. Predicting O-glycosylated sites with confidence therefore implicitly includes prediction of the surface exposure of a given site, which is much more complex than a simple consensus motif, as it necessarily includes long range interactions within the peptide chain.

DATA SOURCES

Proteins with a carbohydrate assigned either to a serine or a threonine residue were extracted from the 1996 version of SWISS-PROT (29) (release 33.00) as well as directly from recently published reports. All entries were checked by consulting the original references and a revision was performed if necessary. One example is glycophorin A, which recently has been revised to contain 16 O-glycosylation sites (22). Revision of O-glycan assignments (either number or position) was performed in 20% of the entries. In general the information of the primary databases was reliable, but care should be taken before concluding that a given serine or threonine is not glycosylated if not mentioned in the feature table. All erroneous assignments in the primary database were communicated to the database administrator.

As discussed earlier by Gooley and colleagues (16) there exist differences between glycosylation sites deduced from *in vitro* glycosylated peptides from a particular glycoprotein and glycosylation sites deduced from Edman degradation of the same *in vivo* glycosylated glycoprotein. This fact has now been confirmed in a very elegant way by Tabak and colleagues (30) using a chimeric reporter protein containing the acceptor site expressed in COS7 cells. However, we have included both *in vitro* and *in vivo* data in the database. When differences appear regarding glycosylation sites as a cause of the experimental technique, this will be noted in the database in the comment field.

DESCRIPTION

Version 2.0 (25/9-1996) of O-GLYCBASE contains 127 glycoprotein entries, 48 057 residues and 726 experimentally determined O-glycosylation sites. The entries consist of 11 fields composed by a field name followed by a colon (:) delimiter and the field values separated by commas. Each entry starts with an identifier and entry date. Empty fields are marked by a hyphen (–). An example of an entry is shown in Figure 2.

FIELDS

IDENTIFIER

The unique identifier is a three digit number preceded by a (>) to mark the beginning of the entry. The entry date is also given here.

GLYCPROT

The name of the glycoprotein, all alternative names and widely used abbreviations are given in this field.

SPECIES

The species is given in this field. This is relevant in order to relate subtle differences in the glycosylation pattern between related species.

DB_REF

The assigned glycoprotein sequence is fully cross-referenced with the major databases: SWISS-PROT, PIR, PROSITE, PDB, EMBL, HSSP, LISTA and MIM. Both accession numbers and entry names are given in this field.

OGLYCAN

The reducing terminal carbohydrate linked to the serine and threonine is mentioned (if known). Carbohydrate structures of the elongated chain are mentioned when known.

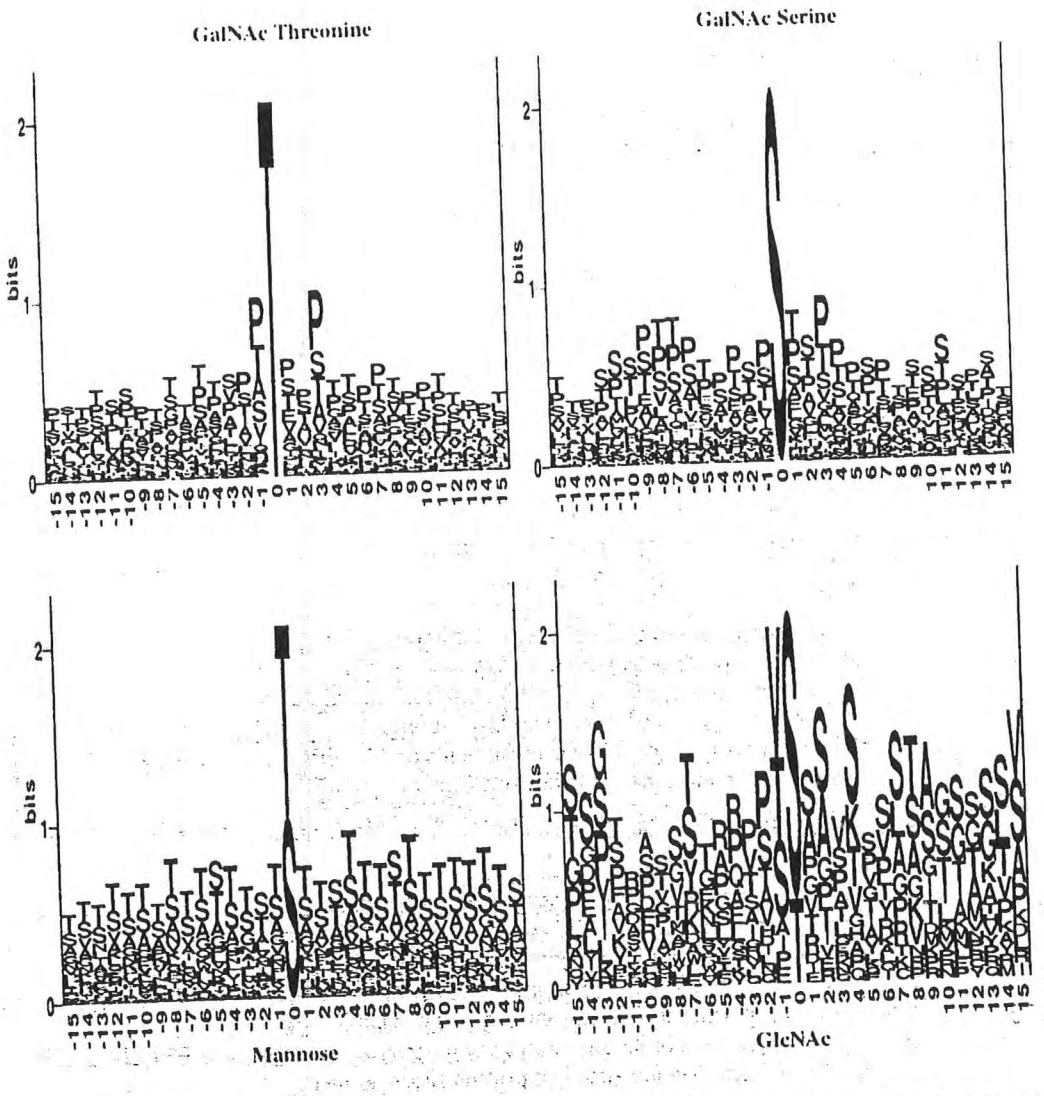


Figure 1. Shannon information content (7,34) shown in logo form (35) for mucin type O-glycosylation of threonine (upper left, 224 sites), serine (upper right, 110 sites), Mannose (lower left, 117 sites) and GlcNAc (lower right, 22 sites). All sites were aligned with the glycosylated residue at position 0. The height of each column reflects the bias from random of particular residues surrounding the glycosylation site. The size of each residue letter reflects the frequency of the residue on that position. Thereby, the logo displays the specificity patterns of GalNAc, Mannose and GlcNAc transferases. Note the high frequency of proline at position –1 and +3 as found Wilson *et al.* (14) for GalNAc glycosylated threonines. The high frequency of serine and threonine in the vicinity of the glycosylated position 0 reflects th t O-glycosylation sites in general are clustered, or found within domains of clustered serines and threonines. Note that the acceptor sequence patterns for GalNAc (upper panel), Mannose and GlcNAc transferases (lower panel) have some features in common: (i) they all have serine and threonine as abundant amino acids in the context of the glycosylatin site (position 0), indicating that clustering is a common feature. (ii) They all display a low frequency of charged amino acids at position –1 as originally found for the GalNAc transferase only by O'Connell *et al.* (36) and recently confirmed by Nehrke *et al.* (30). These common features suggests that the transferases may have a common ancestor. Sequence similarity between human GalNAc transferases and a counterpart from *C.elegans* have recently been reported (17). The differences between the transferase acceptor patterns include: proline in the context is only evident for GalNAc and GlcNAc. For mannose the most frequent hydrophobic residue is alanine on all positions. For GlcNAc a high frequency of valine is found at position –1, although the statistics only include 22 sites. The height of the central serine/threonine has been scaled to magnify the context and is thus non-informative. The neutral and polar amino acids are shown in green, the basic in blue, the acidic in red and the neutral and hydrophobic in black.

SER, THR and ASN

These following fields contain the assignment of the O-linked and N-linked glycosylation sites. If the glycosylation site is verified experimentally by Edman degradation or other experimental techniques it is marked as experimental. Sites marked as potential in the databases are marked as predicted. Glycoproteins containing predicted sites only are not included. The N-linked glycosylation is also marked in order to see the relation between N- and O-linked glycosylation.

REFERENCES

References to the published source of the O-glycan assignment. Additional references to the sequence information are given in the primary databases.

SEQ

Contains first the number of residues in the sequence and then the full sequence including signal sequence in a format of 80 characters per line. The assignment of N-linked and O-linked

```
>105          7 Aug 1996
GLYCPROT:     Myc prot-oncogene protein, c-MYC.
SPECIES:      Human
DB_REF:       SWISS,          P01106, P01107, MYC_HUMAN.
              PIR,            A01349, TVHUM.
              PIR,            A01350, TVHUT.
              PROSITE,        PS00038
OGLYCAN:      GlcNAc
SER:          Experimental:   -
              Predicted:      -
THR:          Experimental:   58
              Predicted:      -
ASN:          Experimental:   -
              Predicted:      -
REFERENCES:   Chou TY., Hart GW., Dang CV. (1995)
              J. Biol. Chem., 270:18961-18965
SEQ:          439
MPLNVSFTNRNYDLDYDSVQPYFYCDEEENFYQQQQQSELQPPAPSEDIWKKFELLPTPPLSPSRRSGLCSPSYVAVTPF
SLRGDNDGGGGSFSTADQLEMVTELLGGDMVNQSFICDPDDETFIKNIIIQDCMWSGFSAAAKLVSEKLASYQAARKDSG
SPNPARGHSVCSTSSLYLQDLSAAASECIDPSVVFPYPLNDSSSPKSCASQDSSAFSPSSDSLLSSTESSPQGSPEPLVL
HEETPPTTSSDSEEEQEDEEEIDVVSVEKRQAPGKRSESGSPSAGGHSKPPHSPLVLKRCHVSTHQHNYAAPPSTRKDYP
AAKRVKLDSVRVLRQISNNRKCTSPRSSDTEENVKRRTHNVLERQRRNELKRSFFALRDQIPELENNEKAPKVVILKKAT
AYILSVQAEEQKLISEEDLLRKRREQLKHKLEQLRNSCA
.....T.....
.....
.....
.....
.....
.....
COMMENT:      Revised according to reference.
```

Figure 2. Example of an entry in O-GLYCBASE.

glycosylation sites is given below, where uppercase (T, S, N) denote experimentally verified glycosylation sites of threonine, serine and asparagine, respectively and lower case (t, s, n) denote predicted sites. Dots (.) denote no glycosylation.

COMMENTS

Contains any relevant comment or observation. Information about expression system and whether the site information is obtained from an *in vivo* glycosylated glycoprotein or from *in vitro* glycosylated peptides is included here.

PREDICTION OF MUCIN TYPE O-GLYCOSYLATION

To predict whether recently sequenced proteins contain putative mucin type O-glycosylation sites we have developed an algorithm to predict mucin type O-glycosylation sites from the primary structure (7). The method is based on artificial neural networks, renowned for their ability to recognize even highly complex and nonlinear sequence patterns (31–33). The networks were trained on a non-redundant mammalian extract of O-GLYCBASE and the performance of the predictions were tested on two independent tests. This was done in order to estimate the range of performance on a given new sequence. As the sequence context around O-glycosylated serines differed from threonine (Fig. 2) these were treated separately. The networks recognized 60–95% of the O-glycosylation sites in the two independent test sets. Details of the network architectures used, training procedures, test set selection and validation are given in Hansen *et al.* (7). The accuracy of the method is mainly dependent on size of the dataset used during training. As new mucin-type glycosylation sites are deduced every year an updated non-redundant dataset can be constructed. We are therefore currently updating the NetOglyc algorithm to version 2.0 using an enlarged dataset (will be published elsewhere). For prediction benchmark studies the non-redundant version of O-GLYCBASE is made available. It contains 53 mammalian glycoproteins with 264 O-glycosylation sites. It can be accessed through anonymous FTP to: ftp.cbs.dtu.dk, file /pub/Oglyc/O-Unique.seq. Prediction of mucin-type O-glycosylation is available from URL: http://www.cbs.dtu.dk/netOglyc/cbsnetOglyc.html ; or by Email to: netOglyc@cbs.dtu.dk. Forward a mail including the word 'help' for obtaining information about sequence formats and other details.

QUALITY AND COMPLETENESS OF DATA AND FUTURE DIRECTIONS

The redundancy has been minimized by removing all identical sequences. For some sequences conflicting data regarding glycosylation assignments appear in the literature for the same glycoprotein. This information is included with notification of the conflict in the comment field. The aim has been to include all O-glycosylated proteins where the exact position of O-glycosylation is known. Comparison with the original papers and comparison with different databases has been used to minimize errors. However, there may still be sites in the database which should be assigned as glycosylated as the recent revision of glycophorin A shows (22). New entries will be included as they are published or submitted directly to the database.

DATABASE ACCESS

Version 2.0 of O-GLYCBASE is accessible via the internet.
URL: http://www.cbs.dtu.dk/databases/OGLYCBASE/ or by anonymous FTP to: ftp.cbs.dtu.dk, file pub/Oglyc/Oglyc.base

Users of the database are encouraged to provide corrections, comments, or new material for the database; unpublished material will be held confidential at the request of the authors. The administrator of the database can be contacted by Email: janhan@cbs.dtu.dk

We encourage users of the database to quote this paper.

ACKNOWLEDGEMENTS

This work was supported by the Danish National Research Foundation and the Danish Medical Research Foundation. Andrew Gooley and Henrik Clausen is thanked for stimulating discussions regarding specificities of various oligosaccharide transferases.

REFERENCES

1 Jentoft, N. (1990) *Trends Biochem. Sci.* **15**, 291–294.
2 Hart, G. (1992) *Curr. Opin. Cell Biol.* **4**, 1017–1023.
3 Fukuda, M. (1991) *Glycobiology* **1**, 337–356.
4 Kinloch, R., Sakai, Y., and Wassarman, P. (1995) *Proc. Natl Acad. Sci. USA* **92**, 263–267.
5 Muramatsu, T. (1993) *Glycobiology* **3**, 294–296.
6 Jenkins, N., Parekh, R., and James, D. (1996) *Nature Biotechnol.* **14**, 975–981.
7 Hansen, J.E., Lund, O., Engelbrecht, J., Bohr, H., Nielsen, J.O., Hansen, J-E.S., and Brunak, S. (1995) *Biochem. J.* **308**, 801–813.
8 Elhammer, A. P., Poorman, R., Brown, E., Maggiora, L., Hoogerheide, J. G., and Kzdy, F. D. (1993) *J. Biol. Chem.* **268**, 10029–10038.
9 Chou, K. C., Zhang, C. T., Kezdy, F. J., and Poorman, R. A. (1995) *Proteins* **21**, 118–126.
0 Parge, H. E., Forest, K. T., Hickey, M. J., Christensen, D. A., Getzoff, E. D., and Tainer, J. A. (1995) *Nature* **378**, 32–38.
11 Dobos, K. M., Swiderek, K., Khoo, K. H., Brennan, P. J., and Belisle, J. T. (1995) *Infect. Immun.* **63**, 2846–2853.
2 Hara, S. and Yamakawa, M. (1995) *Biochem. J.* **310**, 651–656.
3 Bause, A. (1983) *Biochem. J.* **209**, 331–336.
14 Wilson, I. B. H., Gavel, Y., and Heijne, G. (1991) *Biochem. J.* **275**, 529–534.
5 Wang, Y., Abernethy, J., Eckhardt, A., and Hill, R. (1992) *J. Biol. Chem.* **267**, 12709–12716.
16 Gooley, A. and Williams, K. (1994) *Glycobiology* **4**, 413–417.
17 Clausen, H. and Bennett, E. P. (1996) *Glycobiology* (review in press).
18 Sorensen, T., White, T., Wandall, H. H., Kristensen, A. K., Roepstorff, P., and Clausen, H. (1995) *J. Biol. Chem.* **270**, 24166–24173.
9 White, T., Bennett, E. P., Takio, K., Sorensen, T., Bonding, N., and Clausen, H. (1995) *J. Biol. Chem.* **270**, 24156–24165.
20 Carraway, K. and Hull, S. (1991) *Glycobiology* **1**, 131–138.
1 Gooley, A., Classon, B., Marshcalek, R., and Williams, K. (1991) *Biochem. Biophys. Res. Commun.* **178**, 1194–1200.
22 Pisano, A., Redmond, J., Williams, K., and Gooley, A. (1993) *Glycobiology* **3**, 429–435.
23 Haltiwanger, R., Kelly, W., Roquemore, E., Blomberg, M., Dong, L., Kreppel, L., Chou, T., and Hart, G. (1992) *Biochem. Soc. Trans.* **20**, 264–269.
24 Haynes, P. A., Gooley, A. A., Ferguson, M. A., Redmond, J. W., and Williams, K. L. (1993) *Eur. J. Biochem.* **216**, 729–737.
25 Harris, R. J. and Spellman, M. W. (1993) *Glycobiology* **3**, 219–224.
26 Bourdon, M. A., Krusius, T., Campbell, S., Schwartz, N. B., and Ruoslahti, E. (1987) *Proc. Natl Acad. Sci. USA* **84**, 3194–3198.
27 Huber, S., Winterhalter, K. H., and Vaughan, L. (1988) *J. Biol. Chem.* **263**, 752–756.
28 Dahms, N. and Hart, G. (1986) *J. Biol. Chem.* **261**, 13186–13196.
29 Bairoch, A. and Boeckmann, B. (1994) *Nucleic Acids Res.* **22**, 3578–35 0.
30 Nehrke, K., Hagen, F., and Tabak, L. (1996) *J. Biol. Chem.* **271**, 7061–7065.
31 Presnell, S. and Cohen, F. (1993) *Annu. Rev. Biophys. Biomol. Struct.* **22**, 283–298.
32 Brunak, S., Engelbrecht, J., and Knudsen, S. (1991) *J. Mol. Biol.* **220**, 49–65.
33 Brunak, S., Engelbrecht, J., and Knudsen, S. (1990) *Nucleic Acids Res.* **18**, 4797–4801.
34 Shannon, C. E. (1948) *Bell System Tech. J.* **27**, 379–423, 623–656.
35 Schneider, T. D. and Stephens, R. M. (1990) *Nucleic Acids Res.* **18**, 6097–6100.
36 O'Connell, B. C., Tabak, L. A., and Ramasubbu, N. (1991) *Biochem. Biophys. Res. Commun.* **180**, 1024–1030.
37 Strous, G. and Dekker, J. (1992) *Crit. Rev. Biochem. Mol. Biol.* **27**, 57–92.
38 Hardingham, T. and Fosang, A. (1992) *FASEB. J.* **6**, 861–70.
39 Yanagishita, M. and Hascall, V. (1992) *J. Biol. Chem.* **267**, 9451–9454.
40 Nishimura, H., Kawabata, S., Kisiel, W., Hase, S., Ikenaka, T., Takao, T., Shimonishi, Y., and Iwanaga, S. (1989) *J. Biol. Chem.* **264**, 20320–20325.
41 Nishimura, H., Takao, T., Hase, S., Shimonishi, Y., and Iwanaga, S. (1992) *J. Biol. Chem.* **267**, 17520–17525.
42 Harris, R. J., Leonard, C. K., Guzzetta, A. W., and Spellman, M. W. (1991) *Biochemistry* **30**, 2311–2314.
43 Harris, R. J., Ling, V. T., and Spellman, M. W. (1992) *J. Biol. Chem.* **26** 5102–5107.
44 Hausler, A., Ballou, L., Ballou, C., and Robbins, P. (1992) *Proc. Natl Acad. Sci. USA* **89**, 6846–6850.
45 Neustroev, K., Golubev, A., Firsov, L., Ibatullin, F., Protasevich, I., and Makarov, A. (1993) *FEBS Lett.* **316**, 157–160.
46 Allen, A., Desai, N., Neurberger, N., and Creeth, J. (1978) *Biochem. J.* **171**, 665–674.

: J. Biochem. *161*, 273–280 (1986)
FEBS 1986

tructural and functional analysis of a cloned delta endotoxin *Bacillus thuringiensis berliner* 1715

rman HÖFTE[1], Henri de GREVE[1], Jef SEURINCK[1], Stefan JANSENS[1], Jacques MAHILLON[1], Christophe AMPE[1], l VANDEKERCKHOVE[2], Hilde VANDERBRUGGEN[1], Marc van MONTAGU[2], Marc ZABEAU[1] and Mark VAECK[1]

'lant Genetic Systems N.V., Gent
.aboratorium voor Genetica Rijksuniversiteit, Gent

:ceived June 19/August 20, 1986) – EJB 86 0635

A plasmid-encoded crystal protein gene (*bt2*) has been cloned from *Bacillus thuringiensis berliner* 1715. In *Escherichia coli*, it directs the synthesis of the 130-kDa protein (Bt2) which is toxic to larvae of *Pieris brassicae* and *Manduca sexta*. Comparison of the deduced amino acid sequence of this Bt2 protein with the *B. thuringiensis kurstaki* HD1 Dipel, *B. thuringiensis kurstaki* HD73 and *B. thuringiensis sotto* crystal protein sequences suggests that homologous recombination between the different genes has occurred during evolution.

Treatment of the Bt2 protein with trypsin or chymotrypsin yields a 60-kDa protease-resistant and fully toxic polypeptide. The minimal portion of the Bt2 protein required for toxicity has been determined by analysing the polypeptides produced by deletion derivatives of the *bt2* gene. It coincides with the 60-kDa protease-resistant Bt2 fragment and it starts between amino acids 29 and 35 at the N-terminus and terminates between positions 599 and 607 at the C-terminus.

Bacillus thuringiensis is a gram-positive bacterium which oduces endogenous crystals upon sporulation. The crystals composed of proteins which are specifically toxic against tain insect larvae, mainly lepidopteran and dipteran species . Upon ingestion by larvae, the crystals dissolve in the aline conditions of the insect midgut and release proteins 130–160 kDa [2, 3]. Which are proteolytically processed midgut proteases to yield smaller toxic fragments [4].

Most crystal protein genes have been localised on large smids [5, 6]. Some genes have recently been cloned and pressed in *Escherichia coli* [7–10]. However these cloned ne products have not been subjected to a detailed functional aracterisation. On the other hand, all biochemical studies ve been performed on proteins derived from the original stals of *B. thuringiensis* [4, 11]. Generally such crystals are xtures of distinct polypeptides which may exhibit different ictional properties.

This paper describes the cloning of a crystal protein gene m *B. thuriengiensis* subspecies *berliner* 1715. The purified combinant polypeptide was shown to exhibit a toxic activity *Pieris brassicae* and *Manduca sexta* larvae comparable the activity of the original crystal protein. The minimum lypeptide fragment still exhibiting complete toxic activity s mapped within the crystal protein.

Correspondence to H. Höfte, N.V. Plant Genetic Systems, J. Ieaustraat 22, B-9000 Gent, Belgium

bbreviations. B.t., Bacillus thuringiensis; SDS-PAGE, sodium yl sulphate/polyacrylamide gel electrophoresis; kb, 10^3 bases; ISA, enzyme-linked immunosorbent assay; bp, base pairs; NaCl/ phosphate-buffered saline; HPLC, high-performance liquid chrotography; LD_{50}, dose at which 50% lethality is observed.

MATERIALS AND METHODS

Bacterial strains and plasmids

The *B. thuringiensis* strain *berliner* 1715 was kindly provided by Dr A. Klier (Inst. Pasteur, Paris) [8]. Growth and sporulation conditions were described in [12]. *E. coli* strains used were K514 [13] and K-12ΔH1Δtrp [14].

Plasmids used were pLK54, pLK85, pLK63 and pLK94 [15], pUC8 [16] pKM109/90 [17] and pEcoR251; the latter is a suicide vector expressing the *Eco*RI gene which can be inactivated by cloning DNA fragments in a unique *Bgl*II site, thus allowing a positive selection (Botterman et al., unpublished work).

Plasmids carrying a λP_L or λP_R promotor fragment were maintained in the cI-repressor-producing strain K514(λ). Temperature induction of the λP_L or P_R promoter was accomplished in strain K-12ΔH1Δtrp which carries a temperature-sensitive cI repressor, essentially as described by Zabeau and Stanley [14], except that induction was at 38°C.

DNA MANIPULATIONS

Recombinant DNA techniques were as described by Maniatis et al. [18] and DNA sequencing was performed according to Maxam and Gilbert [19].

3'-End deletions in bt2

Fig. 1 shows pLBKM25, the intermediate plasmid used to construct the 3'-end deletion derivatives. This plasmid is composed of the following elements (the numbers between

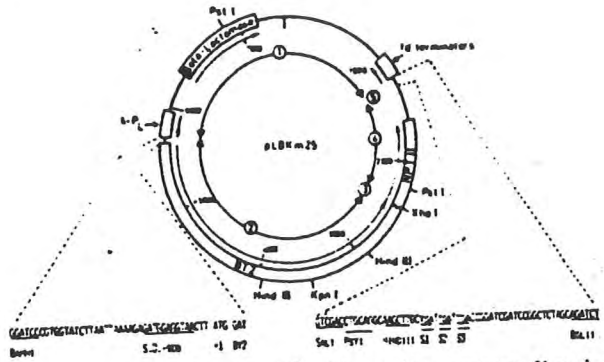


Fig. 1. *Intermediate plasmid pLBkm25, used to construct 3′ end deletions in* bt2. The sequences at the 5′ end of *bt2* and the region around the synthetic stop codons (S1-S2-S3) are shown. The numbers are explained in Materials and Methods

brackets refer to Fig. 1): a fragment of pLK54 carrying the β-lactamase gene, the origin of replication of pBR322, the lambda P_L promotor and two transcription terminators of phage fd in tandem (1); a fragment of pGI502 comprising the *bt2* gene from position −32 to 3240 in the DNA sequence (2); an *Xho*I linker (3); the *Hin*dIII-*Sal*I fragment of pKM109/90 containing the *neo* gene deleted at the 5′ end up to the 5th codon (4); a fragment of pLK94 carrying three stop codons in the three reading frames to prevent read-through translation from an open reading frame fused to the *Sal*I site (5). pLBKM25 was treated with *Kpn*I, *Sal*I, *Pol*Ik and ligase to generate pLB16; with *Hin*dIII and ligase to produce pLB12 and with *Kpn*I, *Bal*31 exonuclease, *Sal*I, the 'Klenow' fragment of *E. coli* DNA polymerase I, and ligase to produce the other 3′ deletion derivatives.

5′-End deletions in bt2

pRB10, derived from pGI502 and pLK63, contains the λP_R promotor, the upstream region and the ATG start codon of the *cro* gene followed by a *Bam*HI linker fused to the second codon of the full-length *bt2* gene. This gives rise to the sequence ATG GAT CCC GAT at the 5′ end of the gene with GAT being the 2nd codon of *bt2*. pRB210 is derived from pRB10. The *bt2* gene has been deleted up to the 37th codon using the *Cla*I site at position 106. The sequence at the 5′ end is: ATG GAT CGA TCC GAT where GAT is the 37th codon of *bt2*.

Fusions of bt2 *gene fragments to the 5′ end of* lacZ

pBZ12 is derived from pLB879 and pLK85, whereby the *lac*Z gene has been fused to the 3′ end of the deleted *bt2* gene, showing the sequence CCG GCA TCG ACC GAT CCC at the junction where T is position 1821 in *bt2* and C position 22 in *lac*Z. Similarly, pBZ13 is derived from pLB834 with the junction sequence being: TAT ATA GCA GAT CCG TCG ACC GAT CCC where G is position 1798 in *bt2* and C position 22 in *lac*Z.

BIOCHEMICAL METHODS

Preparation of crystal proteins

Crystals were isolated and purified from spore preparations as described by Mahillon and Delcour [12]. Crystal proteins were solubilized in 50 mM Na_2CO_3 pH 10, 5 mM dithiothreitol or in 50 mM $NaHCO_3$ pH 9.5, 200 mM thioglycolate. More than 90% of the crystal proteins were solubilized after a 2-h incubation at 37 °C.

Purification of Bt2 protein

Cells (5 g) from a saturated culture of K514 (pGI502) pellet were resuspended in 50 ml 50 mM Tris/HCl pH 7.9, 50 mM EDTA, 15% sucrose, treated with lysozyme (100 μg/ml) and sonicated on ice (10 min, 400 W). The cell debris were washed twice with 200 ml NaCl/P_i containing 2% Triton X-100, incubated for 30 min in the same buffer at 0 °C, centrifuged (10 min, 15000 × *g*) and washed twice with NaCl/P_i. The Bt2 protein, present in the pellet, was solubilized in 50 ml extraction buffer (0.1 M $NaHCO_3$, pH 9.5, 0.2 M thioglycolate) for 2 h at 37 °C.

For structural analysis, the Bt2 was further purified. The pellet was dissolved in buffer A (50 mM Tris/HCl, pH 8.0, 1 M urea, 0.2% 2-mercaptoethanol) and loaded on a column (10 × 2.5 cm) of DEAE-cellulose (Whatman DE 52) pre-equilibrated in the same buffer, washed twice with 50 ml buffer A, and eluted with 100 ml of a 0–0.3 M KCl linear gradient in buffer A. Fractions containing the Bt2 protein were pooled and precipitated at pH 4. The precipitate was collected and redissolved in 3 ml buffer A pH 10. Then the pH was re-adjusted to 8; the sample applied to a Sephacryl S-300 column (100 × 2.5 cm) and eluted with a flow rate of 40 ml/h. Pure Bt2 protein was dialysed exhaustively against 0.1 M NH_4HCO_3 containing 0.2% 2-mercaptoethanol.

Generation of the 60-kDa tryptic fragment of Bt2

Purified Bt2 protein (1 mg/ml) was digested with bovine trypsin or chymotrypsin (Sigma) 1:70 for 45 min at 37 °C.

Protein sequencing

Amino-terminal sequences were determined using a gas-phase sequenator (Applied Biosystems Inc. USA), operated essentially as described by Hewick et al. [20]. Stepwise liberated phenylthiohydantoin derivatives were identified by HPLC analysis as described by Hunkapiller and Hood [21]. The 60-kDa Bt2 fragment was separated on SDS-PAGE, electroblotted onto polybrene-coated glass-fibre sheets [22] and sequenced.

Immunodetection methods

Polyclonal antisera against crystal proteins and against Bt2 were raised in rabbits and mice. These antisera reacted strongly with Bt2 and with solubilized crystal proteins from *B.t. berliner* both in Ouchterlony assay and in ELISA.

Hybridomas producing monoclonal antibodies against *B.t.* crystal proteins will be described elsewhere (Vanderbruggen et al., unpublished results). Immunological screening of bacterial colonies [23] and immunoblotting on nitrocellulose [24] as performed using rabbit anti-(crystal serum) and a peroxidase-labelled goat anti-(rabbit Ig) (Sigma).

Quantification of soluble crystal proteins or derived polypeptides was carried out with the ELISA techniques [25]. Flexible polyvinyl microtiter plates, precoated with goat anti-(*B.t.* crystal protein) antibodies were incubated for 90 min at 4 °C with dilutions of the antigen samples. After washing, diluted anti-(*B.t.* crystal protein) antibody was added, usually a hybridoma culture supernatant (1 h at room temperature

ıle 1. *Toxicity experiments*
: LD_{50} values for *P. brassicae* (3rd larval instar) and for *M. sexta* larval instar) of different toxin preparations, were measured.
= not tested

	LD_{50} for	
	P. brassicae	*M. sexta*
	ng/larva	ng/cm²
crystal from *B.t. berliner*	15	n.t.
above, solubilized	0.6	7.5
protein	1.6	6
protein/trypsin	1.5	5

ıtes were washed and incubated with a dilution of alkaline-osphatase-labeled anti-(mouse IgG) antibodies (Sigma 5153). After washing, *p*-nitrophenyl phosphate (Sigma, 4-105) was added and the reaction monitored by measuring : *A* at 405 nm. The detection limit of the test for purified, ubilized crystal protein, was in the range 0.1–1 ng/ml. :eening of hybridoma supernatants was performed on crotiter plates which were coated directly by an overnight :ubation with the crystal protein solution at 4 °C (10 µg/).

:ect toxicity assays

Toxicity assays were performed on first instar larvae of : tobacco hornworm (*Manduca sexta*). 3 ml of artificial diet without formaldehyde was dispersed in a 4-cm² vial; a l sample was applied and four newly hatched larvae were ıced into each vial. 20 larvae were used per sample dilution. owth and mortality were followed over a 7-day period.

Toxicity assays on larvae of the large white cabbage butter- (*Pieris brassicae*) were done on discs (0.25 cm²), cut from sh cabbage leaves (*Brassica oleracea* var. *gemnifera* D.C.) which 5 µl of a sample dilution was applied. Third instar vae were obtained from a synchronised culture of *P. assicae*. When the first disc was consumed, a fresh disc thout sample was given. Viability of 50 larvae per sample ution were monitored every 24 h over a period of 5 days.

ESULTS

oning of a B. thuringiensis *gene encoding 130-kDa crystal protein exhibiting insecticidal activity*

The SDS/polyacrylamide gel electrophoresis patterns of *thuringiensis berliner* 1715 crystal preparations show two ajor protein bands of about 140 kDa and 130 kDa (Fig. 2A, ıe 10). The dissolved purified crystals from this strain are ghly toxic towards *P. brassicae* and *M. sexta* (larvae able 1).

A library was constructed by cloning size-fractionated 0–15 kb) *Sau*3A-digested plasmid DNA from *B. thurin-ensis berliner* 1715 into the suicide vector pEcoR251. Four lonies producing crystal proteins were identified among 50 clones using a colony immunoblot assay with rabbit anti- tal protein) serum. Restriction enzyme analysis revealed the four clones contained overlapping DNA fragments. om one of these clones, pGI612, a 7.5-kb *Bam*HI-*Pst*I frag-ent, comprising a region shared by the four plasmids, was bcloned into pUC8 to produce plasmid pGI502.

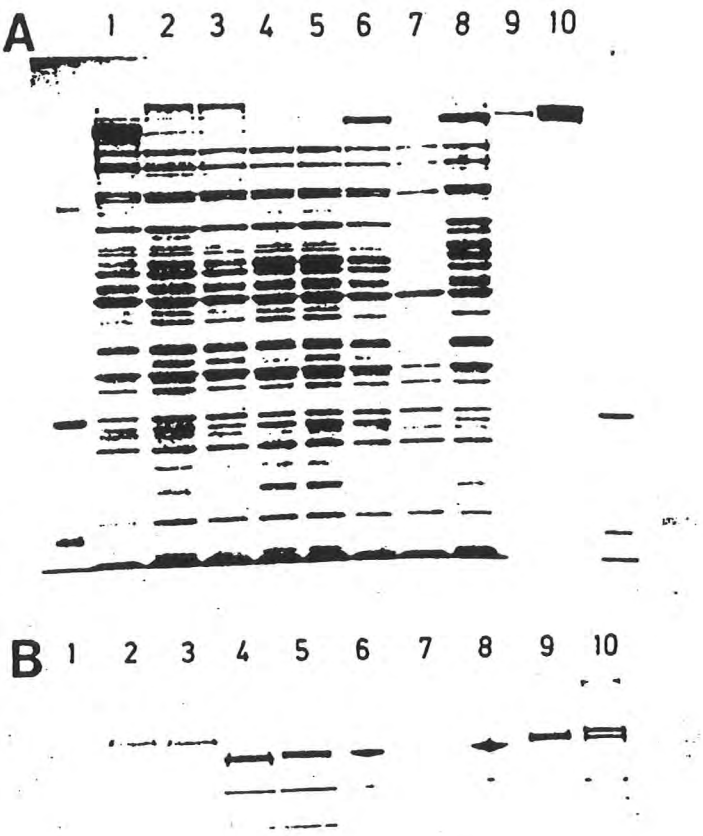


Fig. 2. *SDS-PAGE (A) and Western blot with rabbit antiserum against* B.t. berliner *crystal proteins (B) of clones containing the* bt2 *gene and its derivatives.* (A) A 10% polyacrylamide gel and Coomassie staining were used. Total cellular extracts of 2×10^8 cells were induced for 5 h. Lane 1, K-12ΔH1Δtrp (pHH10 control *lac*Z); lane 2, K-12ΔH1Δtrp (pBZ13); lane 3, K-12ΔH1Δtrp (pBZ12); lane 4, K-12ΔH1Δtrp (pRB210); lane 5, K-12ΔH1Δtrp (pRB10); lane 6, K-12ΔH1Δtrp (pLB10); lane 7, K-12ΔH1Δtrp (pLB10) grown at 28 °C; lane 8, K514 (pGI502); lane 9, 2 µg of purified Bt2 protein from clone K514 (pGI502); lane 10, 3 µg *B.t. berliner* crystals. Standard molecular mass markers (BioRad) are on both side. (B) Lanes as in A but with different amounts: lanes 1–4, extracts from 2×10^8 cells; lane 5, 1×10^8 cells; lanes 6–8, 5×10^6 cells; lane 9, 0.2 µg purified Bt protein; lane 10, 0.3 µg *B.t. berliner* crystal protein

Total cell extract of *E. coli* K514 (pGI502) revealed an intense protein band with apparent molecular mass of 130 kDa on SDS-PAGE which was not present in K514 containing the pUC8 plasmid without insert. This protein, termed Bt2, comigrates with one of the major crystal proteins of *B.t. berliner* in SDS-PAGE (Fig. 2A, lane 8). It represented between 5% and 10% of the total protein content in K514 (pGI502). It was present as a precipitate in *E. coli* and could be solubilized under the conditions required to solubilize *B.t.* crystals. The relationship of purified Bt2 protein with *B.t.* crystal proteins was analysed: in Western blotting, Bt2 reacted

strongly with a rabbit anti-(*B.t. berliner* crystal) serum (Fig. 2B, lane 9) and in an ELISA, 8 out of 16 monoclonal antibodies generated against *B.t. berliner* crystal proteins were reactive with the purified Bt2. Purified Bt2 showed toxicity levels comparable to those of solubilized crystals from *B.t. berliner* against *P. brassicae* and *M. sexta* larvae. (Table 1).

Nucleotide sequence of the toxin gene

Fig. 3 shows the restriction enzyme map of the 7.5-kb *Bam*HI-*Pst*I fragment from pGI502 containing the *bt2* gene. To localize the gene on this fragment, production of Bt2 protein by deletion mutants generated by the enzymes *Hpa*I, *Kpn*I and *Xba*I was monitored, using immunoassays. The *Hpa*I-deleted plasmid still encoded an intact 130-kDa protein indicating that the Bt2 toxin encoding sequence is localized on the 4343-bp *Hpa*I-*Pst*I fragment. The DNA sequence of this fragment (Fig. 3) shows one large open reading frame starting at an ATG codon at position 1 and ending at a TAA termination codon at position 3466. It codes for a protein of 1155 amino acids with a predicted molecular mass of 130533 Da, which agrees well with the molecular mass of Bt2 as determined in SDS-PAGE. Bt2 protein from K514 (pGI502) was additionally purified by DEAE-Sephadex ion-exchange chromatography and Sephacryl gel filtration. The amino acid sequence of the 20 N-terminal residues of this purified protein was determined by gas-phase sequencing [20]. This sequence, Xaa-Asp-Asn-Asn-Pro-Asn-Ile-Asn-Glu-Xaa-Ile-Pro-Tyr-Asn-Xaa-Leu-Xaa-Asn-Pro, is identical to the amino acid sequence deduced from the nucleotide sequence (Fig. 6) (Xaa indicates residues for which unambiguous identification was not possible).

A 60-kDa toxic polypeptide generated through proteolytic degradation of the 130-kDa Bt2 protein

The delta-endotoxins of *B.t.* are believed to be protoxins which are degraded by insect gut proteases into smaller active toxins [11]. We therefore investigated whether smaller toxic polypeptides could be generated from purified Bt2 by proteolytic cleavage with either trypsin or chymotrypsin. At defined time intervals, aliquots were analysed on SDS-PAGE. The 130-kDa Bt2 protein is rapidly degraded by trypsin or chymotrypsin and yields a major polypeptide of 60 kDa after a 10-min digestion at 37°C. This 60-kDa polypeptide is relatively resistant to further degradation by both enzymes (over a 2-h period), indicating that it constitutes a protease-resistant core within the Bt2 protein. The 60-kDa tryptic polypeptide was purified by gel filtration and its insect toxicity was determined. On a molar basis, it was equally toxic to *P. brassicae* larvae as intact 130-kDa Bt2 (Table 1). The N-terminal sequence of the 60-kDa tryptic core was determined by gas-phase sequencing as Ile-Glu-Thr-Gly-Tyr-Thr-Pro-Ile-Asp-Ile-Xaa-Leu allowing its unambiguous location in the original Bt2 sequence, starting from residue number 29 (see Fig. 5).

Delineation of the minimum gene fragment encoding an active toxin

Derivatives of the Bt2 gene that contain different deletions at the 5′ or the 3′ end were constructed and expressed in *E. coli*. We analysed the produced polypeptides to delineate the minimal Bt2 fragment required for insect toxicity.

The toxin gene was placed under transcriptional control of the λP_L promotor. The resulting plasmid pLB10 directed

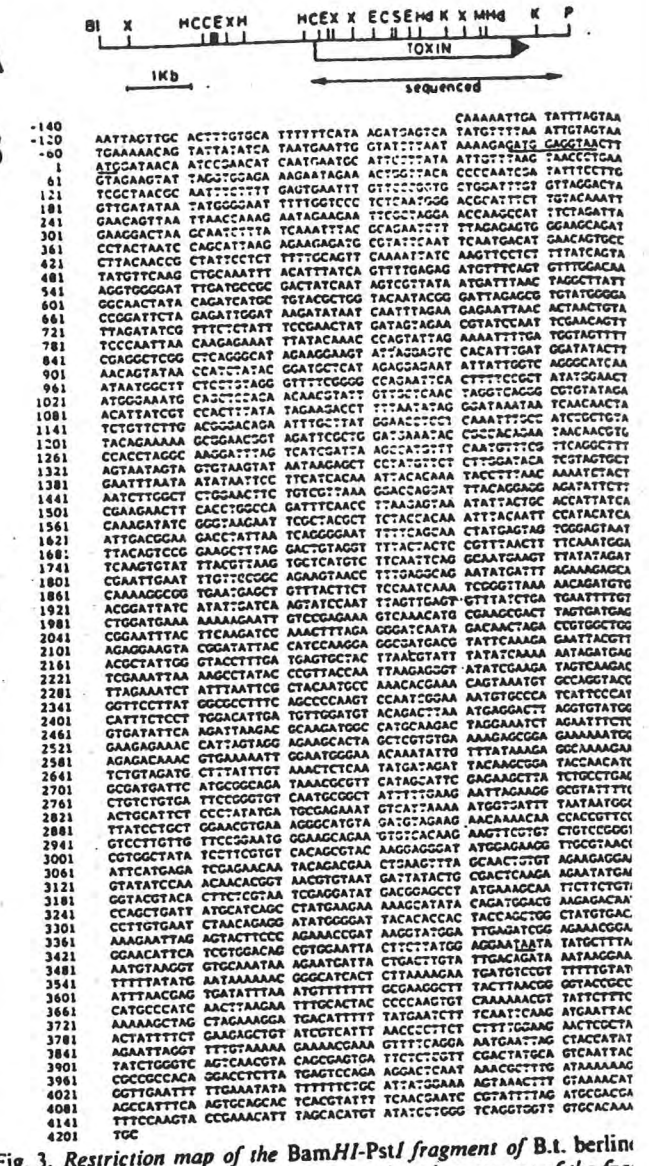
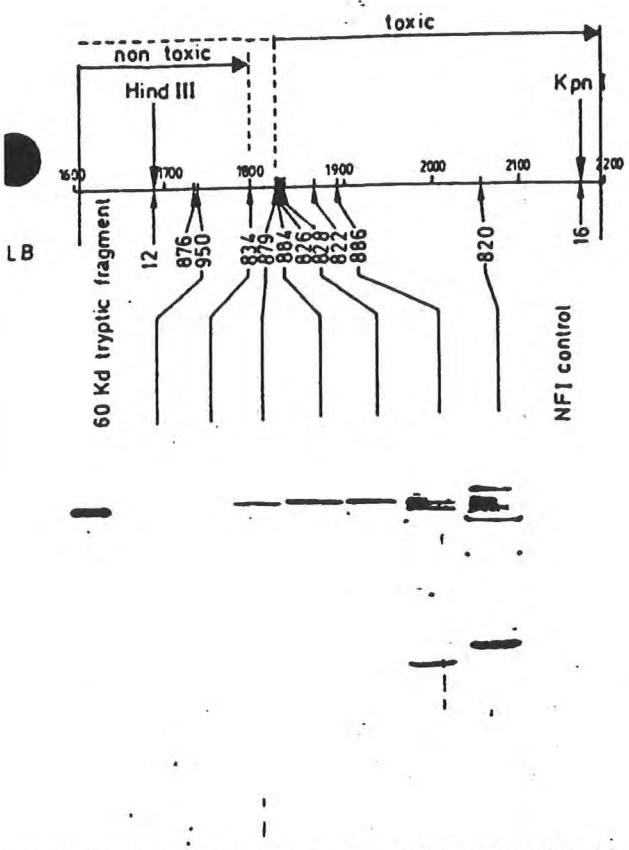


Fig. 3. *Restriction map of the* Bam*HI-*Pst*I fragment of* B.t. berlin *plasmid DNA in pGI502 (A) and the nucleotide sequence of the fra ment (B).* (A) The sequenced region is indicated. The arrow represen the toxin gene. X = *Xba*I; E = *Eco*RI; C = *Cla*I; Bl = *Bam*H H = *Hpa*I; Hd = *Hin*dIII; S = *Sac*I; K = *Kpn*I; M = *Mlu* P = *Pst*I. (B) The ATG initiation codon, the TAA stop codon ar the Shine and Dalgarno sequence of *bt2* are underlined

high level expression of the Bt2 protein after inductic (Fig. 2A, lanes 6 and 7; Fig. 2B, lanes 6 and 7). We sul sequently constructed pLBKm25 (see Materials and Met ods), a derivative of pLB10, which allowed the isolation of series of 3′ deletions of the *bt2* gene. Fig. 4 shows the positio of the different deletion end points. One of these deletio (pLB16), comprises *bt2* gene sequences up to the *Kpn*I si at position 2170. *E. coli* K-12ΔH1Δtrp containing pLB produces a 80-kDa polypeptide reactive with anti-(*B.t. be liner* crystal) serum in Western blotting. When assayed on *brassicae* larvae, this protein exhibited toxicity levels simil to those of intact Bt2 protein (Table 2). In another deleti clone, pLB12, the *bt2* gene extends to the *Hin*dIII site position 1692 and encodes a 60-kDa protein, still detectal as a faint band in Western blotting. However, extracts fr this clone were completely nontoxic towards *P. brassi* larvae (Table 2). These results suggest that the gene fragm



ig. 4. *Localisation of the endpoints of the 3' deletions in* bt2. Positions f the 3' ends of deletion clones are indicated. Also shown is the result f a Western blotting with anti-(*B.t.* crystal) serum on extracts of of these clones

ncoding the active toxin is contained within the *Kpn*I deletion ragment but extends beyond the *Hin*dIII site. To determine he 3' end point of the minimal fragment precisely, random eletion mutants encoding N-terminal fragments of ecreasing size, starting from the *Kpn*I site, were constructed sing exonuclease *Bal*31. Nine derivatives which have their eletion end points in the *Hin*dIII-*Kpn*I region were analysed. xtracts from K-12ΔH1Δtrp (pLB879) and from larger clones ere fully toxic whereas extracts from K-12ΔH1Δtrp pLB834) and from smaller clones were completely non-toxic o *P. brassicae* larvae (Table 2). The presence of Bt2-like olypeptides in the extracts was verified using Western lotting with anti-(*B.t. berliner* crystal) serum (Fig. 4) and evealed that less Bt2-like antigen was present in K-2ΔH1Δtrp (pLB834) than in K-12ΔH1Δtrp (pLB879).

However this quantitative difference could not account or the complete lack of toxicity of K-12ΔH1Δtrp (pLB834) ver a 100-fold dilution range (Table 2). The deletion ndpoints in pLB879 and pLB834 were determined by DNA equencing (Fig. 5) and indicated that the critical end point f a DNA fragment encoding an active toxin maps between ositions 1798 and 1821 on the *bt2* gene.

The lack of toxicity in pLB834 could result from a higher usceptibility to proteolytic degradation of the pLB834-ndoded polypeptide. To investigate this possibility, we onstructed inframe fusions of the *bt2* gene fragments ained in pLB834 and pLB879 to the 5' end of *lacZ*, to uce pBZ12 and pBZ13 respectively. It has been shown hat stable β-galactosidase fusion proteins can be easily pro-luced in high quantities in *E. coli* [14, 27]. SDS-PAGE analysis howed that the *bt2-lacZ* fusion genes directed the production

Table 2. *Toxicity against 3rd instar larvae of* P. brassicae *of cell extracts of K-12ΔH1Δtrp strains containing different plasmids*
Toxicity was measured as percentage mortality after 4 days

Strain	Toxicity at dilution of		
	1	1 10	1/100
	%		
pLB16	100	100	6
pLB820	100	100	54
pLB822	100	100	80
pLB828	100	100	68
pLB826	100	100	96
pLB884	100	100	74
pLB879	98	50	8
pLB834	0	2	2
pLB950	6	0	2
pLB876	2	0	4
pLB12	0	0	0
pBZ12	100	74	22
pBZ13	4	0	0
pHH10	10	0	0
pRB10	100	100	100
pRB210	6	2	4

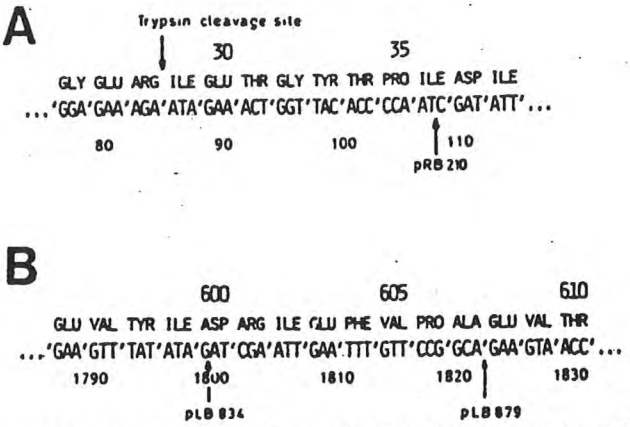


Fig. 5. *The sequence around the 5' end (A) and the 3' end (B) of the smallest toxin-encoding* bt2 *gene fragment.* The end points of *bt2* in the deletion clones and the N-terminus of the 60-kDa trypsin-cleavage product are indicated

of high amounts of protein with a high molecular mass as expected for the fusion protein (Fig. 2A, lanes 2 and 3). Reactivity with anti-(*B.t. berliner* crystal) serum (Fig. 2B, lanes 2 and 3) and anti-(β-galactosidase) serum (not shown) in Western blotting confirmed the presence of both Bt2 and β-galactosidase determinants. The toxicity of pBZ12 was comparable to that of pLB879, whereas the equally stable fusion protein encoded by pBZ13 was completely nontoxic (Table 2).

To delineate the 5' border of the gene fragment encoding an active toxin, we constructed pRB210 which contains a 5' deletion in the *bt2* gene lacking the first 36 codons. A polypeptide encoded by pRB210 would start only eight amino acids beyond residue 29, the N-terminus of the fully active 60-kDa processed toxin. Clones containing pRB210 and pRB10, a control plasmid containing a full-length *bt2* gene, both produced proteins of the expected size that reacted with anti-Bt2 antiserum (Fig. 2A, lanes 4 and 5; Fig. 2B, lanes 4 and 5). Strain K-12ΔH1Δtrp (pRB10) was fully toxic whereas

K-12ΔH1Δtrp (pRB210) was completely non-toxic to *P. brassicae* larvae (Table 2). Thus the N-terminus of the minimal toxic polypeptide is localized between amino acid positions 29 and 37 (Fig. 5).

Taken together, these data show that the minimal toxic fragment of the Bt2 protein is a 60-kDa polypeptide delineated by residues 29 and 37 at the N-terminus and amino acids 601–607 at the C-terminus (Fig. 5).

DISCUSSION

A plasmid-encoded crystal protein gene from *B.t. berliner* 1715 has been cloned and expressed in *E. coli*. The complete nucleotide sequence was determined and found to encode a polypeptide of 1155 amino acids. In *E. coli* this gene directs the synthesis of a 130-kDa protein, Bt2, which shows biochemical properties similar to the 130-kDa *B.t. berliner* 1715 crystal protein. The toxicity of purified Bt2 against *M. sexta* and *P. brassicae* larvae is similar to that of the original crystal proteins produced by *B.t. berliner* 1715.

Comparison of the complete amino acid sequence of Bt2 with other deduced amino acid sequences from cloned crystal proteins genes of *B.t.* strains *kurstaki* HD1 (Bt kur HDI) [28], *kurstaki* HD73 (Bt kur HD73) [10] and *sotto* (Bt sotto) [9] leads to the following observations (Figs 6 and 7).

a) The four toxins show a similar molecular mass (130 kDa) and exhibit extensive sequence homology.

b) Bt kur HD1 and Bt sotto are nearly identical over the total length of the published Bt sotto sequence (12 substitutions in 934 amino acids).

c) The four proteins are almost identical from residues 1 to 282 (I in Fig. 7).

d) From positions 283 to 458, Bt2 and Bt kur. HD73 are nearly identical (two substitutions in 174 amino acids, II_A in Fig. 7 but different from Bt kur. HD1 and Bt sotto at 65 positions (II_B in Fig. 7).

e) From residues 467 to 723, Bt2 is nearly identical to Bt kur. HD1 and Bt sotto (three substitutions in 266 amino acids, III_A in Fig. 7) but different from Bt kur. HD73 (III_B in Fig. 7), particularly between positions 467 and 611 (94 substitutions in 144 residues).

f) From positions 724 to 1155, Bt kur. HDI is almost identical to Bt kur. HD73 (three substitutions 431 amino acids, IV_B in Fig. 7) but different from Bt2 (IV_A at position 793 with an apparent deletion of 25 residues in Bt2 and between amino acids 1054 and 1117 where 18 out of 63 residues differ). Interestingly, the corresponding 75-bp sequence which is deleted in *bt2* is flanked by an 8-bp direct repeat of the sequence (AAAGTGTG) in the other three genes. Such a direct repeat might allow excision of the fragment by homologous recombination, leaving one copy of the repeat as is found in the *bt2* sequence, except for a 1-bp substitution (position 2379). Alternatively, such direct repeats could arise from an insertion event.

The general picture emerging from this sequence comparison is represented in Fig. 7 and shows that stretches of nearly identical sequences can be identified, which are recombined in different ways in the respective 130-kDa polypeptides. Thus, the crystal protein genes in the various *B.t.* strains may have evolved through homologous recombination events giving rise to toxins with distinct structural and possibly also functional properties. Homologous recombination between different crystal genes is likely to occur since they are located on transmissible plasmids [29]. The observation that IS elements flank the crystal genes in *B.t. berliner* 1715 [12] and *B.t. kurstaki* HD1 [30] also suggests a high mobility of these genes.

Despite considerable differences between the coding sequences, the 5′ regions upstream of the initiation codon are completely identical in the four genes over at least 140 bp. This suggests a similar regulation of the expression of all four genes.

Hydrophobicity analysis of the deduced amino acid sequence of Bt2 (Fig. 7) reveals that the whole N-terminal half of the Bt2 protein is more hydrophobic than the C-terminal half. In addition, two highly hydrophobic regions are present near the N-terminus of the protein (amino acid positions 29–80 and 137–172). Whether they play a critical role in the toxin/cell membrane interaction, as was found for other protein toxins [31], remains to be determined.

Large-molecular-mass *B. thuringiensis* crystal proteins have been shown to be 'protoxins' which are processed by insect gut proteases. The *in vitro* digestion of these crystals, using either insect midgut juice or well defined proteolytic enzymes, has been described and a variety of toxic polypeptides ranging in molecular mass from 160 kDa [2] to as small as 1 kDa have been reported [32, 33]. Other data indicate that polypeptides in the 55–70-kDa range can be generated by proteolytic degradation from the original crystal proteins using different enzymes [4, 34], or by spontaneous degradation of solubilized crystal proteins [35]. We have purified a polypeptide of 60 kDa, obtained by tryptic digestion of the Bt2 protein, showing the same toxicity as the intact Bt2 protein. The N-terminus matches the Bt2 sequence starting from amino acid 29. Nagamatsu et al. [34] also isolated a trypsin-resistant toxic core peptide of 58 kDa from *B.t. dendrolimus*. The N-terminal amino acid sequence was identical to the one determined for the 60-kDa Bt2 fragment.

To determine the minimal gene fragment still encoding an active polypeptide toxin we have used deletions of the *bt2* gene. The minimal fragment was found to be an approximately 60-kDa polypeptide, delineated at its N-terminus between amino acid positions 29 and 37 and at its C-terminus between positions 599 and 607. Interestingly, this fragment largely overlaps with the trypsin-resistant 60-kDa polypeptide. Indeed two putative trypsin-cleavage sites are present which could give rise to an approximately 60-kDa polypeptide, starting at position 29, namely Arg-601 and Arg-619. Our data are in agreement with those of Schnepf et al. [28] who also localized the active toxic fragment from a cloned *B.t. kurstaki* crystal protein, in the N-terminal half of the protein.

Based on the data obtained with the different deletions, we conclude that removal of a few amino acids from either the N-terminus or the C-terminus of this 60-kDa polypeptide results in a complete loss of toxicity. The absence of toxicity observed in C-terminal deletions up to position 599 is also associated with a relative protease sensitivity of these truncated Bt2 proteins. This suggests that toxicity requires a specific three-dimensional conformation which is disturbed by deleting only a few terminal residues.

The generation of very small toxic peptides by insect gut proteases has been reported [20, 21]. However, our present data show that it is very unlikely that active toxic fragments smaller than the 60-kDa fragment can arise from the Bt2 protein. Experiments are now in progress to identify the functional domains of the toxin using monoclonal antibodies specific for well defined regions on the Bt2 molecule.

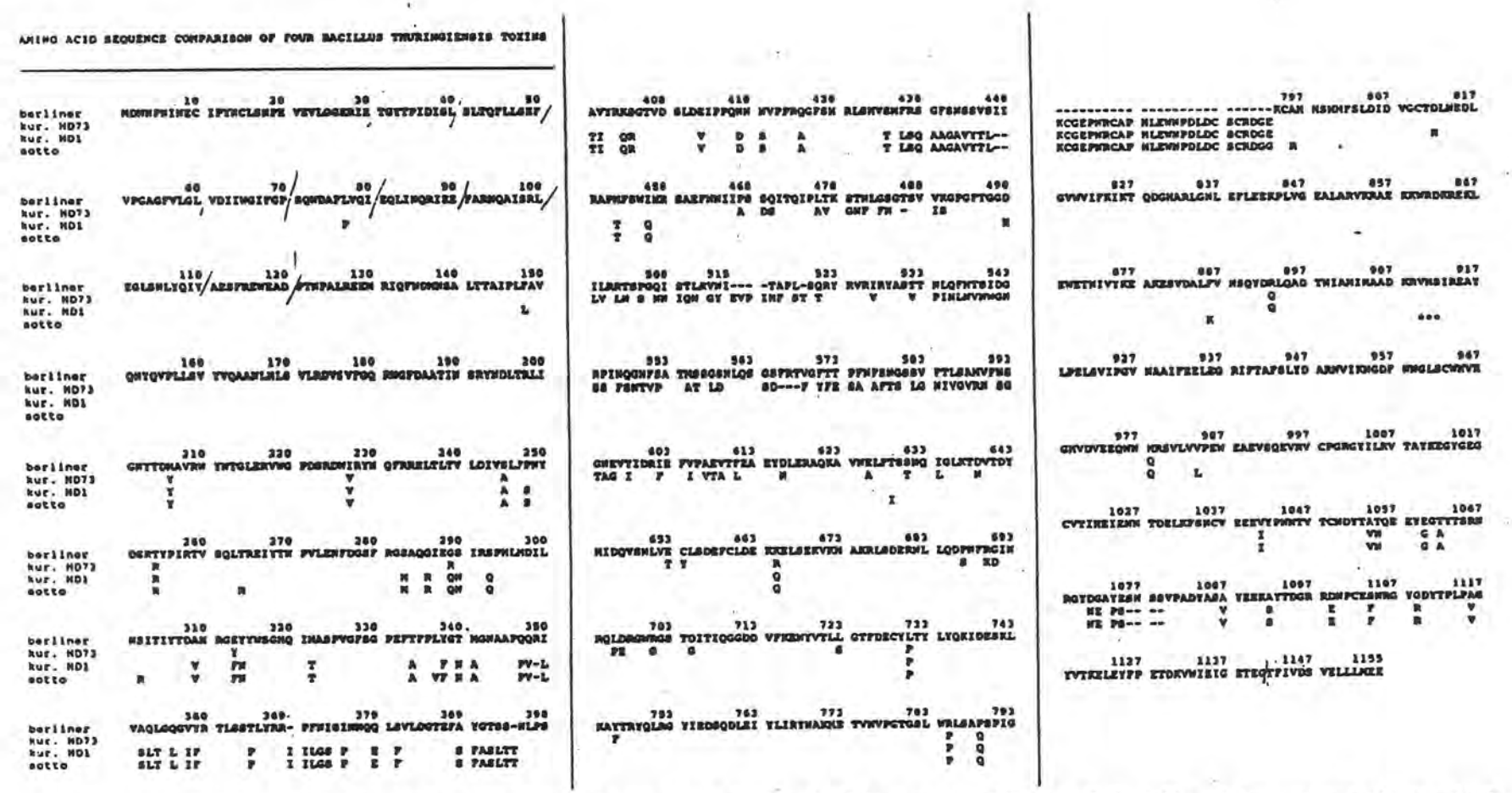


Fig. 6. *Amino acid sequence comparison of Bt2 and three other crystal proteins.* Sequences were aligned to the Bt2 sequence by adding gaps (indicated by −) where necessary. Amino acids in the other proteins are shown only when differing from the corresponding amino acids in Bt2. berliner = Bt2; kur. HD73 = *B.t. kurstaki* HD73 [10]; kur. HD1 = *B.t. kurstaki* HD1 [15]; sotto = *B.t. sotto* [9]

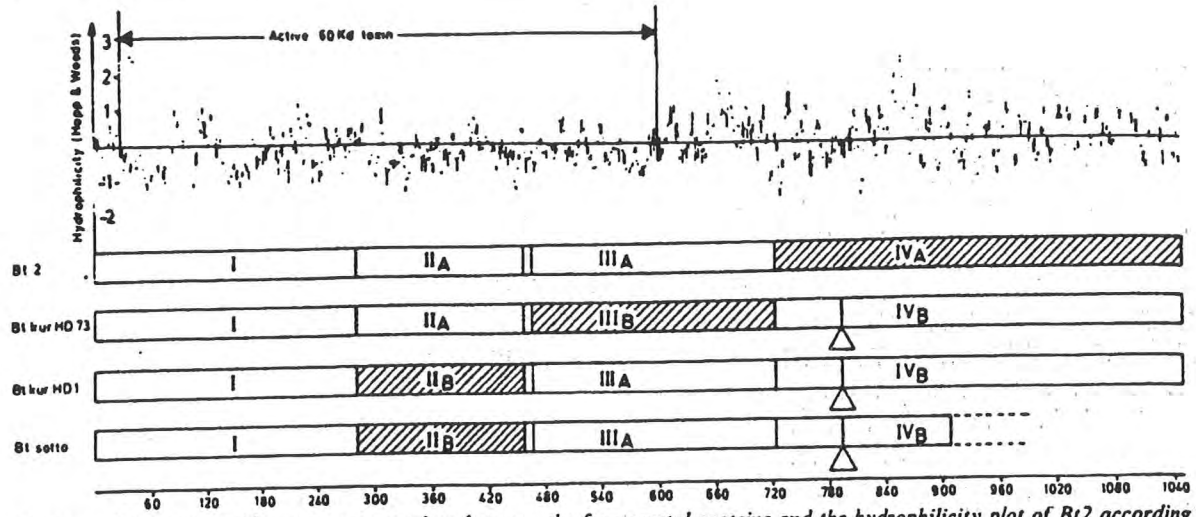


Fig. 7. *Diagram of the amino acid sequence comparison between the four crystal proteins and the hydrophilicity plot of Bt2 according to Hopp and Woods [36].* The location of the smallest active toxin in the amino acid sequence is shown. The roman numbers I–IV refer to the sequence 'blocks' mentioned in the text. The triangle indicates the sequence of 25 amino acids not present in Bt2. The hatched areas represent the amino acid stretches showing features unique for the protein considered

We thank Carine Matthijs and Paul Janssen for technical assistance, Ludo Van Rompuy for the computer analysis, Jan Leemans for critically reviewing and improving the manuscript, Karin Tenning for typing and Karel Spruyt and Albert Verstraete for making the figures. We particularly thank D. Degheele for useful advice concerning the insect testing procedures. J.V. is Research Associate for the Belgian National Fund for Scientific Research; H.V. is indebted to the *Belgian Institut voor wetenschappelijk onderzoek voor nijverheid en landbouw* for a fellowship.

REFERENCES

1. Dulmage, H. T. (1980) in *Microbial control of pests and plant diseases 1970–1980* (Burges, H. D., ed.) pp. 193–222, Academic Press, London.
2. Huber, H. E., Luthy, P., Ebersold, H.-R. & Cordier, J. L. (1981) *Arch. Microbiol. 129*, 14–18.
3. Calabrese, D. M., Nickerson, K. W. & Lane, L. C. (1980) *Can. J. Microbiol. 26*, 1006–1010.
4. Lilley, M., Ruffel, R. N. & Somerville, H. J. (1980) *Gen. Microbiol. 118*, 1–11.
5. Gonzalez, J. M., Brown, B. J. & Carlton, B. C. (1982) *Proc. Natl Acad. Sci. USA 79*, 6951–6955.
6. Kronstad, J. W., Schnepf, H. E. & Whiteley, H. R. (1983) *J. Bacteriol. 154*, 419–428.
7. Schnepf, H. E. & Whiteley, H. R. (1981) *Proc. Natl Acad. Sci. USA 78*, 2893–2897.
8. Klier, A., Fargette, F., Ribier, J. & Rapoport, G. (1982) *EMBO J. 1*, 791–799.
9. Shibano, Y., Yagamata, A., Nakamara, N., Iizuka, T., Sugisaki, H. & Takanami, M. (1985) *Gene 34*, 243–351.
10. Adang, M. J., Staver, M. J., Rochelean, T. A., Leighton, J., Barker, R. F. & Thompson, D. V. (1985) *Gene 36*, 289–300.
11. Thomas, W. E. & Ellar, D. J. (1983) *J. Cell Sci. 60*, 181–197.
12. Mahillon, J. & Delcour, J. (1984) *J. Microbiol. Methods 3*, 69–76.
13. Colson, C., Glover, S. W., Symonds, N. & Stacey, K. (1965) *Genetics 52*, 1043–1050.
14. Zabeau, M. & Stanley, K. R. (1982) *EMBO J. 1*, 1217–1224.
15. Botterman, J. (1986) PhD thesis, State University of Ghent.
16. Vieira, J. & Messing, J. (1982) *Gene 19*, 259–268.
17. Reiss, B. (1984) *EMBO J. 3*, 3317–3322.
18. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory Press, New York.
19. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.
20. Hewick, R. M., Hunkapiller, M. W., Hood, L. E. & Dreyer, W. J. (1981) *J. Biol. Chem. 256*, 7990–7997.
21. Hunkapiller, M. W. & Hood, L. E. (1983) *Methods Enzymol. 91*, 486–493.
22. Vandekerckhove, J., Bauw, G., Puype, M., Van Damme, J. & Van Montagu, M. (1985) *Eur. J. Biochem. 152*, 9–19.
23. Helfman, D. M., Feramisco, J. R., Fiddes, J. C., Thomas, G. P. & Hughes, S. H. (1983) *Proc. Natl Acad. Sci. USA 80*, 31–35.
24. Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl Acad. Sci. USA 76*, 4350–4354.
25. Engvall, E. & Pesce, A. J. (1978) *Scand. Immunol. Suppl. 7*.
26. Bell, R. A. & Joachim, F. G. (1976) *Ann. Entomol. Soc. Am. 69*, 365–373.
27. Guo, L. H., Stepien, P. P., Tso, J. Y., Brousseau, R., Narang, S., Thomas, D. Y. & Wu, R. (1984) *Gene 29*, 251–254.
28. Schnepf, H. E., Wong, H. C. & Whiteley, H. R. (1985) *J. Biol. Chem. 260*, 6264–6272.
29. Gonzalez, J. M., Dulmage, H. T. & Carlton, B. C. (1981) *Plasmid 5*, 351–365.
30. Whiteley, H. R., Kronstad, J. W. & Schnepf, H. E. (1985) in *Molecular biology of microbial differentiation* (Hoch, J. A. & Sellow, P., eds) pp. 225–229, American Society of Microbiology, Washington DC.
31. Middlebrook, J. L. & Dorland, R. B. (1984) *Microbiol. Rev. 48*, 199–221.
32. Fast, P. G. & Martin, W. G. (1980) *Biochem. Biophys. Res. Commun. 95*, 1314–1320.
33. Faust, R. N., Hallam, G. M. & Travers, R. S. (1974) *J. Invertebr. Pathol. 24*, 365–373.
34. Nagamatsu, Y., Itai, Y., Hathnaka, C., Funatsu, G. & Hayashi, K. (1984) *Agric. Biol. Chem. 48*, 611–619.
35. Bulla, L. A., Jr, Bechtel, D. B., Kramer, K. J., Shetna, Y. I., Aronson, A. I. & Fitz-James, P. C. (1980) *Rev. Microbiol. 8*, 147–203.
36. Hopp, T. P. & Woods, K. R. (1981) *Proc. Natl Acad. Sci. USA 78*, 3824–3828.

demonstrated a shielding effect of the ribosome on nascent proteins; ribosome-bound nascent chains 30–35 residues long were found to be protected from external proteolytic attack. Protamine, 31–32 residues long, may therefore be protected from attack by a putative methionine-removing enzyme until the chain is completed and released into the cytoplasm. Other contributory circumstances may be the difficulty of cleavage of the Met–Pro peptide bond; in general, X–Pro bonds are resistant to many proteases with the possible exception of imidodipeptidase[20]; but this enzyme is specific only for dipeptides. Also, as shown by Ling and Dixon[16], at the stage of differentiation at which protamine synthesis is maximal, there has been a considerable reduction of the spermatid cytoplasm and ribosome content. The levels of cytoplasmic enzymes, including the methionine-removing enzyme, may therefore become limiting at this stage.

The involvement of methionine in the initiation of protein synthesis in eukaryotic cells is not yet firmly established although the presence of a formylatable species of Met-tRNA with the properties of an initiator tRNA in mouse ascites cells[31,32] is certainly consistent with such a role. Our observations of methionine involvement in the synthesis of the very unusual, sperm-specific polypeptide, protamine, suggest that such a mechanism may be of widespread importance in eukaryotes, and in the special circumstances of protamine biosynthesis, where removal of the N-terminal methionyl residue may become limiting, the transient incorporation of this amino-acid is readily observed.

Received May 12; revised June 19, 1970.

[1] Marcker, K. A., and Sanger, F., *J. Mol. Biol.*, 8, 835 (1964).
[2] Adams, J. M., and Capecchi, M. R., *Proc. US Nat. Acad. Sci.*, 55, 147 (1966).
[3] Webster, R. E., Engelhardt, D. L., and Zinder, N. D., *Proc. US Nat. Acad. Sci.*, 55, 155 (1966).
[4] *Cold Spring Harbor Symp. Quant. Biol.*, 31 (1966); 34 (1969).
[5] Smith, A. E., and Marcker, K. A., *J. Mol. Biol.*, 38, 241 (1968).
[6] Miller, R. L., and Schweet, R., *Arch. Biochem. Biophys.*, 125, 632 (1968).
[7] Heywood, S. M., *Nature*, 225, 696 (1970).
[8] Caskey, C. T., Redfield, B., and Weissbach, H., *Arch. Biochem. Biophys.*, 120, 119 (1967).
[9] Takeishi, K., Ukita, T., and Nishimura, S., *J. Biol. Chem.*, 243, 5761 (1968).
[10] Caskey, C. T., Beaudet, A., and Nirenberg, M., *J. Mol. Biol.*, 37, 99 (1968).
[11] Ando, T., and Watanabe, S., *Int. J. Protein Res.*, 1, 221 (1969).
[12] Ling, V., Trevithick, J. R., and Dixon, G. H., *Canad. J. Biochem.*, 47, 51 (1969).
[13] Marushige, K., and Dixon, G. H., *Develop. Biol.*, 19, 397 (1969).
[14] Ingles, C. J., Trevithick, J. R., Smith, M., and Dixon, G. H., *Biochem. Biophys. Res. Commun.*, 22, 627 (1966).
[15] Ingles, C. J., and Dixon, G. H., *Proc. US Nat. Acad. Sci.*, 58, 1011 (1967).
[16] Sung, M., and Smithies, O., *Biopolymers*, 7, 39 (1969).
[17] Marushige, K., Ling, V., and Dixon, G. H., *J. Biol. Chem.*, 244, 5953 (1969).
[18] Jergil, B., and Dixon, G. H., *J. Biol. Chem.*, 245, 425 (1970).
[19] Gray, W. R., in *Methods in Enzymology*, 11, 139 (Academic Press, New York, 1967).
[20] Anderson, G. W., Zimmerman, J. E., and Callahan, F. M., *J. Amer. Chem. Soc.*, 86, 1839 (1964).
[21] Sheehan, J. C., and Yang, D. H., *J. Amer. Chem. Soc.*, 80, 1154 (1958).
[22] Baliga, B. S., Pronczuk, A. W., and Munro, H. N., *J. Biol. Chem.*, 244, 4480 (1969).
[23] Kaji, A., Kaji, H., and Novelli, G. D., *J. Biol. Chem.*, 240, 1185, 1192 (1965).
[24] Momose, K., and Kaji, A., *J. Biol. Chem.*, 241, 3294 (1966).
[25] Kaji, K., Novelli, G. D., and Kaji, A., *Biochim. Biophys. Acta*, 76, 474 (1963).
[26] Ling, V., and Dixon, G. H., *J. Biol. Chem.*, 245, 3035 (1970).
[27] Rich, A., Eikenberry, E. F., and Malkin, L. I., *Cold Spring Harbor Symp. Quant. Biol.*, 31, 303 (1966).
[28] Malkin, L. I., and Rich, A., *J. Mol. Biol.*, 26, 329 (1967).
[29] Schmidt, P. J., Mitchell, B. S., Smith, M., and Tsuyuki, H., *Gen. Comp. Endocrinol.*, 5, 197 (1965).
[30] Davis, N. C., and Smith, E. L., *J. Biol. Chem.*, 224, 261 (1957).
[31] Smith, A. E., and Marcker, K. A., *Nature*, 226, 607 (1970).
[32] Brown, J. C., and Smith, A. E., *Nature*, 226, 610 (1970).

Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4

by
U. K. LAEMMLI
MRC Laboratory of Molecular Biology, Hills Road, Cambridge

Using an improved method of gel electrophoresis, many hitherto unknown proteins have been found in bacteriophage T4 and some of these have been identified with specific gene products. Four major components of the head are cleaved during the process of assembly, apparently after the precursor proteins have assembled into some large intermediate structure.

BACTERIOPHAGES of the T-even type are complex structures containing many different proteins and specified by many genes. Using an improved technique of electrophoretic separation I have found that the phage particle contains at least twenty-eight components, eleven of which are in the head. In the course of identifying the genes specifying these proteins I discovered that four major components of the head, the product of gene 22, 23, 24 and a protein called IP of unknown genetic origin are cleaved during the process of assembly. The head of bacteriophage T4 is therefore no longer a self-assembly system in the narrow sense, because the bonding properties of the various components become altered during the assembly process.

The product of gene 23 is the principal protein component of the head of bacteriophage T4 (refs. 1–3). Two minor components of unknown genetic origin have also been found in capsids[2,3]. Besides the product of gene 23, the products of genes 20, 21, 22, 24, 31, 40 and 66 are required to determine the size and shape of the head-shell (refs. 4 and 5 and unpublished work of F. A. Eiserling, E. P. Geiduschek, R. H. Epstein and E. J. Metter). To several of these genes shape-specifying functions have been tentatively assigned[5]. Gene 22 is associated with the diameter selecting (initiation) process of head formation, gene 66 with the elongation of the particle and genes 20 and 40 with the formation of the hemispherical cap. Gene 31 somehow modifies or activates the major subunit for ordered assembly[6]. Ten more proteins, the products of genes 2, 4, 13, 14, 16, 17, 49, 50, 64 and 65, are thought to control later steps in head formation[7].

Structural Components of the Phage

Many phage proteins can be separated with our improved method of disk-electrophoresis in sodium dodecyl sulphate (SDS). This system, to be described in detail elsewhere (U. K. L. and J. V. Maizel), combines the high resolution power of disk-electrophoresis[8] with the capability of SDS to break down proteins into their individual polypeptide chains[9]. The proteins are also separated according to their molecular weight as was first reported for a continuous system[10]. All the proteins the genetic

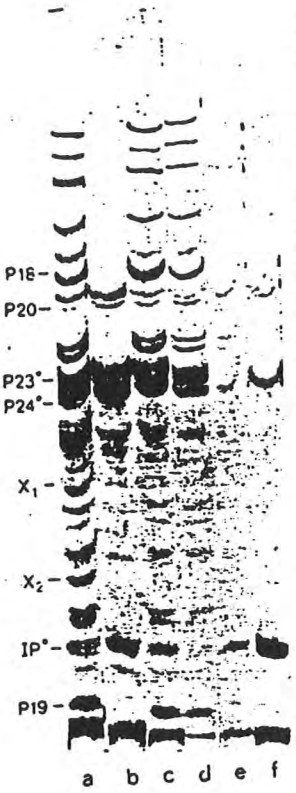


Fig. 1. Autoradiogram of ^{14}C-labelled T4 phage proteins separated in acrylamide gels. ^{14}C-amino-acid-labelled preparations were analysed in 10 per cent acrylamide gels containing SDS. *a*, Wild type lysate; *b*, purified heads; *c*, purified phage particles; *d*, purified "ghosted" phage particles; *e*, supernatant of "ghosted" phage particles; *f*, "early labelled" phage particles. The prefix P is used to designate the protein of a particular gene: for example, P20 stands for the product of gene 20; the asterisk indicates that this protein is derived from a large precursor protein and has become modified during head assembly.

^{14}C-labelled lysates. Ten ml. cultures of *Escherichia coli* B^{b} (the restrictive host for phage carrying amber mutations) in 'M9' medium[18] grown at 37° C to 2×10^{8} cells/ml. were infected with the various phages at a multiplicity of five and superinfected at 8 min with the same phage and the same multiplicity to ensure lysis inhibition. Two μCi of ^{14}C-amino-acid mixture ('CFB 104', Radiochemical Centre, Amersham) with a specific activity of 45 mCi/m atom was added to each culture 13 min after the first infection. Three ml. of 3 per cent 'casamino-acids' mixture (Difco) was added to each sample at 30 min, and the infected cells were concentrated by low speed centrifugation 35 min following infection. The pellets were drained and directly resuspended in "final sample" buffer (see gel electrophoresis).

^{14}C-labelled phage and head particles. Ten ml. cultures were grown and infected as described above (^{14}C-labelled lysates). The double mutant (B255–E18) in genes 10 and 18 was used for production of tailless heads. Ten μCi of the ^{14}C-amino-acid mixture was added 13 min after the first infection to each culture. The infected cells were concentrated by a low speed centrifugation 35 min after infection and the pellet was resuspended in 1 ml. neutral phosphate buffer containing 10^{-3} M $MgSO_4$, 20 μg deoxyribonuclease and a drop of chloroform. The pellet was resuspended by repeated pipetting, and incubated for 15–30 min at room temperature before layering on a CsCl step gradient[3]. The latter was prepared in tubes for a Spinco 'SW50' rotor, with 0·8 ml. layers, and the following densities starting at the bottom of the tube: 1·55, 1·46, 1·38 and 1·29 g/cm^3. Furthermore, a 10 per cent sucrose solution (in neutral phosphate buffer and 10^{-2} M $MgSO_4$) was layered on the last CsCl step to prevent precipitation of soluble proteins at the CsCl interface. Centrifugation was for 1 h at 40,000 r.p.m. The phage and heads, which form a sharp visible band two-thirds down the tube, were collected through the bottom of the tube and dialysed against water. The band containing the heads was always viscous, indicating that the heads lost their DNA in CsCl although the DNA was still confined within the band. Occasionally the heads lost their DNA before centrifugation particularly in more concentrated lysates, which had to be treated with deoxyribonuclease for a much longer time. "Early labelled" phages were prepared identically, but the label was added 1 min and chased 6 min after infection by the addition of 3 ml. of 3 per cent 'casamino-acids'.

"Ghosted" phage particles and supernatant of "ghosted" phage particles. Crystalline NaCl was added to ^{14}C-labelled purified phage particles in water to a final concentration of 5 M. The preparation was then repeatedly frozen in a solid CO_2-acetone bath and thawed in warm water. This procedure releases the DNA and its internal proteins from the phage head. The sample was dialysed against neutral phosphate buffer containing 10^{-3} M $MgSO_4$ and the DNA digested by the addition of a small amount of crystalline deoxyribonuclease. The ghost was finally separated from the soluble proteins (internal proteins) by centrifugation and layered on a step gradient identical to that already described. Centrifugation was at 35,000 r.p.m. for 1 h. The top of the gradient containing the internal proteins was collected with a pipette and the ghosts, which band at an approximate density of 1·3 g/cm^3, were collected through the bottom of the tube. SDS at a final concentration of 2 per cent was added to both samples and the samples were dialysed against water containing 2 per cent SDS.

Gel electrophoresis. Gels containing 3 per cent (stacking gel), 8·0 per cent or 10 per cent acrylamide were prepared from a stock solution of 30 per cent by weight of acrylamide and 0·8 per cent by weight of *N,N'-bis*-methylene acrylamide. The final concentrations in the separation gel were as follows: 0·375 M Tris-HCl (*p*H 8·8) and 0·1 per cent SDS. The gels were polymerized chemically by the addition of 0·025 per cent by volume of tetramethylethylenediamine (TEMED) and ammonium persulphate. Ten cm gels were prepared in glass tubes of a total length of 15 cm and with an inside diameter of 6 mm. The stacking gels of 3 per cent acrylamide and a length of 1 cm contained 0·125 M Tris–HCl (*p*H 6·8) and 0·1 per cent SDS and were polymerized chemically in the same way as for the separating gel. The electrode buffer (*p*H 8·3) contained 0·025 M Tris and 0·192 M glycine and 0·1 per cent SDS. The samples (0·2–0·3 ml.) contained the final concentrations ("final sample buffer"): 0·0625 M Tris–HCl (*p*H 6·8), 2 per cent SDS, 10 per cent glycerol, 5 per cent 2-mercaptoethanol and 0·001 per cent bromophenol blue as the dye. The proteins were completely dissociated by immersing the samples for 1·5 min in boiling water[4]. Electrophoresis was carried out with a current of 3 mA per gel until the bromophenol blue marker reached the bottom of the gel (about 7 h). The proteins were fixed in the gel with 50 per cent trichloroacetic acid (TCA) overnight, stained for 1 h at 37° C with a 0·1 per cent Coomassie brilliant blue solution made up freshly in 50 per cent TCA. The gels were diffusion-destained by repeated washing in 7 per cent acetic acid. Autoradiograms of gels were prepared by a modified version (U. K. L. and J. V. Maizel, unpublished) of Fairbanks *et al.*[19] (autoradiograms are shown in Figs. 1–7).

origin of which was determined are labelled in Fig. 1, and their molecular weights are listed in Table 1. At least 28 bands can be distinguished in the autoradiogram of radioactively labelled, purified phage particles (Fig. 1*c*). The 28 proteins found in dissociated phage particles do not include proteins with molecular weights less than about 15,000. Those proteins are not sieved in gels of 10 per cent acrylamide (unpublished results of U. K. L. and J. V. Maizel) and migrate with the marker dye. In gels of higher acrylamide concentration another three low molecular weight proteins have been separated (results not shown). The largest protein in the phage has an approximate molecular weight of at least 120,000 and no label stays at the top of the gel, indicating complete dissociation of the particles by the method used.

Eleven of these 28 proteins are found in the purified head preparation (Fig. 1*b*). The remaining 17 proteins absent from the head but present in the whole phage pattern are presumably structural proteins of the tail and tail fibres. The complete absence of these proteins from the head gel pattern indicates the high degree of purity of the preparation. The classification of the proteins into tail and head components may not be valid for proteins making up the head to tail junction.

Only two minor proteins besides the major components P23 were found by others[2,3] in phage capsids purified in the same way but fractionated in urea gels. The larger number of proteins found in SDS gels is to be expected, for at least 46 genes are known to affect T4 morphogenesis[4,7] although the proteins of these genes may not all be incorporated into the particle.

The gel pattern of a total lysate of wild type infected cells radioactively labelled at late times is presented in Fig. 1*a*. Note that most of the resolved proteins that are synthesized late in infection are structural phage components. A few rather intense bands (X_1 and X_2), however, are completely missing in the phage particles, also demonstrating that the separation of the phage particles from the soluble proteins is complete.

When phages are subjected to osmotic shock a number of proteins are released[11]. The principal one is IP*, an internal protein (Fig. 1). It is present in complete phage particles, is extracted from the phage particles by freezing and thawing in high salt concentrations and is quantitatively recovered in the supernatant. The supernatant fraction also contains many minor components which are found both in phages as well as phage ghosts, indicating that the

Table 1. MOLECULAR WEIGHTS OF PHAGE PROTEINS DETERMINED BY COMPARING THEIR MOBILITY IN SDS GELS WITH THOSE OF MARKER PROTEINS WITH KNOWN MOLECULAR WEIGHTS[10,11]

Gene product	Observed value in SDS gels	Published value
P18	69,000	50,000[a]
P20	63,000	
P23	56,000	
P23*	46,500	46,000[21]
P24	45,000	
P24*	43,500	
P22	31,000	
IP	23,500	
IP*	21,000	
P19	18,000	

The following marker proteins were used: serum albumin (68,000), γ-globulin, heavy chain (50,000), ovalbumin (43,000), γ-globulin, light chain (23,500) and TMV (17,000), to calibrate 10 per cent acrylamide gels. In this improved gel system the distances of migration of the various marker proteins relative to the distance of migration of bromophenol blue are also a linear function of the logarithm of the molecular weight of the marker proteins, as has been described[10,11]. Radioactively labelled phage proteins were mixed with unlabelled marker proteins before electrophoresis; the distance of migration of the phage proteins was determined from the autoradiogram and those of the marker proteins from the stained gel. It is assumed that the phage proteins also separate in SDS gels solely according to their molecular weight.

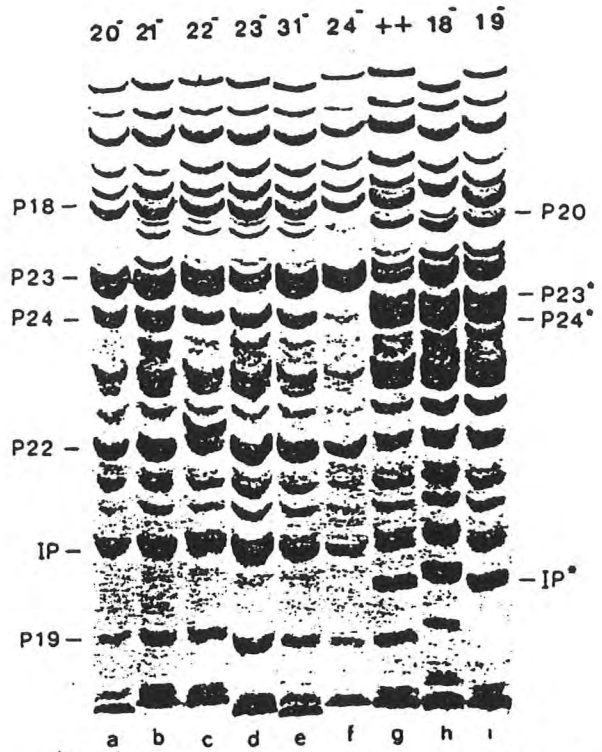


Fig. 2. Identification of gene products on 10 per cent acrylamide gels. The lysates were prepared as described in Fig. 1 and analysed on 10 per cent acrylamide gels. *a*, Gene 20-defective lysate, mutant N50; *b*, gene 21-defective lysate, mutant N90; *c*, gene 22-defective lysate mutant B270; *d*, gene 23-defective lysate, mutant H11; *e*, gene 31-defective lysate, mutant N54; *f*, gene 24-defective lysate, mutant N65; *g*, lysate from wild type infected cell; *h*, gene 18-defective lysate, mutant E18; *i*, gene 19-defective lysate, mutant E1137. Mutants N90 (gene 21), E18 (gene 18) and E1137 (gene 19) carried second mutations in gene 10 (mutant B255). The following amber fragments may be detected. A rather intense band just below P23 is consistently seen in 18-defective lysates, and is presumed to be the amber fragment of mutant E18 in gene 18. Furthermore, the gel pattern of defective lysates of all double mutants in genes 18–10, 21–10, and 19–10 possess a band just above P23 which is probably the amber fragment of mutant B255 in gene 10. It is striking that the amber fragment of mutant B270 gene 22 behaves anomalously in the gel: it migrates more slowly than the wild type product, although SDS gels are known to separate on the basis of molecular weights. This anomalous behaviour of certain proteins will be discussed elsewhere (U. K. L. and J. V. Maizel).

separation of the released proteins and the ghosted particles by centrifugation was not complete. Some of these proteins, however, are extracted into the supernatant quantitatively by the freezing and thawing procedure.

Only the protein IP*, and perhaps some low molecular weight proteins which migrate with the marker dye, are structural phage components synthesized at early times (early proteins). This can be seen in the gel pattern of purified phage labelled at early times only (Fig. 1*f*). IP* is also labelled at late times. All the other structural phage components are synthesized only late in infection. Some of the principal late proteins show up on the autoradiogram, probably because of residual incorporation of radioactive amino-acids following the chase with unlabelled amino-acids.

Identification of Gene Products

The products of genes 18, 19, 20, 22, 23 and 24 were identified by comparing the gel pattern of extracts of cells infected with wild type phage with those infected with amber mutants in various genes. The identification of the tail and tail fibre proteins will be described later (J. King and U. K. L.). Amber mutations produce only fragments of the protein chain of the mutant gene on infection of restrictive bacteria[1]. These fragments migrate differently in the gel from the complete proteins. The molecular weights of the proteins identified are listed in Table 1. Fig. 2 is the autoradiogram from dried and sliced gels for various mutants. The product of gene 20 is identified by its absence in the gel pattern of a 20-defective* lysate (Figs. 2 and 3*a*), and the product of gene 22 by its absence in the gel pattern of a 22-defective lysate (Figs. 2 and 3*c*). Note the amber fragment of mutant B270 in gene 22. This fragment was also identified by Hosoda and Levinthal[12] in urea gels (see legend to Fig. 2).

The product of gene 23 is easily identified by its absence in the gel pattern of a 23-defective lysate (Fig. 2*d*). It can also be seen that P23 overlaps with two minor tail components. If the 23-defective lysate is analysed on gels of lower acrylamide concentration another important observation is made. A band, P23*, which is detected in variable amounts in the other head defective lysates, is completely missing in the 23-defective lysate (Fig. 3*d*). This band, P23*, overlaps with P24 in Fig. 2, but is better separated from P24 in less concentrated gels (Fig. 3). As with the product P23, the product of P24 was identified by its absence in the gel pattern of a 24-defective lysate (Figs. 2*f* and 3*e*).

So far, no missing bands have been found in the gel patterns of 31 or 21-defective lysates (Fig. 2*b* and *e*), but analysis of the 21-defective lysate on gels of lower acrylamide concentration (Fig. 3*b*), which resolves higher molecular weight proteins better, clearly shows that a band is missing. This protein, however, is the product of gene 10, a baseplate gene. The mutant N90 in gene 21 in fact carries a second mutation in gene 10 (mutant B255).

In comparing the gel pattern of the head-defective lysates (Fig. 2*a–f*) with that of wild type (Fig. 2*g*), further important differences are observed, which shed light on the precursor–product relationship of the head components.

The principal fraction of the gene 23 product has a molecular weight of 56,000 in all the head-defective lysates, but in wild type or tail-defective lysates it appears at the position of P23*, with a molecular weight of 46,500. Small but significant amounts of P23* are also observed in head-defective lysates (Fig. 3*a–f*). In lysates prepared identically about 20 per cent of the total P23 is converted to P23* in the 20-defective lysate, 10 per cent in 21 and 2–3 per cent in 22, 24 and 31-defective lysates, as determined from densitometer tracing of the autoradiographs.

The bands P22 and IP are both absent or considerably less intense in the wild type gel pattern (each overlaps with two other proteins). A new band, IP*, is seen at the bottom of the gel, which is completely missing in the head defective lysates (Fig. 2*a–f*).

The band P24, which is found in all head-defective lysates is missing in the wild type pattern, but a new band, P24*, which migrates slightly faster is observed. This is difficult to visualize in Fig. 2, but will become evident in Fig. 6.

Also included in Fig. 2 is the gel pattern of two tail defective lysates. The product of gene 18 (mol. wt 69,000), the principal protein of the tail sheath[13], is identified by its absence in an 18-defective lysate (Fig. 2*h*) and the product of gene 19 (mol. wt 18,000) by its absence in a 19-defective lysate (Fig. 2*i*). P19 is thought to be the chief component of the tail tube[14]. This demonstrates that the differences in the head proteins are not related to tail attachment, for the gel patterns of the tail-defective lysates are identical with that of wild type.

Evidence will be presented that the proteins P23, P22, P24 and IP are cleaved in wild type infected cells and are precursors to proteins P23*, P24* and IP* found in the final head structure. (The cleavage product of P22 was not detected.) This precursor–product conversion is strongly inhibited by mutations in genes 20, 21, 22, 23, 24 and 31. Two important conclusions can be drawn: (*a*) the aberrant head-related structures—single and

* Henceforth, lysates of cells infected at restrictive conditions with amber mutants in various genes will be referred to as, for instance, a "21-defective lysate", where the amber mutant used was in gene 21.

multi-layered polyheads[4,5,15], τ-particles[4,5,16] and lumps[6] known to be produced in these mutant infected cells—chiefly consist of the precursor protein P23; (*b*) because P22 is not cleaved in these mutant infected cells, but is required for polyhead and τ-particle formation, as has been established by genetic means[5], it is strongly suggested that P22 is incorporated into these structures as such.

Kinetics of the Cleavage Reactions

The following experiments were designed to study the precursor relationship of the proteins P23, P24 and IP with P23*, P24* and IP*, respectively. In this experiment the infected cells were pulse-labelled with radioactive amino-acids for a short period (1 min) and the modification of the various proteins was then followed by analysis of the samples taken at intervals in SDS gels. The results of autoradiography of the dried gels are presented in Fig. 4.

Cleavage of P23. It is readily seen in Fig. 4 that most of the 23 protein is at the position of P23 immediately following the pulse of the radioactive amino-acids. It then rapidly disappears, and a new band, P23*, appears simultaneously. That P23 is cleaved and gives rise to P23* is suggested by the fact that both are principal components and this is reinforced by the absence of P23 and P23* from the gel patterns of a 23 defective lysate (Fig. 3*d*). The kinetics of the cleavage reaction P23→P23* are plotted in Fig. 5*a*. Cleavage is very rapid: about 50 per cent of the precursor is cleaved within the first 2 min following chase of the label. P23* appears at about the same rate, in a satisfactorily correlated way. The total labelled protein in P23 and P23* is also plotted in Fig. 5*a*. The total label increases during the first minute, which reflects the completion time of the chase of the labelled amino-acid, but finally the total falls off by 25–30 per cent. This final decrease of the total labelled protein can be nicely explained, for a cleavage from about 56,000 to 46,500 corresponds to a loss of about 20 per cent by weight of protein.

Cleavage of P22 and IP. In the pulse-chase experiment of Fig. 4, two other protein bands, P22 and IP, disappear with time. (The disappearance of P22 in wild type infected cells has also been observed by M. Showe, personal communication.) A new band, IP*, appears at the bottom of the gel pattern. The kinetics of these cleavage reactions are plotted in Fig. 5*b*. P22 disappears with approximately the same initial rate as P23: about 50 per cent is cleaved 2–3 min following chase of the labelled amino-acids. I have not found a band in the gel pattern which may be derived from P22. Hosoda and Levinthal[12] reported indirect evidence that P22 is a structural phage component, but they considered the possibility that P22 might become altered during head formation.

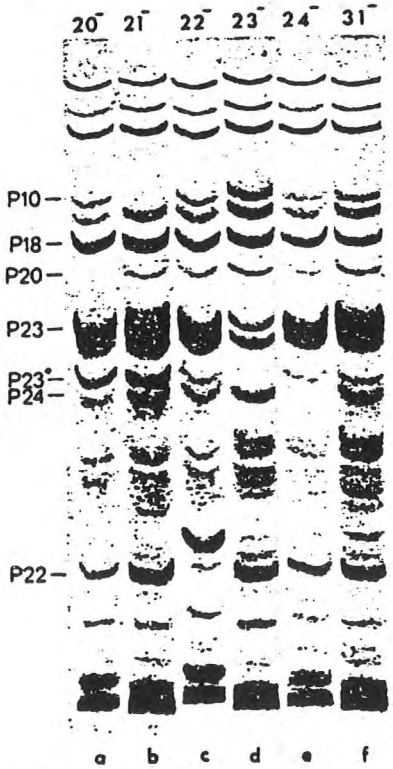


Fig. 3. Identification of gene products on 8 per cent acrylamide gels. ^{35}S-labelled lysates were prepared as described (Fig. 1) and analysed on a 8 per cent acrylamide gels. *a*, Gene 20-defective lysate, mutant N50; *b*, gene 21-defective lysate, mutant N90; *c*, gene 22-defective lysate, mutant B270; *d*, gene 23-defective lysate, mutant H11; *e*, gene 24-defective lysate, mutant N65; *f*, gene 31-defective lysate, mutant N54.

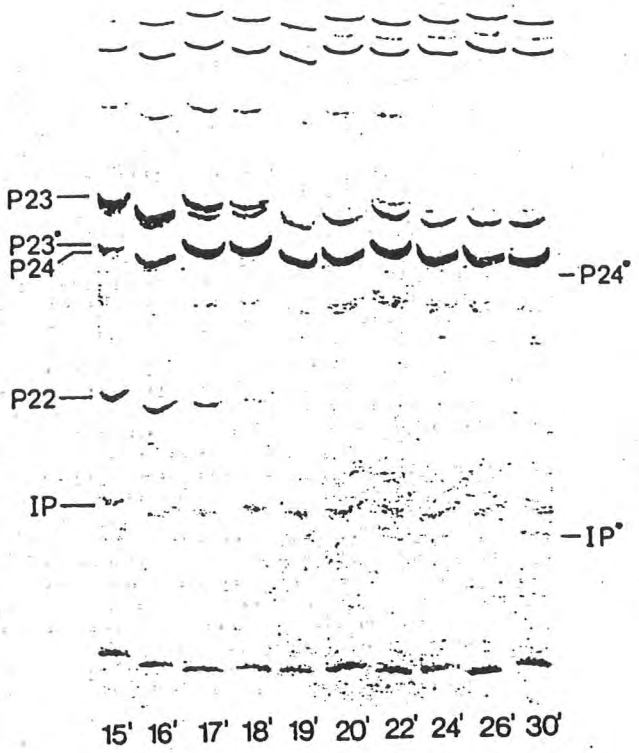


Fig. 4. Cleavage of products of genes 22, 23, 24 and protein IP (10 per cent acrylamide gels). A 10 ml. culture grown at 37° C was infected with a double mutant defective in genes 10 and 18 (mutant B255 and E18) as described (Fig. 1). The radioactive amino-acid mixture (10 μCi) was added 14 min after the first infection, and chased 1 min later with an excess of unlabelled amino-acids (final concentration 1 per cent). The chase of the label was verified by measuring the counts in the total TCA precipitable proteins. One ml. samples were prepared at intervals after the chase and immediately frozen in a solid CO_2-acetone bath. SDS was added after thawing to a final concentration of 2 per cent and the samples were carefully dialysed into 2 per cent SDS in water. The samples were finally mixed with an equal volume of twice concentrated "final sample buffer" and boiled for 1 min before electrophoresis. The sampling time is indicated at the bottom of the gels. All the preliminary experiments were done with wild type phage, but this experiment was performed with the double mutant in genes 10 and 18 in view of plans for future experiments.

Evidence for the precursor–product conversion IP→IP* is provided by the observation that the disappearance of IP and the appearance of IP* is coordinated in time (Fig. 5*b*). Furthermore, the total label lost at the IP position is recovered in the IP* band. The reaction is slower than that of P23 and P22. Only about 50 per cent of the total label at the IP position disappears. This could be explained either by another protein band overlapping with the IP band or synthesis in excess.

IP* cannot arise from P22. IP* must be derived from a precursor which is synthesized early, for I have shown (Fig. 1) that IP* is strongly labelled in the "early labelled" phage preparation. P22, however, is reported to be synthesized at late times only[12], which I have confirmed. The precursor conversion of IP→IP* has also been observed in a pulse-chase experiment in which the label was added between 4 and 5 min following infection, thus labelling only early proteins (results not shown). The band IP was easily recognized in these gels and the disappearance of IP and the appearance of IP* were again correlated in time. This quantitative agreement and the absence of other unaccounted changing bands in the gel pattern support the argument for the IP→IP* relationship. Of course, a final proof awaits chemical analysis.

It was also observed that, although the pulse was performed between 4 and 5 min after infection, cleavage of IP starts only

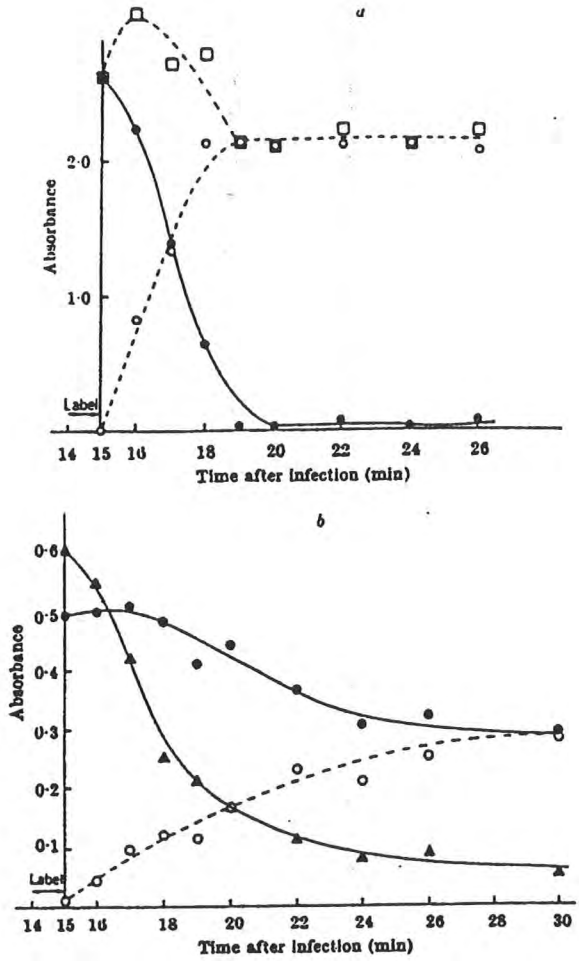


Fig. 5. Kinetics of cleavage of P23, P22 and IP. The kinetics of cleavage were measured using a microdensitometer (double beam recording microdensitometer, Joyce-Loebl) to record the autoradiogram. The exposure of the autoradiogram was chosen so that the absorbance of the band to be measured did not exceed 1 unit. The abscissa represents the integrated absorbance over the relevant peaks. *a*: ●—●, integrated absorbance over the P23 peak; ○---○, integrated absorbance over the P23* peak; □---□, total absorbance in P23 and P23*. *b*: ▲—▲, integrated absorbance over peak P22; ●—●, integrated absorbance over peak IP; ○---○, integrated absorbance over peak IP*.

at late times (after 17 min). Phage assembly starts at about this time and it is therefore thought that the cleavage of IP is linked to phage assembly.

Cleavage of P24. P24* (mol. wt 43,500) appears coordinately with P23* and P22* (Fig. 4) (the kinetics are not plotted). The precursor product relationship P24→P24* is more difficult to demonstrate, because P24 migrates only slightly faster than P23*, but the following experiment proves that P24 is missing in a pulse-chase wild type lysate. P24 separates somewhat better from P23* in a gel of lower concentration. The samples of the pulse-chase experiment were analysed on 8 per cent acrylamide gels and four time points are presented in Fig. 6. P24 is easily distinguished from the small amount of P23* existing immediately after the chase of the radioactive label (Fig. 6, 15 min). P24 disappears at the same time as P24* appears while P23* increases. One might argue that P24 is obscured by the heavy P23* band. This possibility was excluded by adding a lysate (23-defective lysate) containing P24 to the final samples of the pulse-chase experiment. P24 was then detected and it can be concluded that measurements of P24 are reliable, thus showing that P24 most likely gives rise to P24*. The integrated absorbance values over these two bands are about equal, but the small loss of protein weight by the P24→P24* reaction is not likely to be detected by the densitometric measurements. P24 could not give rise to IP*, because the total label in IP* is two or three times larger than that of P24 and P24 is synthesized late. Moreover, the precursor relationship of P24→P24* is considerably strengthened by recent observations on head maturation genes (my unpublished results). P24 does not seem to be cleaved at all in 50-defective cells and, indeed, no P24* is found. Cleavage of P23, P22 and IP does occur in 50-defective cells, although at a reduced rate.

Fate of the Small Fragments

Where are the small fragments of these cleavage reactions? The expected molecular weights for the small fragments stemming from P23, IP and P24 would be about 10,000, 2,500 and 1,500 respectively. Peptides of this size are not sieved on 10 per cent acrylamide gels and migrate with the marker dye (unpublished results of U. K. L. and J. V. Maizel). Attempts to find at least the 10,000 molecular weight fragment from P23 on gels of higher acrylamide concentration have failed. Possibly the fragments are further broken down to undetectable sizes. Fragments of P22 also have not been detected. Acid-soluble components which are derived from an acid-insoluble precursor are known to exist in T4 infected cells[17]. Two are associated with the phage particle and they are released with the DNA from the head upon osmotic shock[17]. The genetic determinant of one of these internal peptides has been mapped recently and lies in the neighbourhood of genes 20 and 21 (ref. 18). My results definitely rule out gene 20, which is incorporated unmodified into normal phage. Unfortunately, I have not discovered the product of gene 21 in the gel pattern. These results do not rule out the possibility that one of the internal proteins is derived from the small cleavage fragment of P22, P23 or P24. The appearance of the internal peptides seems indeed to be coordinated with the cleavage reactions of P23, P22 and P24. Genes 20, 21, 22, 23, 24 and 31, which affect the cleavage of P23, P22 and IP, are known also to affect the appearance of the internal peptides[17].

It has been pointed out to me by S. Brenner and A. Stretton that the cleavage point must occur at the N-terminal end of the P23 protein. In establishing the co-linearity of gene 23 and its polypeptide[1], they observed peptides in 23-defective lysates which contain the amber fragments, but these are absent in wild type lysates or in purified phage particles. These peptides may be derived from the N-terminal end of the amber fragment, which is cleaved off in the protein P23*. These observations also suggest that the small cleavage fragment, with an expected molecular weight of about 10,000, is fragmented to even smaller pieces.

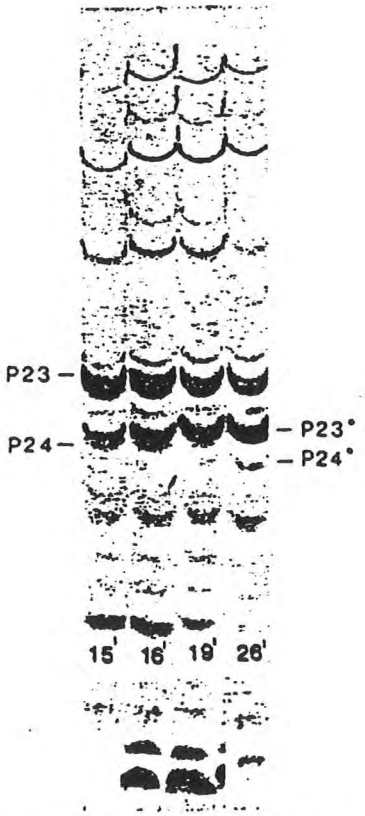


Fig. 6. Cleavage of the product of gene 24 (8 per cent acrylamide gels). Some samples (15, 16, 19 and 26 min) of the pulse chase experiment of Fig. 4 were analysed on 8 per cent acrylamide gels. Only the relevant part of the gel pattern is shown.

Cleavage occurs in a Large Structure

The observation that all the proteins P20, P21, P22, P23, P24 and P31 are required for efficient cleavage of P23, P22, P24 and IP suggests that these precursor proteins aggregate first to form an oligomeric structure, and are cleaved subsequently, rather than being cleaved first, and then assembled. The following experiment supports this view. The experiment is based on the observation that most of the precursor proteins are soluble and monomeric in SDS at room temperature, but that phage particles are not totally disrupted, as is also true in urea[12]. The gel pattern of purified phage treated with SDS at room temperature is compared with completely degraded phage, boiled for 1 min in SDS (Fig. 7*a* and *b*). Only a small fraction of P23* is extracted from phage with SDS at room temperature. Most of the proteins stay at the top of the gel or enter the gel as high molecular weight aggregates. Some proteins are not extracted at all. Note, however, that a few proteins are almost quantitatively extracted from the phage particles. The samples from the pulse-chase experiment treated in SDS at room temperature only are shown in Fig. 7. P23 disappears with time but no P23* appears, suggesting that P23 enters an SDS-resistant structure before being cleaved. Of course, these experiments cannot rule out the possibility that P23* is converted to an SDS-resistant structure so rapidly that it is not detected. A high molecular weight protein appeared at the top of the gel, which could be an aggregate of P23 but only accounts for part of the label which disappears from P23. Most of the label stays at the top of the gel. The molecular weight of this structure must therefore be greater than 300,000, for such a molecular weight is excluded from these gels (unpublished results of U. K. L. and J. V. Maizel). It is possible that this structure has a capsid-like shape. IP* and P24* are not resolved in these gels. This is because of the high salt concentration in these samples ('M9' growth medium) which impairs the resolution of the gel in the low molecular weight region.

Maturation of the Head

My experiments demonstrate that the assembly of the head of bacteriophage T4 is not a simple, straightforward self-assembly, because several structural proteins are chemically altered at some stage of assembly. The uncleaved precursor protein P23 can, however, be polymerized into single and multilayered polyheads and τ-particles, if its cleavage is blocked as a result of mutation.

Investigations of genes 2, 4, 13, 14, 16, 17, 49, 50, 64 and 65, which supposedly control late steps in head formation, are in progress. It is interesting that the cleavage of P23, P22, P24 and IP seems to be normal in cells infected with phage carrying mutations in these genes with the exception of genes 2, 50 and 64 (my unpublished results).

Why are these structural head components cleaved? The finding that IP gives rise to an internal IP* sheds light on a possible consequence of the cleavage reactions. Internal proteins are thought to bind to DNA and the cleavage reactions may possibly trigger the necessary DNA–protein interactions, which result in orderly packing of the DNA within the shell of proteins. The following model may be proposed. The proteins P20, P21, P23, P24, P31 and IP form an intermediate structure (SDS-resistant) which combines with an end of a DNA strand. Cleavage of P23, P22, P24 and IP proceeds from the DNA attachment site. During this process more and more DNA binding sites may be formed at the inside of this structure, perhaps by the formation of IP*, thus winding up the DNA strand successively. The small acid-soluble poptides formed during this process may also interact with the packed DNA[17] to neutralize charges. My results do not decide whether DNA packing proceeds simultaneously with the polymerization of the head membrane or follows thereafter.

I thank Dr A. Klug for encouragement and facilities, my colleagues J. King, J. Maizel and S. Altman for many valuable suggestions (J. Maizel in particular for advice on gel techniques and J. King for help in preparing the manuscript), and S. Brenner (among others) for critically reading the manuscript. U. K. L. holds an EMBO fellowship.

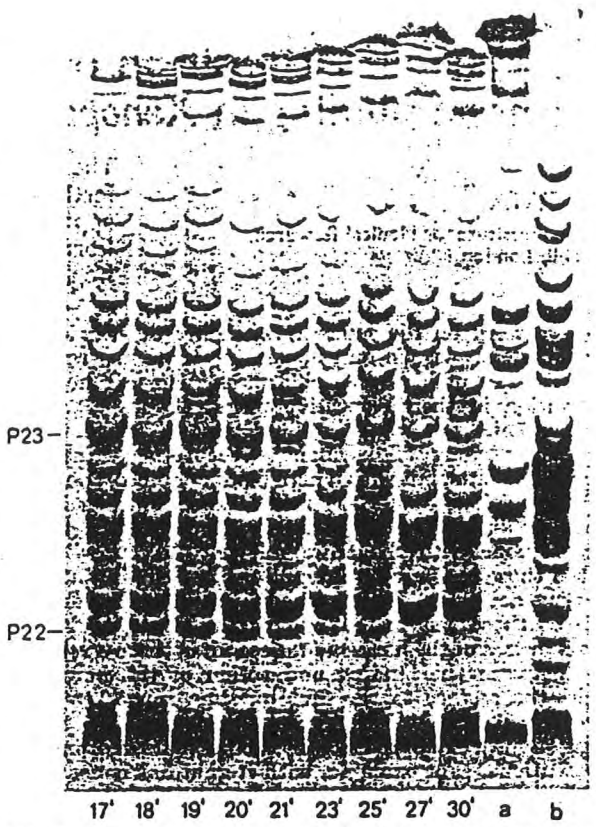


Fig. 7. Chase of P23 into a product stable to SDS at room temperature. A culture was infected and pulse labelled as described in Figs. 1 and 3 with the following changes. The culture was grown at 30° C, infected with wild phage, and labelled for 2 min from 15–17 min after infection. The samples were mixed at room temperature with an equal volume of twice concentrated "final sample buffer" without previous dialysis. The samples were then directly applied to the gels without being boiled. The sampling time is indicated at the bottom of the gels. As a control a purified phage preparation is analysed. *a*, Phage treated in SDS at room temperature only; *b*, phage boiled in SDS for 1·5 min.

Note added in proof. During the preparation of this manuscript I was informed that the alteration of P23 has also been observed by other workers: E. Kellenberger and C. Kellenberger-van der Kamp, *FEBS Lett.*, 8, 3, 140 (1970); R. C. Dickson, S. L. Barnes and F. A. Eiserling, *J. Mol. Biol.* (in the press); and J. Hosoda and R. Cone, *Proc. US Nat. Acad. Sci.* (in the press).

Received May 7, 1970.

[1] Sarabhai, A. S., Stretton, A. O. W., Brenner, S., and Bolle, A., *Nature*, 201, 13 (1964).
[2] Kellenberger, E., *Virology*, 34, 549 (1968).
[3] Baylor, M. B., and Roslansky, P. F., *Virology*, 40, 251 (1970).
[4] Epstein, R. H., Bolle, A., Steinberg, C. H., Kellenberg, E., Boy de la Tour, E., Chevalley, R., Edgar, R. S., Susman, M., Denhardt, G. H., and Lielausis, J., *Cold Spring Harbor Symp. Quant. Biol.*, 28, 375 (1963).
[5] Laemmli, U. K., Molbert, E., Showe, M., and Kellenberger, E., *J. Mol. Biol.*, 49, 99 (1970).
[6] Laemmli, U. K., Beguin, F., and Gujer-Kellenberger, G., *J. Mol. Biol.*, 47, 69 (1970).
[7] Edgar, R. S., and Wood, W. B., *Proc. US Nat. Acad. Sci.*, 55, 498 (1966).
[8] Davis, B. J., *Ann. NY Acad. Sci.*, 121, 404 (1964).
[9] Maizel, J. V., *Fundamental Techniques of Virology* (edit. by Habel, K., and Salzman, N. P.), chap. 32, 334 (Academic Press, New York, 1969).
[10] Shapiro, A. L., Viñuela, E., and Maizel, J. V., *Biophys. Biochem. Res. Commun.*, 28, 815 (1967).
[11] Minigawa, T., *Virology*, 13, 515 (1961).
[12] Hosoda, J., and Levinthal, C., *Virology*, 34, 709 (1968).
[13] King, J., *J. Mol. Biol.*, 32, 231 (1968).
[14] King, J., *FEBS Symp.* (edit. by Ochoa, S., Nachmansohn, D., Asensio, C., and Heredia, C. F.), 21, 156 (Academic Press, London, 1970).
[15] Favre, R., Boy de la Tour, E., Segre, N., and Kellenberger, E., *J. Ultrastruct. Res.*, 13, 318 (1965).
[16] Kellenberger, E., Eiserling, F. A., and Boy de la Tour, E., *J. Ultrastruct. Res.*, 21, 335 (1968).
[17] Eddleman, H. L., and Champe, S. P., *Virology*, 30, 471 (1966).
[18] Sternberg, N., and Champe, S. P., *J. Mol. Biol.*, 46, 337 (1970).
[19] Laemmli, U. K., and Eiserling, F. A., *Molec. Gen. Genet.*, 101, 333 (1968).
[20] Fairbanks, G., Levinthal, C., and Reeder, R. H., *Biochem. Biophys. Res. Commun.*, 20, 393 (1965).
[21] Larcom, L., Bendet, I., and Mumma, S., *Virology*, 41, 1 (1970).
[22] Weber, K., and Osborn, M., *J. Biol. Chem.*, 244, 4406 (1969).

PROTEIN MEASUREMENT WITH THE FOLIN PHENOL REAGENT*

BY OLIVER H. LOWRY, NIRA J. ROSEBROUGH, A. LEWIS FARR, AND ROSE J. RANDALL

(From the Department of Pharmacology, Washington University School of Medicine, St. Louis, Missouri)

(Received for publication, May 28, 1951)

Since 1922 when Wu proposed the use of the Folin phenol reagent for the measurement of proteins (1), a number of modified analytical procedures utilizing this reagent have been reported for the determination of proteins in serum (2–6), in antigen-antibody precipitates (7–9), and in insulin (10).

Although the reagent would seem to be recommended by its great sensitivity and the simplicity of procedure possible with its use, it has not found great favor for general biochemical purposes.

In the belief that this reagent, nevertheless, has considerable merit for certain application, but that its peculiarities and limitations need to be understood for its fullest exploitation, it has been studied with regard to effects of variations in pH, time of reaction, and concentration of reactants, permissible levels of reagents commonly used in handling proteins, and interfering substances. Procedures are described for measuring protein in solution or after precipitation with acids or other agents, and for the determination of as little as 0.2 γ of protein.

Method

Reagents—Reagent A, 2 per cent Na_2CO_3 in 0.10 N NaOH. Reagent B, 0.5 per cent $CuSO_4 \cdot 5H_2O$ in 1 per cent sodium or potassium tartrate. Reagent C, alkaline copper solution. Mix 50 ml. of Reagent A with 1 ml. of Reagent B. Discard after 1 day. Reagent D, carbonate-copper solution, is the same as Reagent C except for omission of NaOH. Reagent E, diluted Folin reagent. Titrate Folin-Ciocalteu phenol reagent ((11), Eimer and Amend, Fisher Scientific Company, New York) with NaOH to a phenolphthalein end-point. On the basis of this titration dilute the Folin reagent (about 2-fold) to make it 1 N in acid. Working standards may be prepared from human serum diluted 100- to 1000-fold (approximately 700 to 70 γ per ml.). These in turn may be checked against a standard solution of crystalline bovine albumin (Armour and

* Supported in part by a grant from the American Cancer Society on the recommendation of the Committee on Growth of the National Research Council.

Company, Chicago); 1 γ is the equivalent of 0.97 γ of serum protein (see below). Dilute solutions of bovine albumin have not proved satisfactory for working standards because of a marked tendency to undergo surface denaturation.

Procedure for Proteins in Solution or Readily Soluble in Dilute Alkali—(Directions are given for a final volume of 1.1 to 1.3 ml., but any multiple or fraction of the volumes given may be employed as desired[1].)

To a sample of 5 to 100 γ of protein in 0.2 ml. or less in a 3 to 10 ml. test-tube, 1 ml. of Reagent C is added. Mix well and allow to stand for 10 minutes or longer at room temperature. 0.10 ml. of Reagent E is added very rapidly and mixed within a second or two (see below). After 30 minutes or longer, the sample is read in a colorimeter or spectrophotometer. For the range 5 to 25 γ of protein per ml. of final volume, it is desirable to make readings at or near $\lambda = 750$ mμ, the absorption peak. For stronger solutions, the readings may be kept in a workable range by reading near $\lambda = 500$ mμ (Fig. 2). Calculate from a standard curve, and, if necessary, make appropriate correction for differences between the color value of the working standard and the particular proteins being measured (see below).

It is unnecessary to bring all the samples and standards to the same volume before the addition of the alkaline copper reagent, provided corrections are made for small differences in final volume. The critical volumes are those of the alkaline copper and Folin reagents.

If the protein is present in an already very dilute solution (less than 25 γ per ml.), 0.5 ml. may be mixed with 0.5 ml. of an exactly double strength Reagent C and otherwise treated as above.

Insoluble Proteins, etc.—Many protein precipitates, *e.g.* tungstate precipitates, will dissolve readily in the alkaline copper reagent. However, after proteins have been precipitated with trichloroacetic or perchloric acid, for example, they will dissolve rather poorly in the 0.1 N alkali of this reagent. They become even harder to dissolve if subsequently extracted with fat solvents, and still more so if dried at 100°.

It is not possible to cover all cases, but the following may be helpful in measuring the protein of acid precipitates. If the amount of protein is not great, so that it is spread rather thinly, it will usually dissolve in ½ hour or so in 1 N NaOH at room temperature. Therefore, one may add, for example, 0.1 ml. of 1 N NaOH to 5 to 100 γ of precipitated protein.

[1] For example, with the Klett colorimeter, transfer 25 to 500 γ of protein in not over 1 ml. volume to a colorimeter tube. Add water if necessary to make 1 ml. Add 5 ml. of Reagent C, and, after 10 minutes, 0.5 ml. of Reagent E. Readings are taken after 30 minutes with the No. 66 filter. If the readings are too high, substitute the No. 54 filter for sample, standards, and blanks.

After ½ hour or more, 1 ml. of Reagent D (no NaOH) is added, followed after 10 minutes by 0.1 ml. of diluted Folin Reagent E as usual.

With larger samples, or very stubborn precipitates, it may be necessary to heat for 10 minutes or more at 100° in 1 N alkali. Although this may lower the readings, they will be reproducible and can be measured with similarly treated standards.[2]

Microanalysis—With a Beckman spectrophotometer adapted to 0.05 ml. volume (12), as little as 0.2 γ of protein may be measured with reasonable precision. Aside from reducing the volumes of sample and reagents, the only necessary change is to use sufficiently slender tubes for the reaction. If the tubes are too large in diameter, low values will result. The following is illustrative of a procedure in which it is desired to precipitate the protein in order, for example, to measure an acid-soluble constituent of the same specimen. In this example, it is assumed that the sample volume is negligible. Otherwise a smaller volume of more concentrated trichloroacetic acid would be used.

To the sample containing 0.2 to 3 γ of protein in a tube of 3 mm. inner diameter and 4 cm. long,[3] are added 10 μl. of 5 per cent trichloroacetic acid.[4] After being mixed and centrifuged, 8 μl. of the supernatant fluid are removed. To the precipitate are added 5 μl. of 8 N NaOH. The sample is thoroughly mixed by tapping or "buzzing,"[5] and is covered

[2] Bovine serum albumin is especially difficult to redissolve after precipitation. Several 40 mg. samples were precipitated with trichloroacetic acid, washed with alcohol and isopropyl ether, and dried. These samples dissolved very slowly in 2 ml. of 1 N NaOH. However, after standing overnight, the protein appeared to be nearly all dissolved and aliquots gave readings 97 per cent of those obtained with non-precipitated samples. Other samples were heated for 30 minutes at 100° in 1 N NaOH. These samples dissolved and the solution turned slightly yellow. The final readings checked well but were only 82 per cent of those obtained with non-precipitated samples. Possibly, heating at lower temperature with the 1 N NaOH would have sufficed, although heating with weaker alkali would not have been effective, judging from other experience. The use of stronger alkali than 1 N did not appear to be an improvement.

[3] These tubes are cleaned by rinsing with dilute NaOH, boiling in half concentrated HNO_3, and rinsing several times in redistilled water. Filling or emptying of a beaker full of tubes (tall form of beaker without lip) is accomplished by slow centrifugation for a few seconds. For emptying, the tubes are transferred upside down to a second beaker with a false bottom of stainless steel screen. With the slow centrifugation required beakers will not be broken.

[4] Suitable micro pipettes are the Lang-Levy variety (13, 14). For entering these narrow tubes the bent tip must be especially short and slender.

[5] The tube is held at an angle against a rapidly rotating flattened rod or nail. Any high speed hand tool mounted in a clamp is satisfactory. The contents of the tube will mix violently without spilling (15). A similar effect may be had with a commercial rubber-tipped massage vibrator.

with a rubber cap or Parafilm. After 30 minutes, 50 μl. of Reagent D are added and the sample is mixed by "buzzing." After 10 minutes or more, 5 μl. of diluted Folin Reagent E are added with *immediate* "buzzing," and the samples are read after 30 minutes. Standards are perhaps best prepared by precipitating 5 μl. of 5, 10, 20, etc., mg. per cent solutions of serum protein with 5 μl. of 10 per cent trichloroacetic acid, with subsequent treatment as for the other samples.

EXPERIMENTAL

There are two distinct steps which lead to the final color with protein: (*a*) reaction with copper in alkali, and (*b*) reduction of the phosphomolybdic-phosphotungstic reagent by the copper-treated protein.

Reaction with Copper in Alkaline Solution—The salient features of this reaction follow. (1) The color obtained in the *absence* of copper is probably attributable entirely to the tyrosine and tryptophan content (16, 17), and this is not greatly increased by alkaline pretreatment ((4–6) and Table I). (2) In the *presence* of copper, alkaline treatment of proteins results in a 3- to 15-fold increase in color, but, in contrast, the presence of copper has only a small effect on the color obtained with free tyrosine and tryptophan (Herriott (17, 18) and Table I). (3) The reaction with copper, although not instantaneous, is nearly complete in 5 or 10 minutes at room temperature under the prescribed conditions. Heating to 100° or increasing the concentration of alkali accelerates the reaction with copper without changing the final color. (4) Pretreatment with alkali alone does not alter the subsequent reaction with copper in alkaline solution. Even pretreatment for an hour at 60° with 2 N NaOH, or for 5 minutes at 100° with 1 N NaOH, when followed by the usual copper treatment, has almost no effect on subsequent color. Prolonged heating with strong alkali will, however, decrease the final color.[2]

Although the alkaline copper reaction and the biuret reaction appear to be related, they are not strictly proportional, nor, with different proteins, is the amount of biuret color directly proportional to the increment caused by copper in the color with the Folin reagent (Table I).

A very small amount of copper is sufficient to give nearly maximum final color (Table II). The action does not appear to be catalytic. Assuming the simple relationship copper + protein $\rightleftharpoons$ copper-protein obtains, the data with low copper concentrations may be utilized to calculate an apparent dissociation constant of 3×10^{-6} mole per liter with a maximum of 1 mole of chromogenic protein-bound copper per 7 or 8 amino acid residues (Table II).

Mehl, Pacovska, and Winzler (19) conclude with Rising and Yang (20) that in the biuret reaction approximately 1 atom of copper is bound for

minutes, 50 μl. of Reagent D
ızzing." After 10 minutes or
: added with *immediate* "buz-
ıutes. Standards are perhaps
0, 20, etc., mg. per cent solu-
cent trichloroacetic acid, with
les.

o the final color with protein:
(*b*) reduction of the phos-
: copper-treated protein.
—The salient features of this
the *absence* of copper is prob-
l tryptophan content (16, 17),
ɔretreatment ((4–6) and Table
treatment of proteins results
ntrast, the presence of copper
ith free tyrosine and tryp-
) The reaction with copper,
te in 5 or 10 minutes at room
Heating to 100° or increas-
reaction with copper without
: with alkali alone does not
alkaline solution. Even pre-
)H, or for 5 minutes at 100°
copper treatment, has almost
:ating with strong alkali will,

l the biuret reaction appear
ɔnal, nor, with different pro-
ɔroportional to the increment
reagent (Table I).
ınt to give nearly maximum
appear to be catalytic. As-
rotein ⇌ copper-protein ob-
s may be utilized to calculate
$^{-6}$ mole per liter with a maxi-
ıd copper per 7 or 8 amino

: with Rising and Yang (20)
ıtom of copper is bound for

TABLE I

Extinction Coefficients of Proteins Variously Treated

The equivalent extinction coefficient ϵ^{N}_{λ} (or 550) is defined as the optical density at λ = 750 (or 550) mμ with 1 atom of N per liter. Nitrogen was measured by the Kjeldahl procedure of Miller and Houghton (24). The biuret color was developed with the reagents of Weichselbaum (25). Source of proteins, crystalline trypsin, crystalline chymotrypsin, and crystalline bovine albumin, Armour and Company, Chicago; cytochrome *c*, Sigma Chemical Company, St. Louis; crystalline zinc insulin, Eli Lilly and Company, Indianapolis; gelatin, Difco Laboratories, Inc., Detroit; L-tyrosine, Eastman Kodak Company, Rochester.

Protein	ϵ^{N}_{750}				
	Copper absent		Alkaline† copper treatment	Increment with Cu	Biuret color
	No pretreatment	Alkali* treatment			
Trypsin	733	910	3,600	2690	26.3
Insulin	989	998	3,000	2002	24.4
Chymotrypsin	278	425	2,930	2505	25.8
Cytochrome *c*	703	738	2,495	1757	‡
Human serum	320	365	2,120	1755	21.4
Bovine serum albumin	312	358	2,050	1692	21.8
Gelatin	79	78	1,145	1067	18.0
Tyrosine	13,700	13,850	15,100	1250	

* 30 minutes in 0.1 N NaOH at room temperature before addition of Folin reagent.

† Regular treatment as described under the procedure.

‡ Not valid because of the color of the cytochrome *c*.

TABLE II

Color Increments from Small Amounts of Copper

Serum protein concentration, 12.1 γ per ml. *K* and chromogenic-bound copper (Cu-protein) calculated from *K* = Cu × protein/Cu-protein = (Cu(total protein − Cu-protein))/Cu-protein = (Cu(maximum ΔO. D.* − ΔO. D.))/ΔO. D. (assuming that chromogenic copper bound to protein is proportioned to ΔO. D.).

Total Cu	O. D.* at 750 mμ	ΔO. D.	ΔO. D., per cent of maximum Δ	*K*	Cu-protein† (calculated)
10^{-6} *M*				× 10^{-6}	
0	78				
8	166	88	42	2.9	0.05
20	237	159	76	3.0	0.10
40	267	189	91	2.8	0.12
2000	286	208	100		0.13

* Optical density.

† Moles per 117 gm. of protein, *i.e.* per amino acid residue.

each 4 amino acid residues, and Mehl *et al.* calculated dissociation constants for the reaction copper + protein ⇌ copper-protein averaging 10 times larger than the one reported herein for the formation of Folin-reactive material. Thus, of the total possible sites for copper combination, it would appear that only about half produce significant reduction of the Folin reagent, and that furthermore this fraction has a considerably greater affinity for copper than the rest.

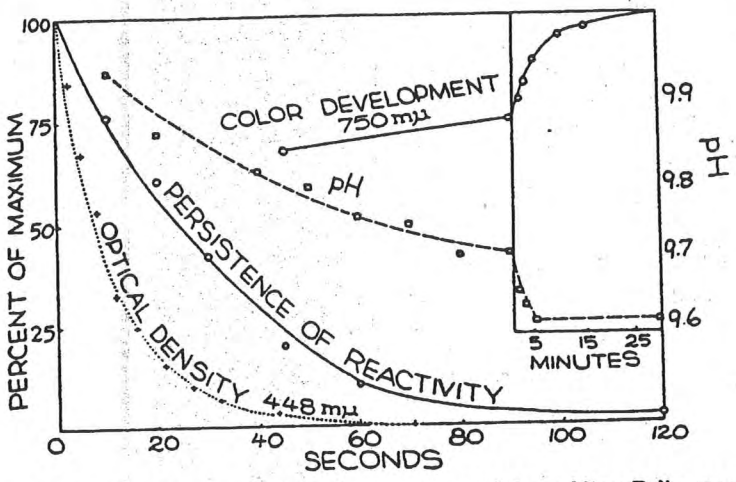


FIG. 1. "Persistence of reactivity" was measured by adding Folin reagent to protein-free alkali; after the given times, copper-treated protein was added in a small volume, and the color at 750 mμ was measured 30 minutes later. "Color development" refers to a sample of serum protein treated in the regular manner. The points on "optical density 448 mμ" are observed (no protein present); the curve is theoretical for a monomolecular reaction with a half time of 8 seconds.

Reduction of Folin Reagent—Three main points may be made as follows: (1) When the Folin reagent is added to the copper-treated protein, maximum color results if the reduction occurs at about pH 10. (2) At this pH the reagent is only reactive for a short time (16). It is for this reason that even a few seconds delay in complete mixing will lessen the amount of color (Fig. 1). The decrease in reactivity of the reagent appears to be a function of the disappearance of the original yellow color of the phosphomolybdate (half time of 8 seconds (Fig. 1)) and is presumably due to dissociation of the phosphate from the molybdate. Surprisingly, the color with protein continues to develop for a number of minutes after the reagent itself has become unreactive to freshly added protein (Fig. 1). Possibly the primary reduction product rearranges, since the absorption spectrum changes in shape between 3 minutes and 30 minutes (Fig. 2). (3)

During the 1st
acid is liberate
the phosphomc

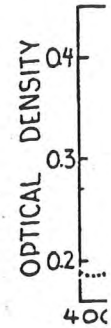


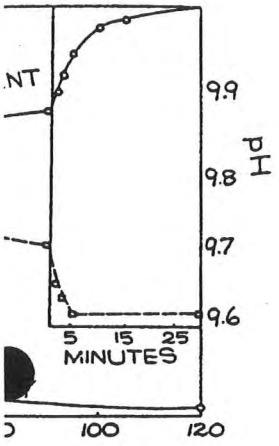
FIG. 2. Absorp
to a solution con

FIG. 3. Effect
centration is calc
cept as noted, fi
cent. All sample
few seconds befor
tration 12 γ per m

be rather well b
to neutralize th
near pH 10, gi
(Fig. 3).

Extinction Co

. calculated dissociation con-
copper-protein averaging 10
· the formation of Folin-reac-
ites for copper combination,
e significant reduction of the
ion has a considerably greater



ed by adding Folin reagent to
reated protein was added in a
1 30 minutes later. "Color de-
treated in the regular manner.
rved (no protein present); the
vith a half time of 8 seconds.

nts may be made as follows:
opper-treated protein, maxi-
out pH 10. (2) At this pH
(16). It is for this reason
xing will lessen the amount
of the reagent appears to be
al yellow color of the phos-
) and is presumably due to
ate. Surprisingly, the color
ter of minutes after the re-
lded protein (Fig. 1). Pos-
, since the absorption spec-
1 30 minutes (Fig. 2). (3)

During the 1st minute or so after the addition of the Folin reagent, extra acid is liberated (Fig. 1), which also may result from the dissociation of the phosphomolybdate. Therefore, for maximum color, the solution must

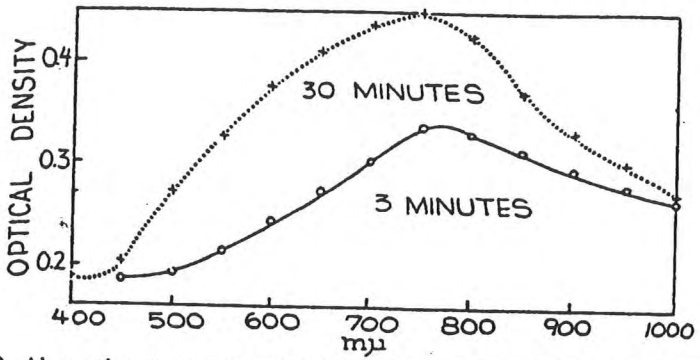


FIG. 2. Absorption spectra 3 and 30 minutes after the addition of Folin reagent to a solution containing 23.3 γ of serum protein per ml.

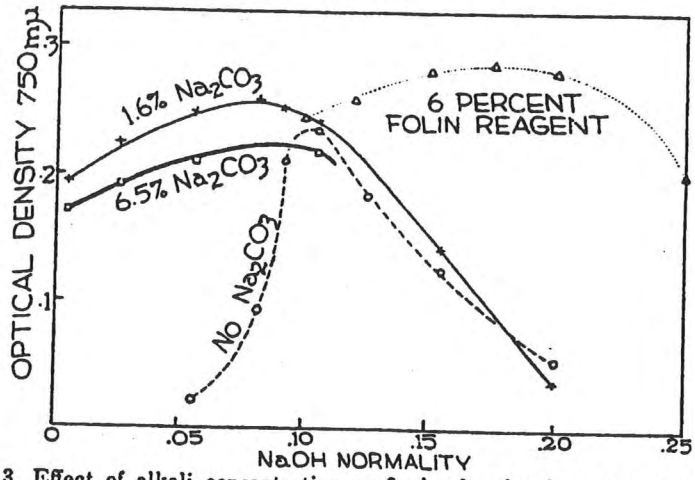


FIG. 3. Effect of alkali concentration on final color development. NaOH concentration is calculated before addition of 0.1 volume of diluted Folin reagent. Except as noted, final concentration of Folin reagent 3 per cent and Na_2CO_3 1.6 per cent. All samples (copper-treated protein) were identical in composition until a few seconds before addition of Folin reagent (see the text). Final protein concentration 12 γ per ml.

be rather well buffered. It was found that a mixture of NaOH, sufficient to neutralize the excess phosphoric acid, and Na_2CO_3, to buffer the mixture near pH 10, gives more color than any amount of either reagent alone (Fig. 3).

Extinction Coefficients and Proportionality—Different pure proteins give

different extinction coefficients with the Folin reagent (Table I). The extremes were observed with trypsin and gelatin which differed by a factor of 3 in chromogenicity. It will be seen that without copper much greater differences occur. The variation in chromogenicity must be kept in mind, but it is much less marked with mixtures of proteins as found in various tissues (Table III), and for many purposes is not a serious drawback.

The relation of color to protein concentration is not quite linear (Table IV).

TABLE III

Apparent Protein Content of Whole Tissues (Rabbit) and Tissue Extracts Calculated from Kjeldahl N and from Folin Color

The tissues were homogenized and precipitated with 5 per cent trichloroacetic acid (TCA), and the lipides removed by successive extraction with 0.1 N potassium acetate in ethanol, ethanol, and isopropyl ether. (The purpose of the acetate is to neutralize the acid and prevent solution of some protein in the ethanol.) The N was determined as in Table I. The extinction coefficients were calculated from the N and color of the extracted precipitates.

Material analyzed		Extracted prot.	TCA extract	Lipide extracted	Whole tissue	N* 670 mμ
		per cent	*per cent*	*per cent*	*per cent*	
Brain	Based on N × 6.25	9.5	1.08	1.05	11.8	1960
	" " Folin color	9.5	0.15	0.28	9.8	[illegible]
Kidney	" " N × 6.25	13.9			15.9	1865
	" " Folin color	13.9	0.30	0.21	14.5	
Liver	" " N × 6.25	17.1			19.9	1875
	" " Folin color	17.1	0.49	0.28	18.0	
Skeletal muscle	" " N × 6.25	17.8			20.1	1942
	" " Folin color	17.8	0.15	0.09	18.8	
Heart	" " N × 6.25	13.0			15.3	1975
	" " Folin color	13.0	0.20	0.17	13.4†	

* See Table I.

† By summation; other values are direct determinations.

Specificity and Interfering Substances—Few substances encountered in biological work cause serious interference. Only a little color was obtained with either acid extracts or the lipides extracted from five different tissues (Table III). Consequently measurements on non-extracted whole tissue would be in error by only 3 to 6 per cent, whereas values based on N determination would be overestimated by 15 to 20 per cent.

Uric acid (16), guanine, and xanthine (21, 22) react with the Folin reagent. Guanine gives about 50 per cent more color than serum protein, weight for weight. The color is not enhanced by copper. Curiously, guanosine does not react appreciably. Hypoxanthine gives no color if purified (21). No more than a trace of color was obtained with adenine, ade-

Folin reagent (Table I). The
latin which differed by a factor
it without copper much greater
ogenicity must be kept in mind,
of proteins as found in various
s is not a serious drawback.
ation is not quite linear (Table

bit) and Tissue Extracts Calculated Folin Color

ed with 5 per cent trichloroacetic
sive extraction with 0.1 N potas-
l ether. (The purpose of the ace-
on of some protein in the ethanol.)
inction coefficients were calculated
es.

ted	TCA extract	Lipide extracted	Whole tissue	N*
ent	*per cent*	*per cent*	*per cent*	
.5	1.08	1.05	11.8	1960
.5	0.15	0.28	9.8	
.9			15.9	1865
.9	0.30	0.21	14.5	
.1			19.9	1875
.1	0.49	0.28	18.0	
.8			20.1	1942
.8	0.15	0.09	18.8	
.0			15.3	1975
.0	0.20	0.17	13.4†	

erminations.

ew substances encountered in

Only a little color was ob-
es extracted from five different
ements on non-extracted whole
ent, whereas values based on N
5 to 20 per cent.

1, 22) react with the Folin re-
nore color than serum protein,
ed by copper. Curiously, gua-
anthine gives no color if puri-
as obtained with adenine, ade-

nosine, cytosine, cytidine, uracil, thymine, or thymidine (see also Funk and Macallum (22)).

Neither color nor interference with protein color development was observed with the following substances at the given *final* concentrations: urea (0.5 per cent), guanidine (0.5 per cent), sodium tungstate (0.5 per cent), sodium sulfate (1 per cent), sodium nitrate (1 per cent), perchloric acid (0.5 per cent neutralized), trichloroacetic acid (0.5 per cent neutralized), ethyl alcohol (5 per cent), ether (5 per cent), acetone (0.5 per cent), zinc sulfate (0.1 per cent), barium hydroxide (0.1 per cent).

Most phenols, except nitrophenols, reduce the reagent (16); therefore thymol and to a lesser degree sulfosalicylic acid interfere, whereas picric acid up to 0.1 per cent is permissible. Glycine (0.5 per cent) decreases the color with protein by 50 per cent. Hydrazine over 0.5 mg. per cent increases the blank.

TABLE IV

Measurement of Small Amounts of Protein from Rabbit Brain

Final volume 0.082 ml.

Optical* density at 750 mμ	$E^{1\%}_{1\text{ cm.}}$ at 750 mμ	Protein		Optical* density at 750 mμ	$E^{1\%}_{1\text{ cm.}}$ at 750 mμ	Protein	
		Found	Present			Found	Present
		γ	γ			γ	γ
0.038		0.13	0.16	0.280	229	0.98	1.00
0.044		0.15	0.16	0.292	236	1.03	1.00
0.040		0.14	0.16	0.283	233	0.99	1.00
0.089	221	0.33	0.33	0.365	227	1.30	1.32
0.095	236	0.35	0.33	0.367	228	1.31	1.32
0.091	226	0.34	0.33	0.365	227	1.32	1.32
0.184	228	0.65	0.66	0.441	219	1.60	1.66
0.191	236	0.67	0.66	0.443	220	1.62	1.66
0.191	236	0.67	0.66	0.444	221	1.61	1.66

* Corrected for blank.

Ammonium sulfate greater than 0.15 per cent final concentration decreases color development. This is partly due to a decrease in alkalinity, and up to 0.25 per cent or so can be tolerated if an equivalent amount of extra alkali is added to the sample. Extra copper does not seem to help.

Microanalysis—With final volumes less than 0.1 ml., the amount of color is proportionately less than on the macro scale, especially if the reaction is carried out in wide tubes. Extensive testing did not definitely identify the cause of the decreased color. Neither oxygen, carbon dioxide, nor glass surface seemed to be involved. The critical step is the period of standing with alkali and copper.

The practical solution to this interesting difficulty seems to be to use slender tubes and to run standards under the same conditions. Table IV illustrates the reproducibility of protein measurements on small brain samples. Rabbit brain was homogenized and diluted 200- to 2000-fold. Aliquots of 3.6 μl. were analyzed for protein at a final volume of 0.082 ml. The amount of protein present was calculated from macro analyses. It is seen that the error is usually not over 0.02 γ.

DISCUSSION

The measurement of protein with copper and the Folin reagent has certain advantages. (1) It is as sensitive as with Nessler's reagent, yet requires no digestion. (2) It is 10 or 20 times more sensitive than measurement of the ultraviolet absorption at $\lambda = 280$ mμ and is much more specific and much less liable to disturbance by turbidities. (3) It is several fold more sensitive than the ninhydrin reaction (23) and is somewhat simpler, as well as much easier to adapt for small scale analyses. Free amino acids give much more color than proteins with the ninhydrin reaction, whereas the reverse is true with the Folin reagent. (4) It is 100 times more sensitive than the biuret reaction.

There are two major disadvantages of the Folin reaction. (*a*) The amount of color varies with different proteins. In this regard it is less constant than the biuret reaction, but more constant than the absorption at $\lambda = 280$ mμ. (*b*) The color is not strictly proportional to concentration.

From a consideration of the advantages and disadvantages, the reasonable applications of the copper-Folin reaction would seem to include (1) measurement of protein during enzyme fractionations, etc., (2) measurement of mixed tissue proteins, particularly when absolute values are not needed, (3) measurement of very small absolute amounts of protein, or highly diluted protein (*e.g.* spinal fluid) or protein mixed with colored substances or other nitrogen-containing substances, and (4) analyses of large numbers of similar protein samples, such as antigen-antibody precipitates.

SUMMARY

1. A study is presented of the measurement of proteins with the Folin phenol reagent after alkaline copper treatment. The basic reactions have certain peculiarities which need to be taken into consideration in using this reagent.

2. Directions are given for measurement of proteins in solution and proteins which have been precipitated with acid, etc. A micro procedure is also described for the measurement of as little as 0.2 γ of protein.

3. The differences in the amount of color obtained with a number of proteins is recorded. Interfering substances are listed.

sting difficulty seems to be to use
er the same conditions. Table IV
measurements on small brain sam-
nd diluted 200- to 2000-fold. Ali-
ein at a final volume of 0.082 ml.
lculated from macro analyses. It
0.02 γ.

ON

per and the Folin reagent has cer-
as with Nessler's reagent, yet re-
mes more sensitive than measure-
280 mμ and is much more specific
turbidities. (3) It is several fold
on (23) and is somewhat simpler,
nall scale analyses. Free amino
ins with the ninhydrin reaction,
lin reagent. (4) It is 100 times

the Folin reaction. (*a*) The
oteins. In this regard it is less
ore constant than the absorption
ly proportional to concentration.
s and disadvantages, the reason-
ction would seem to include (1)
ractionations, etc., (2) measure-
y when absolute values are not
bsolute amounts of protein, or
or protein mixed with colored
ubstances, and (4) analyses of
such as antigen-antibody pre-

nent of proteins with the Folin
ent. The basic reactions have
en into consideration in using

t of proteins in solution and
acid, etc. A micro procedure
little as 0.2 γ of protein.
or obtained with a number of
s are listed.

4. The advantages of simplicity and sensitivity of this reaction recommend it for a number of biochemical purposes.

BIBLIOGRAPHY

1. Wu, H., *J. Biol. Chem.*, **51**, 33 (1922).
2. Wu, H., and Ling, S. M., *Chinese J. Physiol.*, **1**, 161 (1927).
3. Greenberg, D. M., *J. Biol. Chem.*, **82**, 545 (1929).
4. Andersch, M., and Gibson, R. B., *J. Lab. and Clin. Med.*, **18**, 816 (1933).
5. Greenberg, D. M., and Mirolubova, T. N., *J. Lab. and Clin. Med.*, **21**, 431 (1936).
6. Minot, A. S., and Keller, M., *J. Lab. and Clin. Med.*, **21**, 743 (1936).
7. Pressman, D., *Ind. and Eng. Chem., Anal. Ed.*, **15**, 357 (1943).
8. Heidelberger, M., and MacPherson, C. F. C., *Science*, **97**, 405 (1943).
9. Kabat, E. A., and Mayer, M. M., Experimental immunochemistry, Springfield, 321 (1948).
10. Sutherland, E. W., Cori, C. F., Haynes, R., and Olsen, N. S., *J. Biol. Chem.*, **180**, 825 (1949).
11. Folin, O., and Ciocalteu, V., *J. Biol. Chem.*, **73**, 627 (1927).
12. Lowry, O. H., and Bessey, O. A., *J. Biol. Chem.*, **163**, 633 (1946).
13. Levy, M., *Compt.-rend. trav. Lab. Carlsberg, Série chim.*, **21**, 101 (1945).
14. Bessey, O. A., Lowry, O. H., and Brock, M. J., *J. Biol. Chem.*, **164**, 321 (1946).
15. Bessey, O. A., Lowry, O. H., Brock, M. J., and Lopez, J. A., *J. Biol. Chem.*, **166**, 177 (1946).
16. Folin, O., and Denis, W., *J. Biol. Chem.*, **12**, 239 (1912).
17. Herriott, R. M., *J. Gen. Physiol.*, **19**, 283 (1935).
18. Herriott, R. M., *Proc. Soc. Exp. Biol. and Med.*, **46**, 642 (1941).
19. Mehl, J. W., Pacovska, E., and Winzler, R. J., *J. Biol. Chem.*, **177**, 13 (1949).
20. Rising, M. M., and Yang, P. S., *J. Biol. Chem.*, **99**, 755 (1932–33).
21. Hitchings, G. H., *J. Biol. Chem.*, **139**, 843 (1941).
22. Funk, C., and Macallum, A. B., *Biochem. J.*, **7**, 356 (1913).
23. Kunkel, H. G., and Ward, S. M., *J. Biol. Chem.*, **182**, 597 (1950).
24. Miller, L., and Houghton, J. A., *J. Biol. Chem.*, **159**, 373 (1945).
25. Weichselbaum, T. E., *Am. J. Clin. Path.*, **16**, *Tech. Sect.*, **10**, 40 (1946).

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 262, No. 21, Issue of July 25, pp. 10035–10038, 1987
Printed in U.S.A.

Sequence from Picomole Quantities of Proteins Electroblotted onto Polyvinylidene Difluoride Membranes*

(Received for publication, January 13, 1987)

Paul Matsudaira

From the Whitehead Institute for Biomedical Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142

Small amounts (7–250 pmol) of myoglobin, β-lactoglobulin, and other proteins and peptides can be spotted or electroblotted onto polyvinylidene difluoride (PVDF) membranes, stained with Coomassie Blue, and sequenced directly. The membranes are not chemically activated or pretreated with Polybrene before usage. The average repetitive yields and initial coupling of proteins spotted or blotted into PVDF membranes ranged between 84–98% and 30–108% respectively, and were comparable with the yields measured for proteins spotted onto Polybrene-coated glass fiber discs. The results suggest that PVDF membranes are superior supports for sequence analysis of picomole quantities of proteins purified by gel electrophoresis.

Developments in protein sequencing (reviewed in Ref. 1) such as the incorporation of gas-phase methods (2), absorption of proteins onto Polybrene-coated supports (2, 3) or covalent attachment of proteins to derivatized glass (4), and on-line microbore HPLC[1] identification (1) of the PTH-derivatives have combined to permit routine sequence analysis from picomole quantities of protein. Although proteins can be isolated by electroelution from polyacrylamide gels or by HPLC, the yields are often extremely low, and the techniques involved are laborious. One improvement has been to electroblot proteins purified by gel electrophoresis directly onto glass fiber sheets (5, 6). While this approach simplifies sample handling, it is limited in its usefulness because the proteins are difficult to detect, the glass fiber filters must be derivatized, and the reagents specially purified. In addition, a recent study (7) has found that both the repetitive yield and initial yield for coupling from proteins electroblotted onto glass fiber sheets decrease with decreasing amounts of protein. As a result, small quantities of protein (<100 pmol) have still proved difficult to sequence.

Polyvinylidene difluoride (PVDF) membranes are mechanically strong solid phase supports that bind proteins hydrophobically and have been used in immunoblotting applications as a substitute for nitrocellulose filters (8). Because these membranes are inert to most solvents (acetonitrile, trifluoroacetic acid, ethyl acetate, trimethylamine, and hexane), they can be used in automated gas-phase sequenators. Here, I report that 7–250 pmol of protein can be spotted or electroblotted onto PVDF membranes, detected by staining with Coomassie Blue, and then sequenced directly. The initial coupling and repetitive yields are much higher than those reported for proteins electroblotted onto glass fiber filters. Since there is no need for preconditioning steps, elaborate derivatization of the membranes, or special purification of reagents, PVDF membranes are an ideal solid-phase support for sequence analysis, especially when used with electroblotting methods.

EXPERIMENTAL PROCEDURES

Materials and Reagents—PVDF membranes (Immobilon Transfer), 0.45-μm pore size, were obtained from Millipore. Reagents used in this study were reagent grade or electrophoresis grade and were not repurified. Sequencing-grade reagents, glass fiber discs, Polybrene, sperm whale myoglobin, and bovine β-lactoglobulin were purchased from Applied Biosystems. Villin was purified from chicken intestine epithelial cells and cleaved with trypsin as described previously (9).

SDS-Polyacrylamide Gel Electrophoresis—Samples (10–250 pmol) were loaded onto minigels containing a 7.5–20% gradient of polyacrylamide and electrophoresed according to Laemmli (10) at 20 watts constant power (20 min). The minigels (10 × 10 cm, 0.5 mm thick) were modified from the original description of Matsudaira and Burgess (11). After electrophoresis, the gels were soaked in transfer buffer (10 mM 3-[cyclohexylamino]-1-propanesulfonic acid, 10% methanol, pH 11.0) for 5 min to reduce the amount of Tris and glycine. During this time a PVDF membrane was rinsed with 100% methanol and stored in transfer buffer. The gel, sandwiched between a sheet of PVDF membrane and several sheets of blotting paper, was assembled into a blotting apparatus (Mighty Small, Hoeffer) and electroeluted for 10–30 min at 0.5 A in transfer buffer (12, 13). The PVDF membrane was washed in deionized H_2O for 5 min, stained with 0.1% Coomassie Blue R-250 in 50% methanol for 5 min, and then destained in 50% methanol, 10% acetic acid for 5–10 min at room temperature. The membrane was finally rinsed in deionized H_2O for 5–10 min, air dried, and stored at −20 °C.

Sequence Analysis—PVDF membranes were not precycled. Different amounts of protein (10–250 pmol) were spotted either on precycled Polybrene-coated trifluoroacetic acid-activated glass fiber filters or on small rectangles (2 × 4 mm) of PVDF membrane that were initially wetted with 100% acetonitrile. The membranes or filters were then air dried. Glass fiber filters were placed in the sequencing cartridge according to the manufacturer's instructions. Protein electroblotted onto PVDF membranes was stained with Coomassie Blue and the band was cut out with a clean razor. The membrane was centered on the Teflon seal and placed in the cartridge block of the sequenator. Proteins were sequenced on an Applied Biosystems model 470 sequenator equipped with on-line PTH analysis using the regular program O3RPTH. The PTH-derivatives were separated by reverse-phase HPLC over a Brownlee C-18 column (220 × 2.1 mm). The initial yield for the coupling step was calculated from the amount of PTH-derivatives present in the first cycle (corrected for 100% injection) and by the amount of protein spotted or electroblotted onto the membrane. The repetitive yield from myoglobin was determined from the peak heights of valine, leucine, and glutamic acid at positions 1, 2, 4, 11, 13, and 18. The repetitive yield from β-lactoglobulin was

* This work was supported by National Institutes of Health Grant AM35306, the Whitaker Foundation, and the PEW Scholars Program. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] The abbreviations used are HPLC, high performance liquid chromatography; PVDF, polyvinylidene difluoride; PTH, phenylthiohydantoin; SDS, sodium dodecyl sulfate.

calculated for leucine, isoleucine, and valine at residues 1, 2, 3, 10, 12, and 15. The repetitive yields for other proteins and peptides were calculated when possible from the recoveries of valine, alanine, or leucine residues.

RESULTS AND DISCUSSION

Proteins electroblotted onto PVDF membranes can be stained with Coomassie Blue, Amido Black, India Ink, or silver (8). As shown in Fig. 1, a 25-pmol digest of villin (2.4 μg) is easily detected by Coomassie Blue. The two major polypeptides, M_r 44,000 and 51,000, represent the NH_2- and COOH-terminal fragments of villin (9). Densitometry of Coomassie Blue-stained gels show both fragments are present in 5-pmol amounts. These fragments were cut from the PVDF membrane and subjected to 11 cycles of sequence analysis in a gas-phase sequenator. From the recovery of the PTH-derivatives in the first cycle, initial yields for coupling were 25–68%. Since complete transfer of the polypeptides to PVDF membranes was assumed, these values represent minimum estimate of the initial yields. The repetitive yields were 89–94%. The sequences agreed with those previously reported for the fragments.

Because preliminary experiments demonstrated that polypeptides electroblotted onto PVDF membranes can be readily sequenced, quantitative measurements of efficiency from these supports were made. Myoglobin and β-lactoglobulin were electroblotted onto PVDF membranes, and the recovery of PTH-derivatives was quantified. Fig. 2 shows the chromatograms of the PTH-derivatives released from 20 pmol of β-lactoglobulin for the first 15 cycles after electroblotting to PVDF membranes. As reported before (5, 6), Coomassie Blue does not interfere with the sequencing reactions nor produce artifact peaks on the HPLC traces. In the first 2 cycles, a variable amount of PTH-Gly was detected, but this amount is usually <5 pmol and reaches base-line levels after 2 cycles. Very little PTH-Gly is detected if the membranes are washed for 5–10 min in H_2O after transfer. The chromatograms display no DPU peak coeluting with PTH-Trp and the background levels of other PTH-derivatives remain low. Except for PTH-Trp, the recovery of PTH-derivatives was comparable to the amount obtained when Polybrene-coated glass fiber filters were used. Control experiments (not shown) indicate that PVDF membranes do not bind PTH-derivatives. When 100 pmol of PTH standards were applied directly to the membrane and the membrane washed, only 1–2 pmol of PTH-derivatives were detected. Very little PTH-Trp could be recovered from protein electroblotted onto the membranes. In contrast, PTH-Trp was detected in samples spotted onto the membrane. These facts indicate that tryptophan is destroyed either during electrophoresis or electroblotting.

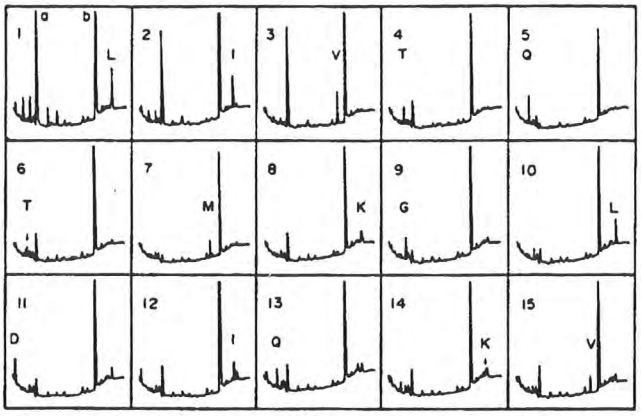


FIG. 2. **Chromatograms of PTH-derivatives, in *single letter abbreviation*, from 20 pmol of β-lactoglobulin electroblotted onto PVDF membranes and stained with Coomassie Blue.** The band was excised from the membrane and sequenced directly. The *abscissa* shows OD_{270}, the ordinate elution time. The peak height of leucine in cycle 1 corresponds to 6.7 pmol of injected material. Only 40% of the total PTH-derivatives is analyzed at every cycle. *Peak a* is *N'N*-dimethyl-*N'*-phenylthiourea; *peak b* is *N'N*-diphenylthiourea.

TABLE I

Comparison of the initial yield for coupling and repetitive yields for β-lactoglobulin and myoglobin from glass fiber and PVDF membranes

Initial and repetitive yields are calculated as described under "Experimental Procedures." The amount of protein listed in column 1 represents either the amount spotted onto the support or the calculated amount on the membrane after blotting (corrected for incomplete transfer).

	Support	Coupling yields	Repetitive yields
pmol		%	%
β-Lactoglobulin			
100	Polybrene-coated glass fiber	35	94.5
100	Spotted onto PVDF, unstained	108	90.7
200	Blotted onto PVDF, stained	75	93.0
50	Blotted onto PVDF, stained	100	93.3
20	Blotted onto PVDF, stained	84	92.8
Average		89	92.8
Myoglobin			
100	Polybrene-coated glass fiber	63	90.5
250	Spotted onto PVDF, unstained	88	87.8
100	Spotted onto PVDF, unstained	106	88.0
10	Spotted onto PVDF, unstained	86	88.0
66	Blotted onto PVDF, unstained	64	87.5
66	Blotted onto PVDF, stained	30	88.8
7	Blotted onto PVDF, stained	100	88.2
Average		79	88.1

95 ■ VELSKKVTGKL (68,94)

51 ■ VAKVEQVKFDA (29,90)
44 ■ VELSKKVTGKL (25,86)

FIG. 1. **A tryptic digest of villin (2.4 μg of total protein) electroblotted to PVDF membranes and stained with Coomassie Blue.** Densitometry of digest shows that the intact villin band contains 12.5 pmol of protein, while the M_r 44,000 and 51,000 bands each contain 5 pmol of material. The bands were cut out and placed in the gas-phase sequenator. Their NH_2-terminal sequences are displayed in *1-letter code*. The initial yield for coupling and the repetitive yield are shown in *parentheses*. The repetitive yields were calculated from the recoveries of PTH-Val, PTH-Ala, and PTH-Leu.

The amount of material that yielded sequence was calculated from the initial yield for coupling, and the efficiency of the sequencing reaction was measured by the repetitive yield. Both values provide quantitative standards for sequence analysis from glass fiber discs and PVDF membranes. Table I shows that the average initial yield for proteins spotted onto PVDF membranes is 97% and for proteins electroblotted onto PVDF membranes is 76%. The latter value is more variable because of difficulties in accurately estimating the amount of protein present on the membrane. More important, the initial yields remain high even when the smallest amount of protein

TABLE II

Proteins and peptides sequenced from PVDF membranes

Proteins and peptides submitted to the Whitehead Protein Sequencing Facility were blotted or spotted on PVDF membranes. The sequenceable amount of material was determined from the recovery or the PTH-derivatives from the first cycle. The repetitive yield was calculated from the recovery of PTH-leucine, valine, or alanine. In some cases the repetitive yield was not determined (ND).

	Sequenceable amount		
	No. of cycles	pmol	Repetitive yield
			%
Proteins blotted to PVDF			
Villin (95 kDa)	11	8	87
Intestine 72-kDa protein	39	9	92
Intestine 105-kDa protein	6	3	96
Anion channel (105 kDa)	9	3	96
Anion channel 68-kDa fragment	10	15	ND
Anion channel 10-kDa fragment	16	17	94
Villin peptides blotted onto PVDF membranes			
44T (44 kDa)	12	2	92
44T-28C (28 kDa)	5	16	ND
51T (51 kDa)	11	2	90
51T-40C (40 kDa)	5	3	ND
51T-29C (29 kDa)	12	30	84
51T-28T (28 kDa)	5	8	ND
51T-16C (16 kDa)	12	16	ND
51T-8C (8 kDa)	5	16	ND
Proteins spotted onto PVDF membranes			
mutS (95 kDa)	5	50	ND
Thioredoxin (12 kDa)	10	9	ND
Homoserine kinase peptide (32 kDa)	10	30	ND
Methanogen hydrogenase (subunit a)	10	1000	98
Methanogen hydrogenase (subunit b)	18	300	ND
Methanogen hydrogenase (subunit c)	10	15	ND
Villin peptides spotted onto PVDF membranes[a]			
13V15 CNBr	8	20	ND
13V18 CNBr	3	25	ND
13V19 CNBr	3	25	ND
13V8 CNBr	7	60	ND
13V10 CNBr	6	63	ND
Synthetic peptides spotted onto PVDF membranes[b]			
CDTAGQEDYDRLR	5	ND	ND
RPDFSLEPPYTGPCK	12	ND	ND
PREDVDY	7	ND	ND

[a] Peptides from reverse-phase HPLC were generated by cyanogen bromide cleavage of M_r 13,000 peptide from villin.

[b] The sequences of the number of cycles during which the recovery of PTH-derivatives is at least 20% of the amount recovered in the first cycle are shown. In all cases, peptides which could not be completely sequenced were reanalyzed after absorption onto Polybrene-coated glass fiber discs or Polybrene-coated PVDF membranes.

is sequenced. The values compare extremely favorably with the 15–25% initial yields (6–7) measured for proteins electroblotted onto glass fiber filters. Since the binding capacity of PVDF membranes (170 $\mu g/cm^2$) (8) is greater than that of Polybrene-coated glass fiber sheets (15–28 $\mu g/cm^2$) (5) or derivatized glass fiber sheets (7–25 $\mu g/cm^2$) (6, 7), the high initial yields measured here were expected.

The high initial yields also suggest that NH_2-terminal modification during electrophoresis was not a major problem as no precautions were taken to remove contaminating chemicals. The gels were prepared in batch and stored for several days (1–7) before use, and samples were solubilized before electrophoresis by boiling in SDS sample buffer. Scavengers like thioglycolate were not included in the electrophoresis.

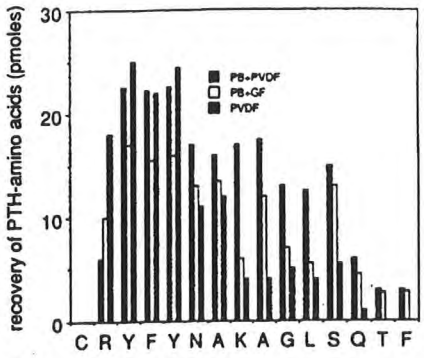


FIG. 3. **The recovery of PTH-derivatives is plotted from the analysis of peptide 8c which was spotted onto a PVDF membrane (*stippled bar*), PVDF membrane coated with Polybrene (*solid bar*), and glass fiber discs coated with Polybrene (*clear bar*).** Cysteine was not modified, and thus its recovery was not measureable. The amount of PTH-Ser was calculated from the recoveries of PTH-Ser and PTH-dehydro alanine.

The repetitive yields from proteins absorbed to glass fiber filters and PVDF membranes were similar, although the repetitive yield from PVDF membranes was 1–2% lower. The slightly reduced repetitive yield could result from a combination of incomplete cleavage, incomplete coupling, or peptide washout. The first two possibilities might be explained if the membrane interferes with the sequencing chemistry. However, the latter possibility is also likely since the hydrophobic binding of proteins and peptides to the membranes will be affected by the duration of the organic solvent washes in the Edman cycle. Because PVDF membranes have a 0.45-μm pore size and are thinner than a glass fiber disc, it is possible that the solvent wash steps in the sequencing programs are much longer than necessary for sequencing proteins absorbed to PVDF membranes. If true, then decreasing the duration of the solvent washes should improve the repetitive yields and shorten the cycle times.

The repetitive yields from myoglobin (90.5%) and β-lactoglobulin (94.5%) in the control experiments are lower than normal (usually 94 and 96%, respectively) and indicate that the sequenator was not operating at maximum efficiency. In addition to the test proteins reported here, these membranes have been used successfully to sequence 10- to 1000-pmol amounts of various proteins and peptides submitted to the Whitehead Facility. Table II shows that NH_2-terminal sequences can be obtained from other proteins (cytoplasmic and membrane) and polypeptides with a range of molecular weights. The repetitive yields ranged between 84–98%. In most cases repetitive yields were not calculated because the sequence lacked a repeated residue or the repetitive yield would be inaccurate due to poor recovery of some PTH-derivatives. The initial yields could not be measured since the band that was sequenced was from proteolytic digests composed of a mixture of peptides.

Short synthetic peptides were sequenced to examine whether peptide washout affects the recovery of PTH-derivatives. The results showed that complete analysis of short peptides can be problematic and depends on the hydrophobicity of the peptide. The recovery of PTH-derivatives from synthetic peptide 8c decreased by 80% after the Tyr-Phe-Tyr portion of the sequence when compared with the recovery from the same peptide spotted onto Polybrene-coated glass fiber discs or Polybrene-coated PVDF membranes (Fig. 3).

short peptides could be completely sequenced but the y of PTH-derivatives was variable and in poor yields. mmon feature of the analysis of peptides was a large lag ry over of each residue into the subsequent cycle. More le results were obtained by spotting the peptides on rene-coated PVDF membranes or Polybrene-coated fiber discs. As a general rule, peptides that are too short ctionate by polyacrylamide gel electrophoresis should be nced from Polybrene-coated supports. Proteins can be d in the presence of SDS and sequenced. Removal of before analysis is advisable to prevent blockage of the ery lines in the sequenator. The level of SDS is reduced sing the membrane with H_2O.

is paper demonstrates that proteins can be efficiently nced after electroblotting and staining on PVDF mem-es. There are several practical consequences that result sequencing on these membranes. First, sample handling nplified because the proteins can be stained after elec-otting. This eliminates the need to radiolabel or chemi-modify the protein prior to electrophoresis. Second, the precycling step employed for conditioning Polybrene-d glass fiber filters is eliminated, and thus the deadtime e sequencing machine is reduced. Third, since Coomassie can detect nanogram amounts of material and the re-ive yields and initial yield for coupling are not dependent rotein quantity, it should be possible to sequence femto-amounts of proteins efficiently. However, improved sensitivity in detecting the cleaved residues will be needed to sequence these amounts of protein.

Acknowledgments—I would like to thank M. Wallek and S. Blackmon for their technical assistance and Dr. P. Kim for stimulating discussions.

REFERENCES

1. Elizinga, M. (ed). (1982) *Methods in Protein Sequence* Humana Press, Clifton, NJ
2. Hewick, R. M., Hunkapiller, M. W., Hood, L. E. & Dreyer, W. J. (1981) *J. Biol. Chem.* **256,** 7990–7997
3. Tarr, G. E., Beecher, J. F., Bell, M. & McKean, D. J. (1978) *Anal. Biochem.* **84,** 622–627
4. Laursen, R. A. (1971) *Eur. J. Biochem.* **20,** 89–102
5. Vandekerckhove, J., Bauw, G., Puype, M., Van Damme, J. & Van Montagu, M. (1985) *Eur. J. Biochem.* **152,** 9–19
6. Abersold, R. H., Teplow, D. B., Hood, L. E. & Kent, S. B. H. (1986) *J. Biol. Chem.* **261,** 4229–4238
7. Yuen, S., Hunkapiller, M. W., Wilson, K. J. & Yuan, P. M. (1986) *Applied Biosystems User Bulletin #24,* Foster City, CA
8. Pluskal, M. F., Przekop, M. B., Kavonian, M. R., Vecoli, C. & Hicks, D. A. (1986) *BioTechniques* **4,** 272–282
9. Matsudaira, P., Jakes, R. & Walker, J. E. (1985) *Nature* **315,** 248–250
10. Laemmli, U. K. (1970) *Nature* **227,** 680–685
11. Matsudaira, P. T. & Burgess, D. R. (1978) *Anal. Biochem.* **87,** 386–396
12. Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76,** 4350–4354
13. Burnette, W. N. (1981) *Anal. Biochem.* **122,** 195–203

THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1990 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 265, No. 19, Issue of July 5, pp. 11397-11404, 1990
Printed in U.S.A.

β Protein C Is Not Glycosylated at Asparagine 329

THE RATE OF TRANSLATION MAY INFLUENCE THE FREQUENCY OF USAGE AT ASPARAGINE-X-CYSTEINE SITES*

(Received for publication, November 8, 1989)

Joseph P. Miletich‡§ and George J. Broze, Jr.¶

From the ‡Division of Laboratory Medicine, Departments of Medicine and Pathology, Washington University School of Medicine, St. Louis and the ¶Division of Hematology/Oncology, Department of Medicine, The Jewish Hospital of St. Louis, St. Louis, Missouri 63110

About 30% of human plasma protein C is smaller than the predominant form as judged by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. It has been suggested that this species, referred to as β protein C, is a degraded molecule. However, β protein C is secreted in culture by the HepG2 cell line and is present in plasma collected directly into numerous proteinase inhibitors; the percentage of β protein C does not change with time during culture or after blood collection. Neither thrombin, activated protein C, nor activated factor X converts the α form to β in the presence or absence of calcium and phospholipids. The NH_2-terminal sequences of the heavy chains of both forms are identical, and both release the same dodecapeptide and develop a functional active site when cleaved by thrombin. Both also react with antibodies to a synthetic COOH-terminal peptide. Timed digests with *N*-glycosidase are consistent with the interpretation that β protein C has three *N*-linked oligosaccharide chains whereas α protein C has four. It is asparagine 329 that is not glycosylated in β protein C since antibodies to a synthetic peptide based on the sequence around this amino acid react only with β protein C. This site is unique in having cysteine instead of serine or threonine 2 residues distal. It is likely that the sulfhydryl group can substitute for the usual hydroxyl group as a hydrogen bond acceptor for the glycosylation reaction only until it forms a disulfide bond. The percentage of protein C that is glycosylated at this site may therefore depend at least in part on the rate of disulfide bond formation which may in turn be related to the rate of protein synthesis.

Protein C is a 62,000-dalton plasma glycoprotein that circulates at an average concentration of 65 nM (1). It is the precursor to a serine proteinase known as activated protein C. Activation results from the thrombin-mediated cleavage of the Arg^{169}-Leu^{170} bond, releasing a 12-amino acid peptide from the two-chain molecule. The membrane cofactor thrombomodulin enhances the action of thrombin on protein C (2). Activated protein C proteolytically inactivates the procoagulant cofactors $VIII_a$ and V_a (3), which are necessary for the efficient formation of activated factor X and thrombin, respectively. Therefore, protein C is the central factor in a naturally occurring autoregulatory anticoagulant pathway.

The zymogen is synthesized in the liver as a 461-amino acid precursor protein that undergoes extensive co- and/or post-translational modification (4). The prepro leader sequence of 42 amino acids (Fig. 1) is removed by signal peptidase and a processing proteinase that cleaves the Arg-Ala bond between −1 and 1. The 9 glutamic acid residues between positions 6 and 29 in the numbering scheme of Fig. 1 are carboxylated in a vitamin K-dependent reaction. One or more determinants in the sequence from −1 to −24 probably constitute a recognition site for the carboxylase (5). Deficiency or antagonism of vitamin K (as with warfarin) blocks the modification of these glutamic residues and renders the synthesized protein C molecule incapable of assembling with high affinity into the calcium-mediated membrane complexes necessary for proper activation and function. Asp^{71} can be hydroxylated; this modification may influence calcium binding (6). The Lys^{156}-Arg^{157} dipeptide is excised by an unidentified proteinase producing a two-chain molecule; this processing step is incomplete, however, and 5–15% of plasma protein C is single chain (7).

Nearly 25% of the mass of protein C is carbohydrate. *O*-Linked glycosylation sites have not been identified, but there are four attachment sites at asparagines 97, 248, 313, and 329. The fourth site does not have the typical Asp-*X*-Ser/Thr sequence; a cysteine replaces the usual serine or threonine residue (8). About 30% of plasma protein C is smaller than the major form (by about 4 kDa, as assessed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis). The basis for the structural difference in this material, known as β protein C, has not been determined previously although it was thought possibly to be a degraded molecule (9). The human protein C gene has been mapped to a single locus on chromosome 2 (10). In this paper we show that β protein C is not derived from α protein C; it differs in not being glycosylated at asparagine 329.

EXPERIMENTAL PROCEDURES

Materials

Aprotinin, benzamidine, diisopropyl fluorophosphate (DFP),[1] soybean trypsin inhibitor, bovine serum albumin, ovalbumin, heparin-

* This work was supported in part by National Institutes of Health Grants HL 41963 and HL 14147 and by the Monsanto Chemical Corporation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: Box 8118, Division of Laboratory Medicine, Washington University School of Medicine, 660 S. Euclid Ave., St. Louis, MO 63110. Tel.: 314-362-1623; Fax: 314-362-1461.

[1] The abbreviations used are: DFP, diisopropyl fluorophosphate; dansyl, 5-dimethylaminonaphthalene-1-sulfonyl; GGACK, L-glutamyl-L-glycyl-L-arginine chloromethyl ketone; PPACK, phenylalanyl-L-propyl-L-arginine chloromethyl ketone; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; MES, 2-(*N*-morpholino)ethanesulfonic acid.

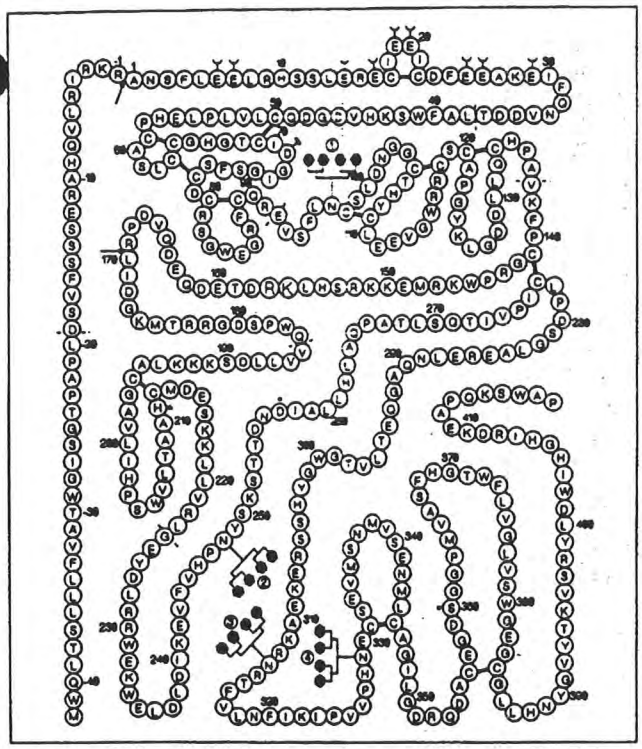


FIG. 1. **Two-dimensional representation of human protein C structure.** *Dashed lines* indicate exon boundaries, and *asterisks* denote the active site histidine, aspartic, and serine residues. See Introduction for modifications at other positions.

garose, bovine γ-globulin, affinity-purified rabbit anti-mouse IgG, rabbit brain phospholipids, 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide, bromoacetic acid *N*-hydroxysuccinimide ester, and acrylamide and bisacrylamide were from Sigma. Polystyrene latex particles (0.8–1.0-μm diameter) were purchased from Interfacial Dynamics Corporation. Fluorescein isiothiocyanate was obtained from Molecular Probes, Inc., and dansyl-L-glutamyl-L-glycyl-L-arginine chloromethyl ketone (dansyl-GGACK) and D-phenylalanyl-L-propyl-L-arginine chloromethyl ketone (PPACK) were from Calbiochem. Kits for detection of mouse or rabbit primary antibodies on Western blots using second antibody alkaline phosphatase conjugates were purchased from Promega Biotec. Immobilized recombinant protein A was purchased from Repligen. HepG2 hepatoma cells were obtained from the American Type Culture Collection. Molecular weight standards (unstained and prestained) and Affi-Gel 10 were from Bio-Rad. Sodium warfarin was from Du Pont. Vitamin K_1 (Mephyton) was from Merck Sharp and Dohme. Heparin (Liquaemin) was from Organon. IODO-GEN was from Pierce Chemical Co. Sodium [^{125}I]iodide, carrier free, was from Du Pont-New England Nuclear. Dulbecco's modified Eagle's medium, L-glutamine, penicillin/streptomycin, and fetal calf serum were obtained from KC Biological. *N*-Glycosidase F (*N*-glycanase) was obtained from Genzyme. All other chemicals were reagent grade or better and came from Sigma or Fisher. [^{3}H]DFP was obtained from Amersham Corp.

Methods

Protein C antigen was quantitated by a two-site particle concentration fluorescence immunoassay using monoclonal antibodies hC-2 and hC-4 (1).

Peptides were synthesized on an Applied Biosystems 430A automated peptide synthesizer using *t*-butoxycarbonyl chemistry. Standard cycles and conditions were employed. Amino-terminal sequence analyses of zymogen and activated heavy chains of α and β protein C were performed using approximately 50 pmol of each on a gas-phase sequenator (Applied Biosystems), and the phenylthiohydantoin derivatives were identified by high performance liquid chromatography.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed on 12% gels (acrylamide to bisacrylamide ratio of 37.5:1) by the method of Laemmli (11) using 4% stacking gels.

Monoclonal antibody hC-2 was labeled using IODO-GEN to a specific radioactivity of 1,000–2,000 cpm/ng.

Proteins—Protein C was isolated from 10 liters of fresh frozen plasma purchased from the St. Louis Bi-State Chapter of the American Red Cross. A barium citrate precipitate and eluate were prepared as described previously (12) and were dialyzed against 100 mM NaCl, 1 mM EDTA, 10 mM benzamidine, 1 mM DFP, and 50 mM Tris-HCl, pH 7.5. After clarification by centrifugation, this solution (≈110 ml) was applied (at 4 °C) at a flow rate of 20 ml/h to a 2.5 × 5-cm column of Affi-Gel 10 to which had been coupled 125 mg of hC-2 monoclonal antibody. The column was washed with 5 volumes of the buffer used for dialysis, and then the protein was eluted with 2 M sodium thiocyanate, pH 7.0. Fractions containing more than 100 μg/ml protein were pooled immediately, dialyzed against 100 mM NaCl and 50 mM Tris-HCl at pH 7.5, and concentrated to ≈2 ml using an Amicon concentrator with a PM-10 membrane. Small amounts of high molecular weight contaminants were removed by applying 1.0-ml aliquots to two Pharmacia LKB Biotechnology Inc. Superose 12 columns run in tandem on a Pharmacia fast protein liquid chromatography system at a flow rate of 0.4 ml/min. Under these conditions, single-chain protein C elutes slightly faster than two-chain protein C, and α protein C elutes earlier than β protein C. However, since all four species overlap substantially, total protein C was pooled for the experiments described in this paper.

Bovine protein C was prepared from a barium sulfate eluate of bovine plasma purchased from Sigma by a modification of the method described by Stenflo (13). DEAE-Sepharose was used instead of DEAE-Sephadex A-50, and contaminants were further removed by pooling and dialyzing fractions containing protein C against 50 mM NaCl and 20 mM MES, pH 6.0, and passing the solution through a heparin-agarose column in the same buffer. The flow-through was made 50 mM in Tris base and adjusted to pH 7.5, applied to a Mono Q fast protein liquid chromatography column (Pharmacia), and eluted with a NaCl gradient. The final material was >98% pure as judged by 12% SDS-PAGE with and without reduction. When activated by thrombin, this material prolonged the activated partial thromboplastin time of bovine plasma.

Plasmas—Human plasmas were obtained from volunteers using the two-syringe technique to draw 9 volumes of blood into 1 volume of a freshly prepared solution containing 10 mM DFP, 1,000 units/ml aprotinin, 10 units/ml heparin, 100 μg/ml soybean trypsin inhibitor, 100 mM benzamidine, 10 μM PPACK, and 130 mM sodium citrate, pH 6.5. Cells were removed by immediate centrifugation at 12,000 × *g* for 5 min, and the supernatant plasma was collected. Barium-adsorbed plasma was prepared by adding 0.08 volume of 1.0 M $BaCl_2$, incubating for 1 h at 4 °C, and centrifuging at 12,000 × *g* for 8 min at 4 °C.

HepG2 Cells—Conditioned medium and cell lysates were prepared by culturing HepG2 cells to confluence in T-75 flasks with Dulbecco's modified Eagle's medium plus 4 mM L-glutamine, 10% fetal calf serum, and 1 μg/ml vitamin K_1. The medium was removed, and adherent cells were gently washed twice with 15 ml of medium without calf serum or vitamin K_1. The medium was then replaced with Dulbecco's modified Eagle's medium and 4 mM L-glutamine with either 1 μg/ml vitamin K_1 or 20 μg/ml sodium warfarin and the cells incubated overnight. The cells in each flask were gently washed twice the next morning with 15 ml of the appropriate medium and incubated in 5 ml of the same medium for an additional 24 h. The medium from each flask was collected, and 0.10 volume of an inhibitor mixture (100 mM benzamidine, 1 mg/ml leupeptin, 400 μg/ml soybean trypsin inhibitor, 20 mM DFP, 50 mM EDTA, and 20 mM Tris-HCl, pH 8.1) was added; debris was removed by centrifugation. Adherent cells were washed twice with 100 mM NaCl, 5 mM EDTA, and 50 mM Tris-HCl, pH 7.5, and then lysed (in 5 ml of 100 mM NaCl, 10 mM EDTA, 40 μg/ml soybean trypsin inhibition, 10 mM benzamidine, 100 μg/ml leupeptin, 2 mM DFP, 2% (v/v) Triton X-100, and 50 mM Tris-HCl, pH 7.5) by gentle rotation for 1 h at room temperature. This solution was then clarified by centrifugation at 15,000 × *g* for 10 min.

Antibodies—The murine monoclonal antibodies hC-2 and hC-4 have been described (1); hC-2 is specific for an epitope near the activation site of human protein C (Arg^{169}-Leu^{170}); it does not bind to activated protein C or to the activation peptide. The antibody hC-4 recognizes an epitope on the light chain and binds with equal affinity to zymogen or activated protein C. The monoclonal antibody hC-6 was obtained by immunizing BALB/c mice with a synthetic peptide (Lys-Ile-Pro-Val-Val-Pro-His-Asn-Glu-Cys-Ser-Glu-Val-Met-Ser) coupled to keyhole limpet hemocyanin using bromoacetic acid *N*-hydroxysuccinimide ester as described by Bernatowicz and Matsueda (14). Standard methods for fusing splenic white cells to Sp 2/0-Ag 14 mouse myeloma cells with polyethylene glycol and isolating hybri-

domas with selective growth media were followed.

Screening for antipeptide antibodies was accomplished by incubating 20 μl of conditioned medium with 20 μl of 20 nM ovalbumin that had been labeled with fluorescein at a ratio of 5.8:1 before coupling to the synthetic peptide (ratio ≈ 5:1) using bromoacetic acid *N*-hydroxysuccinimide ester. After a 1-h incubation, 20 μl of carboxyl polystyrene latex (5 mg/ml), to which 100 μg/ml affinity-purified rabbit anti-mouse IgG had been coupled using 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide, was added for 15 min. The particles were collected and washed by filtration and front-surface fluorescence quantitated using a Pandex Screen Machine. Positive lines were subcloned twice by limiting dilution and subjected to at least one freeze/thaw cycle. The cell line hC-6 was selected for further study on the basis of an estimated affinity constant of 2×10^8 M (15).

Rabbits were immunized with this same peptide-keyhole limpet hemocyanin conjugate or with a COOH-terminal peptide (Cys-Gly-His-Ile-Arg-Asp-Lys-Glu-Ala-Pro-Glu-Lys-Ser-Trp-Ala-Pro)-keyhole limpet hemocyanin conjugate. In each case, the initial immunization was 1 mg of conjugate in complete Freund's adjuvant given intradermally at approximately 20 sites (50 μl each). Booster injections of 200 μg of the same conjugate in incomplete adjuvant were given subcutaneously at 4–6-week intervals.

Western Blots—Plasma samples for Western blots were prepared by initial purification and concentration of protein C by absorption and elution from monoclonal antibody hC-2 coupled to Affi-Gel 10. A volume of plasma calculated to contain 300–350 pg of protein C was added to 25 μl of hC-2-Affi-Gel (packed volume) in a final volume of 1 ml (with 100 mM NaCl, 50 mM Tris-HCl, pH 7.5) and mixed for 1 h at room temperature. The beads were collected by centrifugation, washed with 1 ml of 100 mM NaCl, 50 mM Tris-HCl, pH 7.5, and the bound protein C was then eluted by adding 25 μl of a twice concentrated solution of the sample buffer used for SDS-PAGE and heating for 3 min at 65 °C. Following centrifugation, 20 μl of the supernatant was removed for analysis; 2 μl of 10% (v/v) β-mercaptoethanol was then added to reduce disulfide bonds. Samples from conditioned medium and cell lysates were prepared in a similar manner except the absorption step was extended overnight at 4 °C since the total volume was 5 ml.

Electrophoretic transfer to nitrocellulose was performed according to Towbin *et al.* (16) at 60 V for 90 min with constant cooling to 10 °C. The nitrocellulose was then incubated at room temperature with gentle orbital agitation in the following solutions (all containing 100 mM NaCl, 50 mM Tris, pH 7.5): 10 mg/ml bovine serum albumin and 1 mg/ml bovine γ-globulin for 1 h; 0.5 μg/ml ^{125}I-hC-2 monoclonal antibody and 1 mg/ml bovine serum albumin overnight; the buffer alone three times for 15 min. The paper was dried with cool air and placed in a film cassette fitted with Cronex (Du Pont) intensifying screens with XAR-5 film (Eastman Kodak) for 6–18 h at −75 °C. The bands were quantitated using an LKB scanning laser densitometer. In control experiments it was established that the percentage of β protein C detected did not vary when the amount of a control plasma subjected to all of these procedures was varied from 50 to 150 μl.

When purified protein C was analyzed directly using other monoclonal or polyclonal antibodies and alkaline phosphatase-conjugated second antibodies, the protocol suggested by the supplier for antigen detection was followed.

Treatment with N-Glycosidase F—Human or bovine protein C was treated using the reagents and conditions recommended by the supplier. Briefly, the protein was denatured at 100 °C in 0.5% SDS and 100 mM β-mercaptoethanol and brought to a final concentration of 0.67 mg/ml by the successive addition of the following (final concentrations indicated): 200 mM sodium phosphate, pH 8.6; 10 mM 1,10-phenanthroline hydrate (in methanol); 1.25% Nonidet P-40; and 10–60 units/ml *N*-glycosidase F. Samples were incubated at 37 °C for varying times; digests were stopped by addition of 2 volumes of twice concentrated SDS sample buffer containing 5% (v/v) β-mercaptoethanol and analyzed by SDS-PAGE.

RESULTS

β Protein C Is Not a Degradation Product—A typical preparation of immunoaffinity-purified human protein C analyzed by 12% SDS-PAGE under reducing conditions is shown in Fig. 2, *lane A*. Roughly 30% of the material is β protein and can be seen in both the single-chain protein C and the heavy chains of two-chain protein C. Blood drawn into a mixture of multiple inhibitors (see "Methods") and prepared immediately for Western blotting has both α and β protein C present (see Fig. 6). In other experiments not shown, serum or plasma prepared from whole blood that was incubated at room temperature for several days after collection showed no change in the percentage of β protein C although small amounts of even faster migrating material could be detected.

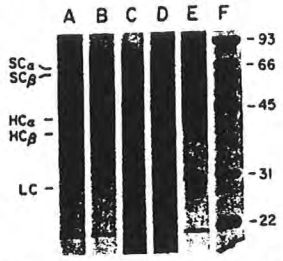


FIG. 2. **Purified human protein C analyzed by SDS-PAGE under reducing conditions.** Gels were stained with Coomassie Brilliant Blue R. *Lane A*, total protein C; *lanes B–E*, polypeptides obtained by electroelution of bands cut from gels similar to *lane A* were re-run to check for purity; *lane F*, molecular weight markers. The approximate amount of protein in each lane was *A*, 4 μg; *B*, 0.8 μg; *C*, *D*, and *E*, 2 μg. SC_α and SC_β, α and β single-chain protein C; HC_α and HC_β, heavy chain of two-chain α and β protein C; *LC*, light chain.

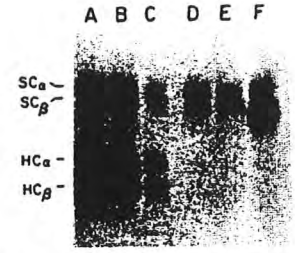


FIG. 3. **Protein C made by HepG2 cells.** Conditioned medium or cell lysate from cultured cells was analyzed for protein C by Western blotting from reduced SDS-polyacrylamide gels and probing with labeled monoclonal antibody hC-2. *Lanes A* and *B*, medium from cells grown with vitamin K; *lane C*, medium from cells grown in the presence of warfarin; *lanes D* and *E*, lysate from cells grown with vitamin K; *lane F*, lysate from cells grown in the presence of warfarin. Abbreviations are as in Fig. 2.

Human thrombin, human factor X_a, and activated protein C itself at concentrations as high as 1 mg/ml do not convert any α protein C to β protein C in the presence of either EDTA or 5 mM $CaCl_2$ and rabbit brain phospholipid (data not shown). Autoproteolysis of human factor X_a to a smaller species sometimes referred to as β factor X_a was clearly evident in the presence of calcium and phospholipid.

HepG2 Cells Secrete Protein C in Culture—When conditioned medium is analyzed by Western blotting, both α and β protein C (single-chain and two-chain) are present (Fig. 3, *lanes A* and *B*). The percentage of β protein C made by the cells in culture is higher than the percentage found in plasma; in various experiments, β protein C ranged from 45 to 55% of the total protein C. These cells also secrete a higher fraction of protein C as single chain. When the intracellular material from these cultures was analyzed (Fig. 3, *lanes D* and *E*), it was not possible to distinguish α and β protein C clearly since a spectrum of molecules of different sizes was present. Presumably, this represents various stages of processing. The range of molecules found is consistent with intracellular protein C being exclusively single-chain material. Excision of the Lys^{156}-Arg^{157} dipeptide must be a late processing event, possibly coupled to secretion.

In other experiments not shown, purified radiolabeled protein C was added to HepG2 cells in culture. At times ranging

from 30 min to 18 h, medium was removed and analyzed by SDS-PAGE and autoradiography. The percentage of β protein C did not change, demonstrating that β protein C does not derive from extracellular proteolysis of α protein C.

The effect of warfarin on HepG2 cells is also shown in Fig. 3, *lanes C* and *F*. The amount of secreted protein C is dramatically decreased. Even though the cells were washed, incubated overnight, and then washed again before beginning the 24-h timed collection of conditioned medium in the presence of warfarin, it cannot be determined from this experiment whether or not the small amount of secreted protein C detected represents synthesis in the presence of warfarin or runoff of protein translated prior to the addition of warfarin. Intracellularly, single-chain material of lower apparent molecular weight than either mature α or β protein C accumulates. This probably represents incompletely processed protein C and suggests that either the vitamin K-dependent carboxylation of the NH_2-terminal glutamic acid residues is a relatively late event or is required prior to some other relatively late processing step.

Both α and β Protein C Are Activated by Thrombin and Have Catalytically Intact Active Sites—Both α and β protein C are specifically cleaved by thrombin (in the absence of calcium), releasing a dodecapeptide from the NH_2-terminal end of the heavy chains (see below). Protein C activated by thrombin was incubated with either dansyl-GGACK or [^{3}H]DFP and then analyzed by SDS-PAGE. The gels were examined under short ultraviolet light (dansyl-GGACK) or by autofluorography ([^{3}H]DFP), and in both cases the bands corresponding to the heavy chains of both α and β-activated protein C were detected. Therefore, the catalytic site must be intact for both.

The Heavy Chains of α and β Protein C Have the Same NH_2-terminal Sequence—Attempts to separate α and β protein C which were only partially or not at all successful include (*a*) size exclusion chromatography; (*b*) ion exchange chromatography; (*c*) lectin affinity chromatography; (*d*) preparative isoelectric focusing; and finally (*e*) reverse phase chromatography following reduction and alkylation. Failure in many instances was due to the fact that protein C is quite heterogeneous apart from the difference between the α and β forms. For example, at least three distinct light chains, apparently differing by as much as 2 kDa as determined by SDS-PAGE, and similarly, doublets separated by ≈1 kDa can be seen for both the α and β heavy chains. The nature of these differences is not known. Isolated α and β protein C heavy chains were therefore obtained by electroelution of bands excised from SDS-polyacrylamide gels run under reducing conditions (Fig. 2, *lanes C* and *D*). The NH_2-terminal sequences were the same for both: Asp-Thr-Glu-Asp-Gln-Glu-Asp-Gln-Val-Asp-Pro-Arg-Leu-Ile-Asp-Gly-Lys-Met-Thr-Arg-Arg-Gly-Asp-Ser-Pro-Trp.

Following complete activation by thrombin, the heavy chains of both forms were obtained in a similar manner, sequenced and again found to be identical: Leu-Ile-Asp-Gly-Lys-Met-Thr-Arg-Arg-Gly-Asp-Ser-Trp-Gln-Val-Val-Leu-Leu-Asp.

Antibodies Obtained by Immunization with a COOH-terminal Synthetic Peptide Bind to Both α and β Protein C—A synthetic peptide corresponding to residues 405–419 was coupled to keyhole limpet hemocyanin through an NH_2-terminal cysteine and used to immunize a rabbit. Human protein C was transferred to nitrocellulose following SDS-PAGE under reducing conditions and incubated with the antiserum. As illustrated in Fig. 4, *lane A*, following incubation with an alkaline phosphatase-conjugated second antibody and color

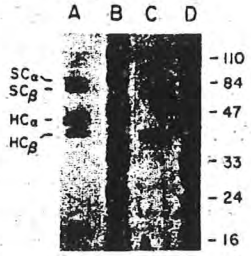


FIG. 4. **Western blots of human protein C.** Approximately 1 μg of protein C (*lanes A* and *C*) or prestained molecular weight standards (*lanes B* and *D*) was subjected to 12% SDS-PAGE and transferred to nitrocellulose. *Lanes A* and *B* were incubated with rabbit antibody against a synthetic COOH-terminal peptide; *lanes C* and *D* were incubated with the murine monoclonal antibody hC-6 "Exraised against a synthetic peptide centered around Asn^{329} (see "Methods" for details). After washing, they were incubated with anti-rabbit and anti-mouse IgG alkaline phosphatase conjugates, respectively, and the color was developed. Abbreviations are as in Fig. 2.

development, it is apparent that the α and β forms (whether single-chain or two-chain) were both detected. As expected, activation by thrombin does not affect reactivity of either species with this antisera (data not shown). It is possible that β protein C is missing just a few amino acids from the COOH terminus and can therefore still react with antibodies that recognize more proximal epitope(s) corresponding to the NH_2-terminal end of the synthetic peptide. However, this would not explain the large apparent molecular weight difference between α and β protein C.

Timed Digests with N-Glycosidase F Are Consistent with One Less Oligosaccharide Side Chain on β Protein C—Since the structural difference between α and β human protein C could not be attributed to a significant difference in the polypeptide sequence, attention was focused on the carbohydrate content. There is direct evidence for glycosylation at all three typical potential *N*-linked sites in bovine protein C (17, 18) as well as at a unique fourth site (Asn-Ala-Cys) corresponding to Asn^{329} of the human protein. All four analogous sites are present in human protein C. To determine whether all four sites of the human protein are actually glycosylated, purified protein C was incubated with different concentrations of *N*-glycosidase F for times ranging from 15 min to days. When the digests were analyzed by SDS-PAGE, the patterns were consistent with removal of one carbohydrate side chain from the light chain and three from the α heavy chain. Moreover, loss of the most labile oligosaccharide from the α heavy chain produces a band that comigrates with the β heavy chain. It was not possible, however, to define a treatment with *N*-glycosidase F which quantitatively removed all of the *N*-linked carbohydrate from human protein C; more extensive digests than those shown in Fig. 5 degraded the protein to an extent that precluded straightforward interpretation. Therefore, even though the pattern is highly suggestive that β protein C is a form of the protein which is missing the single (and most accessible) *N*-linked carbohydrate side chain, it remained formally possible that α protein C has some other structural modification that causes a nearly identical apparent molecular weight difference.

If this were the case, then α and β human protein C might both have only two *N*-linked side chains attached to the heavy chain region. This seemed unlikely for two reasons. First, the apparent molecular weight differences between the heavy chain bands that appear following treatment with *N*-glycosidase F should then all be nearly the same, but close inspection reveals that they are not. The apparent molecular mass difference between the α (≈40 kDa) and β (≈36 kDa) bands is greater (≈4kDa) than the ≈1.5-kDa difference between the β

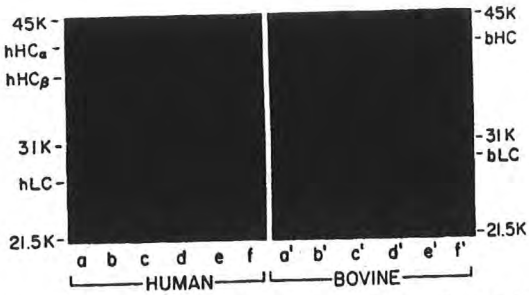


FIG. 5. **Cleavage of *N*-linked carbohydrate.** Human (*h, lanes a–f*) or bovine (*b, lanes a'–f'*) protein C were digested with *N*-glycosidase F (see "Methods") at 37 °C and analyzed by 12% SDS-PAGE with Coomassie Brilliant Blue R staining. *Lanes a, f, a', f'*, 2.5 μg of each protein handled in the same manner as digests but with no *N*-glycosidase F added; *lanes b* and *b'*, 10 μg of protein incubated for 30 min with 10 units/ml enzyme; *lanes c* and *c'*, 10 μg of protein incubated for 1 h with 30 units/ml enzyme; *lanes d* and *d'*, 10 μg of protein incubated for 4 h with 60 units/ml enzyme; and *lanes e* and *e'*, 10 μg of protein incubated for 53 h with 60 units/ml enzyme. Abbreviations are as in Fig. 2.

band and the next band lower (≈34.5 kDa), which is in turn different from the 2-kDa difference between it and the lowest band (≈32.5 kDa). This lowest band is still ≈3 kDa larger than the molecular mass calculated on the basis of the 262 amino acids in the heavy chain (29.4 kDa), suggesting within the limitations of this type of analysis that other modifications may also occur.

Second, the human and bovine heavy chain regions are extremely homologous. Since purified bovine protein C is more homogeneous (it is all α), the pattern of digestion with *N*-glycosidase, although still incomplete, is straightforward in interpretation and clearly shows that cleavage of the three *N*-linked side chains from the heavy chain region produces bands that migrate in the same positions as the bands seen following the digests of human protein C. Therefore, it seemed very likely that β human protein C is synthesized without one of the three *N*-linked carbohydrate side chains usually attached to asparagines 248, 313, and 329.

It is interesting that the bovine light chain (≈28 kDa) has an apparent molecular mass that is ≈2 kDa greater than human protein C light chain (≈26 kDa) although both run ≈4 kDa lower after digestion with *N*-glycosidase F. Since they are very homologous 155-amino acid polypeptides with a calculated molecular mass of 17.8 kDa, this suggests that some other substantive modification(s) may be made to the light chains and that this is somehow different for the human and bovine proteins.

Antibodies Developed by Immunization with a Synthetic Peptide Centered around Asn^{329} *Bind to β but Not to α Protein C*—Since Asn^{248} and Asn^{313} are typical glycosylation sites, Asn^{329} was investigated in detail. Rabbits and mice were immunized with a peptide corresponding to residues 322–336 of human protein C. Purified human protein C was transferred to nitrocellulose following SDS-PAGE under reducing conditions and incubated with either the rabbit antisera or the highest affinity murine monoclonal antibody (hC-6) that reacted with the synthetic peptide and with native human protein C and survived long term culture (see "Methods"). In both cases only β protein C was detected. This is illustrated in Fig. 4, *lane C*, for the monoclonal antibody. These results demonstrate that α human protein C is modified in the sequence from residues 322 to 336 in a manner that prevents recognition by antibodies reactive with that polypeptide sequence. However, β human protein C reacts well, indicating no significant structural difference from the synthetic peptide. Taken together with the results of the digests with *N*-glycosidase F, it seems reasonable to conclude that α human protein C is glycosylated at Asn^{329} but that β human protein C is not. Since there is only one gene for protein C and since α protein C is not converted to β protein C, the unanswered question is why this glycosylation site is normally used only 70% of the time.

The Percentage of Human Protein C Produced without Glycosylation at Asn^{329} *May Depend on the Overall Rate of Protein Synthesis*—To determine whether individuals differ in the efficiency of glycosylation at Asn^{329}, the fraction of β protein C in various plasmas was estimated. The method involved immunoabsorbtion, elution, and separation by SDS-PAGE under reducing conditions, transfer to nitrocellulose, incubation with [^{125}I]iodine-labeled hC-2 antibody, autoradiography, and laser densitometry (see "Methods"). By this technique, 20 normal fasting donors (11 male and 9 female, taking no medications) between the ages of 20 and 40 have a constant amount of β protein C, 31 ± 2%. This fraction must not be dictated by the primary protein structure *per se* since hepG2 cells in culture produce about 50% β protein C. Further, the highly homologous bovine protein is synthesized as with all α protein C.

Not all subjects have 30% β protein C. Plasma from a patient (M. R.) with a 40-year history of thromboembolic disease (numerous deep vein thromboses and several pulmonary emboli diagnosed both clinically and by angiography) was also analyzed. On three separate occasions between events, while receiving either no treatment or prophylactic subcutaneous heparin, this patient's total protein C was only 19% (0.8 mg/ml) of normal for her age (1). As shown in Fig. 6, less than 5% (detection limit) of her protein C is the β form. Note that the volume of plasma analyzed was adjusted to keep the total amount of detected antigen roughly constant. Seven of the patient's eight relatives (T. R., D. C., J. R., J. L., K. C., A. W., and L. R.) who were available for testing had total protein C antigen levels consistent with a heterozygous defect (51–69% of normal for their ages); the percentage of their protein C that was β ranged from 16 to 21%. A niece (K. S.) who was also tested had 104% antigen with 29% β protein C. In this kindred with an inherited deficiency of protein C, lower than normal concentrations of total protein C antigen were associated with an increased fraction of antigen having all four *N*-linked carbohydrate side chains, suggesting a relationship to the rate of synthesis.

Since there are other possible explanations for a decreased

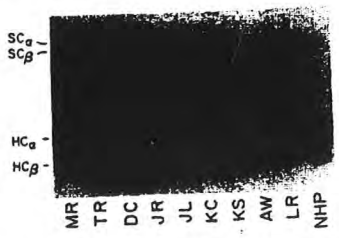


FIG. 6. **Western blots of normal human plasma (*NHP*) and plasmas from a family with protein C deficiency.** Plasmas were incubated with monoclonal antibody hC-2 coupled to Affi-Gel 10 for 1 h; the beads were washed, and the protein was eluted, separated by 12% SDS-PAGE, transferred to nitrocellulose, incubated with labeled hC-2 antibody and analyzed by autoradiography (see "Methods"). The amount of total protein C antigen in each *lane* was kept constant (≈120 ng) by adjusting the volumes of the plasmas. The total concentration of protein C antigen for each subject expressed as a percentage of normal is *M. R.* (propositus), 19%; *T. R.* (son), 58%; *D. C.* (daughter), 50%; *J. R.* (son), 57%; *J. L.* (daughter), 59%; *K. C.* (niece), 66%; *K. S.* (niece), 104%; *A. W.* (sister), 59%; and *L. R.* (brother), 69%. Other abbreviations are as in Fig. 2.

fraction of β protein C in a family with a genetic abnormality, plasma that had been collected from a volunteer (J. P. M.) following administration of large dosages of warfarin for experiments on factor VII (19) were analyzed. In this situation there was no abnormality in the gene for protein C, but warfarin does prevent the synthesis of normal amount of protein C (and other vitamin K-dependent factors). The total protein C (96%) and the percentage of β (30%) in the plasma of the volunteer were normal and stable for the 32 h before administration of 60 mg of warfarin. Immediately thereafter, the concentration of total protein C antigen fell as the synthesis of new molecules was inhibited. A second 30-mg dose of warfarin at 56 h ensured continued decline. As shown in Fig. 7, the *percentage* of β protein C decreased with warfarin administration. By 80 h, total protein C antigen was 35% of the starting concentration; β protein C was only about one-sixth of its initial concentration (less than 14% of the total antigen) at this time. The amount of plasma adsorbed was adjusted to keep the total amount of protein C analyzed approximately the same for each sample. This was done in case the transfer and detection steps were not absolutely linear over the wide range of β protein C which would have resulted from using a fixed volume of plasma; the percentage of β might have appeared to be even lower otherwise. The apparent change in migration of the protein C antigen under the influence of warfarin is most likely an artifact resulting from the increased amount of total protein that was loaded onto the gel for samples adsorbed from larger volumes of plasma. The concentration of total antigen for each time point and the estimated percentage of α and β protein C are shown in Fig. 8.

If this treatment had blocked the synthesis of new molecules completely, the clearance of α and β protein C would have to be different ($t_{1/2}$ ≈ 34 and 16 h, respectively). This was not the case, however. Normally, more than 95% of protein C made will adsorb to barium citrate and can therefore be extracted from citrated plasma by the addition of barium chloride and removal of the resultant precipitate by filtration or centrifugation. In the presence of warfarin, molecules are made which are undercarboxylated and will not bind to barium salts. The same plasmas were analyzed following barium citrate adsorption. As shown in Fig. 8, substantial amounts of nonadsorbable protein C were detected during warfarin treatment. At the 80-h time point, more than half of the total antigen present was apparently undercarboxylated and must represent new synthesis. Surprisingly, nearly all of this undercarboxylated protein C was the α form (fully glycosylated). β Protein C was always roughly 30% of the barium-adsorbable molecules. It is not obvious from this experiment whether or not some barium-adsorbable molecules (either α or β) were also made during warfarin treatment, so estimates of the half-lives of the adsorbable α and β may underestimate their real clearance rates. Nonetheless, it is interesting that the barium-adsorbable α and β proteins did disappear at approximately the same rate ($t_{1/2}$ ≈ 15 h). This experiment demonstrates unequivocally that the fraction of protein C made without carbohydrate at Asn^{329} can vary *in vivo* since the nonadsorbable protein C had to represent new synthesis and was fully glycosylated. The rate of synthesis of new molecules is lower in the presence of warfarin, suggesting that this may influence the likelihood of glycosylation. It is also possible that undercarboxylated molecules are structurally (rather than temporally) more accessible for glycosylation at Asn^{329}.

As the concentration of total protein C returned to normal in response to 20 mg of phytonadione (vitamin K_1) taken at 80 h followed by 5 mg every 8 h thereafter, the overall percentage of β protein C rapidly returned toward normal. A small amount of nonadsorbable protein C was still present at 144 h, probably an effect of residual warfarin not completely overcome by the vitamin K_1 therapy although a slower clearance of undercarboxylated molecules could also be a factor.

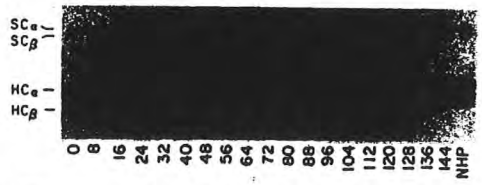


FIG. 7. **Western blots of plasmas from a volunteer taking warfarin.** At 32 and 56 h, 60 and 30 mg of warfarin was ingested; at 80 h, 20 mg of vitamin K was taken, followed by 5 mg every 8 h thereafter. The total amount of protein C analyzed (≈120 ng) was kept constant by varying the volume of plasma. See "Methods" and Fig. 6 for additional details. Abbreviations are as in Fig. 2.

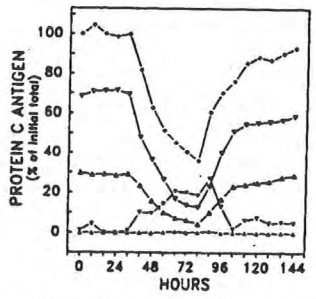


FIG. 8. **Effect of warfarin on the plasma concentrations of different forms of protein C.** The initial concentration of total antigen was 3.3 μg/ml (citrated plasma). *Solid diamonds*, total protein C; *open inverted triangles*, barium-adsorbable α protein C; *open triangles*, barium-adsorbable β protein C; *solid inverted triangles*, nonadsorbable α protein C; *solid triangles*, nonadsorbable β protein C. See "Methods" for details of analysis and the legend to Fig. 7 for times and dosages of warfarin and vitamin K.

DISCUSSION

Asparagine-linked glycosylation begins with the transfer of the precursor oligosaccharide $Glc_3Man_9GlcNAc_2$ to newly synthesized protein. The structure is then later modified by Golgi enzymes. The mature structure of *N*-linked oligosaccharides can vary considerably among different proteins, between different sites on the same protein, and to a lesser extent at the same site on a given protein (for reviews, see Refs. 20–22). Although this processing heterogeneity almost certainly occurs for human protein C, as evidenced by multiple bands on isoelectric focusing gels and minor bands on SDS-PAGE, it does not explain the large difference observed in the α and β forms. β protein C is the result of failure to transfer any oligosaccharide to the fourth potential site. Although direct proof of this hypothesis by carbohydrate and amino acid analysis of the appropriate peptide was not feasible because of limitations in material and lack of an efficient way to separate the α and β forms prior to digestion, the indirect evidence based on reactivity of the monoclonal antibody hC-6 and the pattern of digestion with *N*-glycosidase F is compelling.

There are some reported examples of proteins that are synthesized by a single cell type both with and without oligosaccharide at a given potential *N*-linked site. Castellino and co-workers (23–26) have shown that human plasminogen can be separated into two variant forms by affinity chromatography on lysine-Sepharose. The first form (type I) has both an *N*-linked and an *O*-linked oligosaccharide, but type II has only the *O*-linked sugar side chain. There is no difference in

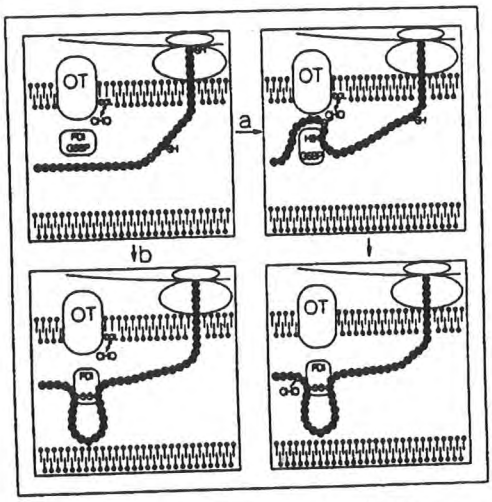


FIG. 9. **Model for glycosylation at Asn^{329}.** In *path a*, the free sulfhydryl of Cys^{331} is recognized by glycosylation site binding protein (*GSBP*), and carbohydrate (*CHO*) is transferred to Asn^{329} before protein disulfide isomerase (*PDI*) forms a disulfide bond between Cys^{331} and Cys^{345}. In *path b*, protein disulfide isomerase catalyzes the disulfide bond formation first; without a free sulfhydryl at Cys^{331}, carbohydrate is not attached to Asn^{329}. *OT*, oligosaccharyl transferase.

the amino acid sequence in these variants. These authors postulate that the rate of synthesis of plasminogen may be so rapid as to preclude *N*-linked glycosylation all the time, but no data to support this hypothesis were presented. There is nothing obviously unusual about the sequence at this site (Asn-Arg-Thr). A second example is tissue plasminogen activator (27) purified from a human melanoma cell line. It is of interest that the variants in this case were also separated by binding to lysine-Sepharose. The type I tissue plasminogen activator is reported to be glycosylated at three *N*-linked glycosylation sites (Asn-Ser-Ser, Asn-Gly-Ser, Asn-Arg-Thr), but type II is only glycosylated at the first and third sites.

Human protein C is more unusual in that the glycosylation site in question has the sequence Asn-*X*-Cys instead of the typical Asn-*X*-Ser/Thr. We know of only two other examples of glycosylation at a similar site. The first is the highly homologous bovine protein C molecule as noted above; the second is von Willebrand factor (28). There are eight occurrences of the sequence Asn-*X*-Cys in von Willebrand factor; in two places, *X* is proline, which might preclude recognition (29). In the six other possible sites, *X* is arginine, threonine, serine, histidine, aspartic acid, and glutamic acid, with only the Asn-Ser-Cys site glycosylated. The other sites may simply not be accessible for glycosylation. However, since seven of eight Asn-*X*-Thr and four of four Asn-*X*-Ser (excluding an Asn-Pro-Ser) sequences in the same protein are glycosylated, it is tempting to speculate that the cysteines play a role in the likelihood of glycosylation. Bause and Legler (30) have proposed that the hydroxyl group of the threonine or serine residue normally found at *N*-linked glycosylation sites is required for recognition by glycosylation site binding protein, functioning as a hydrogen bond acceptor in a proton relay system of glycosylation. These authors used synthetic peptides to demonstrate that the free sulfhydryl group of cysteine can substitute for the hydroxyl group of theonine or serine but at a lower relative activity in their model system. It is possible that glycosylation at Asn^{329} in protein C (or at other Asn-*X*-Cys sites) is intrinsically simply less favorable than at the Asn-*X*-Ser/Thr sites and is therefore the most likely site skipped when protein is being made rapidly. However, it is also of interest that Cys^{331} eventually forms a disulfide bond with Cys^{345}. If the hydrogen bond acceptor theory is correct, then Asn^{329} should not be a potential glycosylation site after the disulfide bond between Cys^{331} and Cys^{345} has formed. The length of time that Asn^{329} is in the lumen of the endoplasmic reticulum and available for glycosylation before Cys^{345} is also in the lumen and available for disulfide bond formation will be longer if protein is being made less rapidly. This theory is more intriguing given that glycosylation site binding protein and protein disulfide isomerase are reportedly the same molecule or at least highly related molecules that might compete (31). Glycosylation at Asn^{329} would not prevent disulfide bond formation. This speculation is shown schematically in Fig. 9 where *path a* represents slower protein synthesis than *path b*. If this is true, then glycosylation should not be observed at Asn-*X*-Cys sites in other proteins where the cysteine is the distal partner of an intramolecular disulfide bond.

The significance of glycosylation at Asn^{329} is not known. Since most of human and all of bovine protein C normally have this modification, β protein C is the atypical molecule. The plasminogen and tissue plasminogen activator variants that lack carbohydrate at one *N*-linked site have not been reported to differ in any major way in functional properties. It has not been possible to determine if there is a functional difference between α and β protein C up to this point because there has not been a way to separate them under nondenaturing conditions. This should be possible by immunoaffinity chromatography with the monoclonal antibody hC-6. If further studies support the hypothesis that the percentage of protein C made without carbohydrate at Asn^{329} reflects the rate of synthesis, there may be clinical situations in which the ratio of β to α protein C would be useful. For example, a higher than normal β/α ratio with a normal or reduced level of total antigen might indicate increased synthesis in the face of rapid turnover whereas a lower than normal ratio might signal a production deficit.

Acknowledgments—We would like to thank Drs. Rosalind and Stuart Kornfeld for helpful discussions, Jane Huth for preparing this manuscript, and the Protein Chemistry Laboratory at Washington University for preparing synthetic peptides and for amino acid sequence analyses.

REFERENCES

1. Miletich, J. P., Sherman, L., and Broze, G. J., Jr. (1987) *N. Engl. J. Med.* **317**, 991–996
2. Esmon, N. L., Owens, W. G., and Esmon, C. T. (1982) *J. Biol. Chem.* **257**, 859–864
3. Marlar, R. A., Kleiss, A. J., and Griff, J. H. (1982) *Blood* **59**, 1067–1072
4. Foster, D. C., Yoshitake, S., and Davie, E. W. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 4673–4677
5. Foster, D. C., Rudinski, M. S., Schach, B. G., Berkner, K. L., Kumar, A. A., Hagen, F. S., Sprecher, C. A. Insley, M. Y., and Davie, E. W. (1987) *Biochemistry* **26**, 7003–7011
6. Öhlin, A.-K., Landes, G., Bourdon, P., Oppenheimer, C., Wydro, R., and Stenflo, J. (1988) *J. Biol. Chem.* **263**, 19240–19248
7. Miletich, J. P., Leykam, J. F., and Broze, G. J., Jr. (1983) *Blood* **62**, 306 (abstr.)
8. Foster, D., and Davie, E. W. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81**, 4766–4770
9. Kisiel, W. (1979) *J. Clin. Invest.* **64**, 761–769
10. Plutzka, J., Hoskins, J. A., Long, G. L., and Crabtree, G. R. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 546–550
11. Laemmli, U. K. (1970) *Nature* **227**, 680–685
12. Broze, G. J., Jr., and Miletich, J. P. (1984) *J. Clin. Invest.* **73**, 933–938
13. Stenflo, J. (1976) *J. Biol. Chem.* **251**, 355–363
14. Bernatowicz, M. S., and Matsueda, G. R. (1986) *Anal. Biochem.* **155**, 95–102
15. Miletich, J. P., and Broze, G. J., Jr. (1987) *Blood* **69**, 1580–1586
16. Towbin, H., Staehelin, T., and Gordan, J. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76**, 4350–4354

17. Fernlund, P., and Stenflo, J. (1982) *J. Biol. Chem.* **257,** 12170–12179
18. Stenflo, J., and Fernlund, P. (1982) *J. Biol. Chem.* **257,** 12180–12190
19. Broze, G. J., Jr., Hickman, S., and Miletich, J. P. (1985) *J. Clin. Invest.* **76,** 937–946
20. Hubbard, S. C., and Ivatt, R. J. (1981) *Annu. Rev. Biochem.* **50,** 555–583
21. Kornfeld, R., and Kornfeld, S. (1985) *Annu. Rev. Biochem.* **54,** 631–664
22. Rademacher, T. W., Parekh, R. B., and Dewk, R. A. (1988) *Annu. Rev. Biochem.* **57,** 785–838
23. Powell, J. R., and Castellino, F. J. (1983) *Biochemistry* **22,** 923–927
24. Hayes, M. L., and Castellino, F. J. (1979) *J. Biol. Chem.* **254,** 8768–8771
25. Hayes, M. L., and Castellino, F. J. (1979) *J. Biol. Chem.* **254,** 8772–8776
26. Hayes, M. L., and Castellino, F. J. (1979) *J. Biol. Chem.* **254,** 8777–8780
27. Bennett, W. F. (1983) *Thromb. Haemostasis* **50,** 106 (abstr.)
28. Titani, K., Kumar, S., Takio, K., Ericsson, L. H., Wade, R. D., Ashida, K., Walsh, K. A., Chopek, M. W., Sadler, J. E., and Fujikawa, K. (1986) *Biochemistry* **25,** 3171–3184
29. Bause, E. (1983) *Biochem. J.* **209,** 331–336
30. Bause, E., and Legler, G. (1981) *Biochem. J.* **195,** 639–644
31. Geetha-Habib, M., Noiva, R., Kaplan, H. A., and Lennarz, W. J. (1988) *Cell* **54,** 1053–1060

Proc. Natl. Acad. Sci. USA
Vol. 74, No. 1, pp. 134–138, January 1977
Biochemistry

Enzymatic conversion of proteins to glycoproteins

(lipid-linked saccharides/protein unfolding/oligosaccharide-lipid/Asn-X-Ser or Asn-X-Thr tripeptide/glycosyl transferase)

DOROTHY D. PLESS AND WILLIAM J. LENNARZ

Department of Physiological Chemistry, The Johns Hopkins University School of Medicine, 725 N. Wolfe Street, Baltimore, Maryland 21205

Communicated by Albert L. Lehninger, October 21, 1976

ABSTRACT **The enzymatic transfer of the oligosaccharide moiety from an oligosaccharide-lipid to denatured forms of three secretory proteins—ovalbumin, α-lactalbumin, and ribonuclease A—has been demonstrated utilizing a membrane fraction from hen oviduct. Based on a survey of 10 proteins denatured by sulfitolysis, the presence of the tripeptide sequence -Asn-X-$^{Thr}_{Ser}$- (X represents a variable amino acid) appears to be necessary but not sufficient for the protein to serve as acceptor *in vitro*. The results of this investigation also suggest that unfolding of the polypeptide chain is required in order to expose sites for carbohydrate attachment.**

Studies in this laboratory have established that membranes prepared from hen oviduct catalyze synthesis of an oligosaccharide-lipid, and transfer of its oligosaccharide moiety to endogenous, membrane-bound protein acceptors (1–3). The participation of similar lipid-linked saccharides in the synthesis of endogenous, membrane-bound glycoproteins has been demonstrated in a variety of other tissues (see ref. 4 for a review). Evidence that oligosaccharide-lipid might be involved in glycosylation of secretory proteins was provided by the findings of Eagon *et al.* (5) that the oligosaccharide moiety of oligosaccharide-lipid can be incorporated into an exogenous, carbohydrate-free form of a kappa-type immunoglobulin light chain. Aside from this study in the myeloma system, no other information has been reported on the possible participation of oligosaccharide-lipid as an oligosaccharide donor for the glycosylation of well-defined, exogenous proteins.

Ovalbumin has a single carbohydrate side chain similar or identical to that found in the oligosaccharide-lipid (3, 6, 7). However, using either oligosaccharide-lipid or GDP-mannose and a variety of incubation conditions, it has not been possible to demonstrate the glycosylation of either endogenous ovalbumin in the membrane preparation (1, 8) or exogenous, carbohydrate-depleted ovalbumin (D. D. Pless, D. Struck, and W. J. Lennarz, unpublished data). To test the possibility that a site for glycosylation might be inaccessible in the native protein, ovalbumin was subjected to denaturation and sulfitolysis. Addition of this modified protein to the membrane preparation in the presence of labeled oligosaccharide-lipid resulted in formation of a new labeled polypeptide with an apparent molecular weight similar to that of ovalbumin (9). Studies of this labeled protein (K. Kronquist and W. J. Lennarz, unpublished data) are complicated by the fact that less than 20% of the amino acid sequence of ovalbumin is known (10, 11). For this reason it was of interest to identify other exogenous protein acceptors of known amino acid sequence in order to determine which structural features of the protein might correlate with its ability to serve as substrate. We have found that, in addition to ovalbumin, denatured forms of bovine ribonuclease A (RNase A; EC 3.1.4.22) and bovine α-lactalbumin serve as acceptors of the oligosaccharide chain of oligosaccharide-lipid.

Abbreviations: RNase, ribonuclease; DNase, deoxyribonuclease; Man, mannose; $NaDodSO_4$, sodium dodecyl sulfate; LA, lactalbumin.

MATERIALS AND METHODS

Materials. Bovine pancreas RNase A (RASE grade) and deoxyribonuclease (DNase; EC 3.1.4.5; DP grade) were purchased from Worthington Biochemical Corp. Carboxypeptidase A (EC 3.4.12.1; bovine pancreas), elastase (EC 3.4.21.11; porcine pancreas), and alcohol dehydrogenase (EC 1.1.1.1; horse liver) were purchased from Boehringer Mannheim, and papain (EC 3.4.22.2; papaya) was obtained from ICN Pharmaceuticals, Inc. α-Lactalbumin (Grade II, bovine milk), lysozyme (EC 3.2.1.17; Grade 1, chicken egg white), ovalbumin (Grade V, chicken egg albumin), and RNase B (type XII-B, bovine pancreas) were all products of the Sigma Chemical Corp. RNase B contained about equal amounts of RNase A and RNase B. RNase A was quantitatively removed by elution through a column of concanavalin-A-Sepharose; RNase B, which bound to the column, was eluted with α-methylmannoside. Sodium [^{35}S]sulfite (42.5 mCi/mmol) and GDP-[^{14}C]mannose (221 mCi/mmol) were obtained from New England Nuclear Co.

General Methods. The oviduct membrane preparation and [*Man*-^{14}C] oligosaccharide-lipid were prepared as previously described (8). Protein was determined by the method of Lowry *et al.* (12) or by absorbance measurement at 280 nm. Extinction values $E_{1\,cm}^{1\%}$ used were: 7.36 for ovalbumin (13), 20.5 for α-lactalbumin (14), and 6.98 for RNase A (15).

Assay for Synthesis of [*Man*-^{14}C]Oligosaccharide-Protein. Standard reaction mixtures contained membrane enzyme (190 μg of protein), [*Man*-^{14}C]oligosaccharide-lipid (1 to 2 × 10^4 cpm) in 5 μl of 0.2% sodium deoxycholate, 10 mM Tris-maleate, pH 7.24, and 15 mM $MnCl_2$ in a total volume of 50 μl. Where indicated exogenous protein (100–250 μg) was added. After incubation for 2 hr at 37° aliquots (40–100%) of the total incubation mixture were analyzed by disc gel electrophoresis in the presence of sodium dodecyl sulfate ($NaDodSO_4$), urea and mercaptoethanol on 12.5% polyacrylamide gels (16). After electrophoresis, the gels were stained for protein or for carbohydrate (17) and then sliced and measured for radioactivity (16). The staining–destaining procedure removed any low molecular-weight, water-soluble radioactive compounds.

Sulfitolysis Procedure. Proteins were subjected to the sulfitolysis procedure described by Pechère *et al.* (18). Reagents were removed by a series of dialyses (19) or by gel filtration on Bio-Gel P-4. The sulfitolyzed proteins were concentrated by lyophilization and dialyzed against 1 mM Tris·HCl, pH 7.2. Completeness of the reaction was tested using $Na_2{}^{35}SO_3$ (18). The number of moles of $^{35}SO_3$ found to be incorporated per mole of protein were as follows: RNase A, 7.74 [known number of half-cystine/mole = 8.0 (10)]; α-lactalbumin, 8.43 [half-cystine/mole = 8.0 (10)]; and ovalbumin, 8.22 [half-cystine/mole = 5 to 8 (20)].

RESULTS

Rationale for sulfitolysis

Since glycosylation of native ovalbumin could not be detect

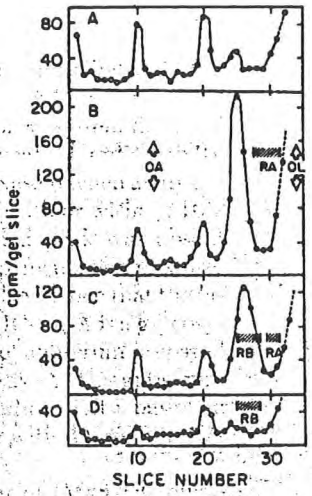


FIG. 1. Comparison of [*Man*-^{14}C]oligosaccharide-polypeptides synthesized in the absence of exogenous protein or in the presence of sulfitolyzed RNase. The standard assay was used with 12,000 cpm [*Man*-^{14}C]oligosaccharide-lipid as substrate. Sulfitolyzed proteins tested as exogenous acceptors were (A) none; (B) sulfitolyzed RNase A (224 μg); (C) a mixture of sulfitolyzed RNase A and B (226 μg of protein); or, (D) sulfitolyzed RNase B (204 μg). An aliquot (30 μl) of each reaction mixture was analyzed by $NaDodSO_4$-polyacrylamide gel electrophoresis. The mobilities of ovalbumin (OA, endogenous in the enzyme preparation) and of [*Man*-^{14}C]oligosaccharide-lipid (OL) are indicated. Where added, the protein bands corresponding to RNase A (RA) and RNase B (RB) are shown.

under a variety of conditions, it seemed possible that potential acceptor proteins had to be unfolded to expose the carbohydrate acceptor site. Based on the finding of Tarentino *et al.* (19) that an endoglycosidase cleaved a portion of the oligosaccharide chain of ovalbumin only after the protein was subjected to sulfitolysis, this procedure was employed for denaturation and disulfide cleavage of a variety of proteins. Using $^{35}SO_3$, it was determined that the sulfitolysis reaction went to completion with RNase A, α-lactalbumin, and ovalbumin (see *Materials and Methods*).

Enzymatic conversion of RNase A to RNase B

Two forms of bovine RNase differ only by the absence (RNase A) or presence (RNase B) of a carbohydrate side chain (21) with the structure $(Man)_6(GlcNAc)_2$-linked to asparagine residue 34. RNase A and RNase B were sulfitolyzed and then tested as acceptors of the oligosaccharide chain of oligosaccharide-lipid. Analysis of the products by $NaDodSO_4$–polyacrylamide gel electrophoresis is shown in Fig. 1. In the absence of exogenous protein (Fig. 1A), two endogenous polypeptides with apparent molecular weights of 25,000 and 55,000 are labeled. In the presence of sulfitolyzed RNase A (Fig. 1B), incorporation of a relatively high level of [*Man*-^{14}C]oligosaccharide into a new polypeptide is observed. The major labeled polypeptide has a larger apparent molecular weight than RNase A and is distinct from the products derived from endogenous acceptors. An identical product is formed when a mixture (1:1) of sulfitolyzed RNase A and RNase B is added (Fig. 1C), instead of sulfitolyzed Rnase A only. As shown, the internal standards for the carbohydrate-free and the glycosylated form of ribonuclease are clearly separated and the new labeled polypeptide coincides in mobility with RNase B. As expected, pure sulfitolyzed RNase B does not serve as an acceptor (Fig. 1D), since it already contains an oligosaccharide chain at asparagine-34. The labeled product formed upon incubation of sulfitolyzed RNase A with enzyme and oligosaccharide-lipid also comigrates with RNase B on $NaDodSO_4$ gels containing 7.5%, instead of the usual 12.5% polyacrylamide.

To further characterize the major product labeled in the presence of sulfitolyzed RNase A, reaction mixtures were analyzed by a detergent-free gel electrophoresis system designed for basic proteins (22, 23). A distinct peak of radioactivity coincided with samples of sulfitolyzed RNase B when analyzed in the presence of 10 M urea and either with or without 10% (vol/vol) 2-mercaptoethanol. The labeled peak was absent under both conditions in samples of control incubation mixtures lacking sulfitolyzed RNase A. Additional evidence that the labeled product formed from sulfitolyzed RNase A is a glycosylated form of RNase was obtained by use of antiserum prepared from rabbits injected with sulfitolyzed RNase A. This antiserum precipitated over 75% of the expected radioactivity, based on the amount of radioactivity comigrating with RNase B in $NaDodSO_4$–polyacrylamide gels.

Conditions for the enzymatic transfer of [*Man*-^{14}C]oligosaccharide from [*Man*-^{14}C]oligosaccharide-lipid to sulfitolyzed RNase A were examined. The formation of labeled RNase is dependent on time and is linear for 50 min. The reaction is also dependent on the concentration of membrane enzyme (up to 200 μg), [*Man*-^{14}C]oligosaccharide-lipid (up to 30,000 cpm), and sulfitolyzed RNase A (up to 200 μg). With limiting concentrations of [*Man*-^{14}C]oligosaccharide-lipid (8000 cpm), up to 35% of the oligosaccharide moiety can be transferred to protein. The reaction is stimulated 10-fold by the addition of 15 mM $MnCl_2$, but only 2-fold by 15 mM $MgCl_2$, and the pH optimum is about 7 (data not shown).

One of the protein modifications resulting from the sulfitolysis treatment, namely the introduction of the negatively charged sulfite, is reversed by mercaptans (24). Accordingly, the effect of dithiothreitol on the acceptor activity and on the sulfite content of sulfitolyzed ribonuclease A was examined as shown in Fig. 2. First, it is clear that dithiothreitol *per se* does not inhibit the enzyme, since the addition of low concentrations (less then 10 mM) to standard incubations had little effect on the incorporation of [*Man*-^{14}C]oligosaccharide into the endogenous polypeptides. However, the presence of 2.5 mM dithiothreitol markedly inhibits labeling of sulfitolyzed RNase A. The lability of the linkage of sulfite to protein in the presence of dithiothreitol was confirmed by incubating [^{35}S]sulfitolyzed RNase A under identical conditions, followed by analysis of the amount of protein-bound ^{35}S. The results show that addition of increasing concentrations of dithiothreitol causes a parallel loss of both [^{35}S]sulfite and the acceptor activity of sulfitolyzed RNase A.

Enzymatic incorporation of oligosaccharide into α-lactalbumin

Native α-lactalbumin was found to be inactive as an acceptor (Fig. 3A). In contrast, a major, new radioactive product is formed in the presence of sulfitolyzed α-lactalbumin (Fig. 3B). The new labeled product has a mobility distinctly lower than that of unglycosylated α-lactalbumin, which is consistent with the increased molecular weight expected after incorporation of the labeled oligosaccharide.

As in the case of sulfitolyzed RNase A, dithiothreitol was found to inhibit markedly the glycosylation of sulfitolyzed α-lactalbumin, presumably because of the loss of sulfite groups. These results suggested that introduction of negative charges might be essential to obtain active acceptor protein. To test this, native α-lactalbumin was reduced in the presence of excess dithiothreitol, and then aliquots were alkylated with either iodoacetic acid or iodoacetamide. The results of assays com-

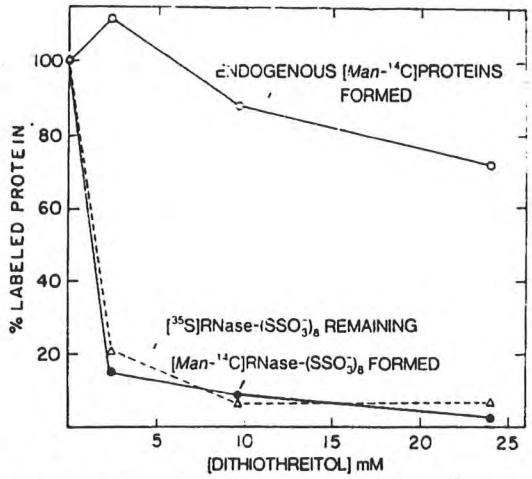


FIG. 2. The effect of mercaptan on acceptor activity and sulfite content of sulfitolyzed RNase A [RNase(SSO_3^-)$_8$]. Synthesis of [*Man*-^{14}C]oligosaccharide-protein was carried out in standard incubations with the addition as indicated of dithiothreitol (freshly prepared in deaerated water). NaDodSO$_4$–polyacrylamide gel electrophoresis was used to distinguish between labeled endogenous protein (○) and labeled sulfitolyzed RNase A (●). The effect of dithiothreitol on protein-bound $^{35}SO_3$ was determined in parallel incubations carried out under identical conditions except that [*Man*-^{14}C]oligosaccharide-lipid was omitted and [^{35}S]RNase-(SSO_3^-)$_8$ (259 μg, 1.36 × 10^5 cpm) replaced the unlabeled protein. After the incubation, carrier protein [110 μg of unlabeled RNase-(SSO_3^-)$_8$] was added. Labeled protein was precipitated by 0.5 ml of ice-cold 0.3 M $HClO_4$, collected, and washed on a Millipore filter, and radioactivity was determined (△).

paring the three forms of α-lactalbumin are presented in Table 1. Reduced α-lactalbumin that was alkylated with either the neutral or the negatively charged reagent was comparable to sulfitolyzed α-lactalbumin in acceptor activity. Reduced α-lactalbumin was relatively inactive, but it should be noted that possible reoxidation of the sulfhydryl groups of reduced α-lactalbumin during the incubation has not been excluded. The results presented in Table 1 also indicate that denaturation prior to reduction of disulfide bonds is not required in the case of α-lactalbumin, since reduction was carried out under nondenaturing conditions.

Survey of sulfitolyzed proteins as exogenous acceptors

A variety of sulfitolyzed proteins have been tested for acceptor activity, as summarized in Table 2. The sulfitolyzed proteins

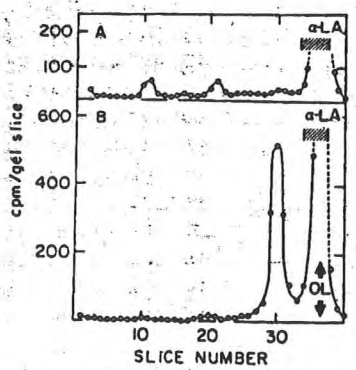


FIG. 3. Comparison of native and sulfitolyzed α-lactalbumin as acceptors. The standard assay was used with 1.2 × 10^4 cpm of [*Man*-^{14}C]oligosaccharide-lipid and a 150 min incubation. Either 160 μg of native α-lactalbumin (A) or 154 μg of sulfitolyzed α-lactalbumin (B) was present as exogenous acceptor. The mobilities of unglycosylated α-lactalbumin (α-LA) and of excess [*Man*-^{14}C]oligosaccharide-lipid (OL) are indicated.

Table 1. Comparison of modified α-lactalbumin derivatives as exogenous glycosyl acceptors

Exogenous protein	[*Man*-^{14}C]Oligosaccharide-α-lactalbumin (cpm)
None	<95
α-Lactalbumin-(SSO_3^-)$_8$	588
α-Lactalbumin -(SH)$_8$	237
α-Lactalbumin-(S-$CH_2CO_2^-$)$_8$	690
α-Lactalbumin-(S-CH_2CONH_2)$_8$	720

Sulfitolyzed α-lactalbumin [α-LA-(SSO_3^-)$_8$] was prepared as described in the *Materials and Methods*. α-Lactalbumin was reduced by incubating for 2.5 hr at 40° with a 50-fold excess (relative to disulfide) of dithiothreitol in 0.2 M Tris, pH 8.0, containing 2 mM EDTA (14). One sample of reduced α-lactalbumin [α-LA-(SH)$_8$] was dialyzed overnight against 2 mM dithiothreitol. To a second sample of reduced α-lactalbumin (10 mg), 60 mg of iodoacetamide was added, and the pH was readjusted to 8.0 with NH_3. After allowing 20 min at room temperature for reaction, the solution, containing carboxamidomethylated α-lactalbumin [α-LA-(SCH_2CONH_2)$_8$] was dialyzed overnight against water. Alkylation with iodoacetic acid to form carboxymethylated α-lactalbumin [α-LA-($SCH_2CO_2^-$)$_8$] was carried out in an analogous manner. α-LA-(SSO_3^-)$_8$ (164 μg), α-LA-(SH)$_8$ (125 μg), α-LA-(SCH_2COO^-)$_8$ (132 μg), and α-LA-(S-CH_2CONH_2)$_8$ (120 μg) were tested for acceptor activity by incubation with [*Man*-^{14}C]oligosaccharide-lipid under standard assay conditions. Aliquots of the reaction mixture were assayed for glycosylated products by NaDodSO$_4$–polyacrylamide gel electrophoresis.

active as exogenous acceptors—ovalbumin, RNase A, and α-lactalbumin—all contain at least one tripeptide of the form -Asn-X-$^{Thr}_{Ser}$-, in which X represents a variable amino acid. None of these proteins, when tested in its native form, served as an acceptor. Sulfitolyzed RNase B, in which the asparagine of this tripeptide sequence is already occupied by a carbohydrate chain, as well as sulfitolyzed lysozyme and sulfitolyzed papain, which lack this tripeptide, are inactive. No labeling of sulfitolyzed preparations of deoxyribonuclease (DNase), elastase,

Table 2. Survey of sulfitolyzed proteins as potential acceptors

Protein	Acceptor activity	Number of free -Asn-X-$^{Thr}_{Ser}$ sites*
Ovalbumin	+	>1
α-Lactalbumin	+	2
RNase A	+	1
RNase B	—	0
Papain	—	0
Lysozyme	—	0
Elastase	—	2
DNase	—	1
Carboxypeptidase A	—	1
Alcohol dehydrogenase	—	1

Proteins were sulfitolyzed as described in *Materials and Methods*. Incubations with oviduct enzyme and [*Man*-^{14}C]oligosaccharide-lipid were carried out under standard conditions and the products were analyzed by NaDodSO$_4$–polyacrylamide gel electrophoresis. Appearance of a peak of radioactivity of mobility distinguishable from that of labeled endogenous proteins and slightly slower than the Coomassie-blue-staining band corresponding to the exogenous protein was used as the criterion for positive acceptor activity.

* Sequences were taken from ref. 10 and 11 except for DNase (25).

carboxypeptidase A, or alcohol dehydrogenase could be detected, although one or more free potential sites for glycosylation occur in each of these proteins.

DISCUSSION

These studies show that hen oviduct membranes catalyze the transfer of oligosaccharide from oligosaccharide-lipid to modified and denatured forms of bovine ribonuclease A and α-lactalbumin. In both cases, the labeled protein product has a higher apparent molecular weight than the protein substrate, as would be expected from attachment of an oligosaccharide side chain. These results, as well as the finding that the labeled product formed in the presence of sulfitolyzed RNase A is electrophoretically identical to authentic glycosylated ribonuclease (RNase B) in two different systems, argue against the possibility that the new labeled products are derived from minor impurities in the protein preparations. Additional support for this conclusion is provided by the findings that: (*i*) sulfitolyzed RNase B does not serve as substrate, and (*ii*) the glycosylated product crossreacts with antiserum prepared against sulfitolyzed RNase A. Since the two forms of ribonuclease appear to differ only by the carbohydrate side chain that blocks asparagine-34 of RNase B (21), the data presented constitute strong evidence that the site for *in vitro* glycosylation of RNase A is asparagine-34. It should be noted that Khalkhali and Marshall have reported that RNase A serves as an acceptor of *N*-actylglucosamine from UDP-*N*-acetylglucosamine in the presence of rabbit liver extracts (26) or human serum (27); the possible participation of lipid intermediates was not examined.

The finding that oligosaccharide-lipid can serve as substrate for the glycosylation of three secretory glycoproteins significantly extends the scope of oligosaccharide-lipid as the precursor for the carbohydrate moiety of glycoproteins. Initially, oligosaccharide-lipid was shown to serve as substrate for glycosylation of endogenous membrane proteins (1, 28). Later, Eagon *et al.* (5) reported the important finding that a carbohydrate-free form of the kappa-type immunoglobulin light chain secreted by mouse myeloma tumor MOPC 46 could be glycosylated by oligosaccharide-lipid in the presence of membrane enzymes from the same tissue. Our results indicate that three soluble proteins can, after sulfitolysis, compete effectively with endogenous membrane proteins as polypeptide acceptors in the oviduct membrane system. Pretreatment that converts ovalbumin into a form that serves as an acceptor also converts two proteins that are completely unrelated in amino acid sequence, and are foreign to the oviduct, into active acceptors. Thus, the enzyme specificity for glycosylation appears to involve, at most, a limited region of the polypeptide chain of the acceptor.

Based on examination of the amino acid sequence of numerous glycoproteins, Marshall proposed that the tripeptide sequence -Asn-X-$^{\text{Thr}}_{\text{Ser}}$- is a necessary, but not sufficient, condition for attachment of an *N*-glycosidically-linked carbohydrate side chain (29). Some proteins, e.g., elastase and carboxypeptidase A, contain one or more potential glycosylation sites, but are not found in a glycosylated form (30). Others, like bovine RNase (21) and probably α-lactalbumin (31), are secreted in both glycosylated and unglycosylated forms. Still others, e.g., ovalbumin (10, 11) and deoxyribonuclease (25), are essentially fully glycosylated at a single site, but contain at least one additional unglycosylated tripeptide. As noted by others (22, 30, 32), a variety of factors may account for the marked variation in the degree of glycosylation *in vivo*, e.g., the presence and the specificity of the appropriate transferase(s), subcellular compartmentalization, and restrictions imposed by polypeptide folding.

Use of the *in vitro* system from hen oviduct described in this study eliminates many of these factors. Thus, it should be possible to focus on the specificity of a particular enzyme and to begin to correlate the acceptor activity of a protein with specific features of its primary and tertiary structure. From the limited number of proteins so far examined several tentative correlations can be drawn. First, all three of the proteins that are acceptors (RNase A, ovalbumin, and α-lactalbumin) contain an -Asn-X-$^{\text{Thr}}_{\text{Ser}}$- tripeptide; none of the proteins tested that lack this tripeptide showed any acceptor activity. Examination of the amino acid sequence of the active proteins has not revealed any common features of sequence in the region adjacent to the required tripeptide. Second, four proteins tested that contained one or more suitable tripeptides were found to be inactive. Thus, the mere presence of the -Asn-X-$^{\text{Thr}}_{\text{Ser}}$- tripeptide does not determine that a protein will serve as an acceptor. As in the case of the active proteins, examination of the amino acid sequence of the inactive proteins in the region of the tripeptide site does not reveal any common feature that would account for their lack of activity.

Finally, acceptor activity has been detected only after the exogenous protein is modified either by sulfitolysis or, in the case of α-lactalbumin, by reduction and alkylation under relatively mild conditions. Thus, in this *in vitro* system, disruption of the native protein conformation appears necessary for recognition of the tripeptide site. Although the temporal relationship between synthesis and glycosylation of a polypeptide chain destined to be a glycoprotein has not been precisely defined, it is clear from the recent work of Kiely *et al.* (33) that glycosylation of ovalbumin can occur while it is still attached to tRNA. Our findings, indicating that unfolding of polypeptide chains is required to expose the appropriate asparagine site, coupled with the observation of Kiely *et al.*, suggest that polypeptide folding may play an important role in regulation of protein glycosylation.

Discussions and collaboration with Dr. Frank Maley as well as with Mr. Douglas Struck were very helpful. We thank Ms. Ann Fuhr for her skilled assistance in preparing the manuscript. This work was supported by grants from the National Institutes of Health (GM 21451) and the National Science Foundation (BMS74-11986).

1. Lucas, J. J., Waechter, C. J. & Lennarz, W. J. (1975) *J. Biol. Chem.* **250**, 1992–2002.
2. Waechter, C. J., Lucas, J. J. & Lennarz, W. J. (1973) *J. Biol. Chem.* **248**, 7570–7579.
3. Chen, W. W., Lennarz, W. J., Tarentino, A. L. & Maley, F. (1975) *J. Biol. Chem.* **250**, 7006–7013.
4. Waechter, C. J. & Lennarz, W. J. (1976) *Annu. Rev. Biochem.* **45**, 95–112.
5. Eagon, P. C., Hsu, A.-F. & Heath, E. C. (1975) *Fed. Proc.* 34, 678.
6. Tai, T., Yamashita, K., Ogata-Arakawa, M., Koide, N., Muramatsu, T., Iwashita, S., Inoue, Y. & Kobata, A. (1975) *J. Biol. Chem.* **250**, 8569–8575.
7. Tarentino, A. L., Plummer, T. H., Jr. & Maley, F. (1972) *J. Biol. Chem.* **247**, 2629–2631.
8. Pless, D. D. & Lennarz, W. J. (1975) *J. Biol. Chem.* **250**, 7014–7019.
9. Pless, D. D. (1976) *Fed. Proc.* **35**, 1539.
10. Dayhoff, M. O., ed. (1972) *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Washington, D.C.), Vol. 5.
11. Dayhoff, M. O., ed. (1973) *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Washington, D.C.), Vol. 5, Suppl. 1.

12. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) *J. Biol. Chem.* 193, 265–275.
13. Cunningham, L. W., Nuenke, B. J. & Strayhorn, W. D. (1957) *J. Biol. Chem.* 228, 835–845.
14. Iyer, K. S. & Klee, W. A. (1973) *J. Biol. Chem.* 248, 707–710.
15. Eaker, D. L., King, T. P. & Craig, L. C. (1965) *Biochemistry* 4, 1473–1478.
16. Kent, C., Krag, S. S. & Lennarz, W. J. (1973) *J. Bacteriol.* 113, 874–883.
17. Fairbanks, G., Steck, T. L. & Wallach, D. F. H. (1971) *Biochemistry* 10, 2606–2617.
18. Pechère, J.-F., Dixon, G. H., Maybury, R. H. & Neurath, H. (1958) *J. Biol. Chem.* 233, 1364–1371.
19. Tarentino, A. L., Plummer, T. H., Jr. & Maley, F. (1974) *J. Biol. Chem.* 249, 818–824.
20. Marshall, R. D. & Neuberger, A. (1972) in *Glycoproteins, Part B*, ed. Gottschalk, A. (Elsevier, New York), p. 740.
21. Plummer, T. H., Jr. & Hirs, C. H. W. (1964) *J. Biol. Chem.* 239, 2530–2538.
22. Beintema, J. J., Gaastra, W., Scheffer, A. J. & Welling, G. W. (1976) *Eur. J. Biochem.* 63, 441–448.
23. Bonner, J., Chalkley, R., Dahmus, M., Fambrough, D., Fujimura, F., Huang, R. C., Huberman, J., Jensen, R., Marushige, K., O[illegible] enbusch, H., Olivera, B. & Widholm, J. (1968) in *Method[illegible] Enzymology*, eds. Grossman, L. & Moldave, K. (Academic P[illegible] New York), Vol. 12B, pp. 32–37.
24. Dixon, G. H. & Wardlaw, A. C. (1960) *Nature* 188, 721–72[illegible]
25. Liao, T.H., Salnikow, J., Moore, S. & Stein, W. H. (1973) *J. B[illegible] Chem.* 248, 1489–1495.
26. Khalkhali, Z. & Marshall, R. D. (1975) *Biochem. J.* 146, 2[illegible] 307.
27. Khalkhali, Z. & Marshall, R. D. (1976) *Carbohydr. Res.* [illegible] 455–473.
28. Hsu, A.-F., Baynes, J. W. & Heath, E. C. (1974) *Proc. Natl. Ac[illegible] Sci. USA* 71, 2391–2395.
29. Marshall, R. D. (1974) *Biochem. Soc. Symp.* 40, 17–26.
30. Hunt, L. T. & Dayhoff, M. O. (1970) *Biochem. Biophys. R[illegible] Commun.* 39, 757–765.
31. Hill, R. L. & Brew, K. (1975) *Adv. Enzymol.* 43, 419–425.
32. Sox, H. C. & Hood, L. (1970) *Proc. Natl. Acad. Sci. USA* [illegible] 975–982.
33. Kiely, M. L., McKnight, S. & Schimke, R. T. (1976) *J. Biol. Ch[illegible]* 251, 5490–5495.

[20] The Porcine Pepsins and Pepsinogens

By A. P. RYLE

Nomenclature

The major acid gastric protease of the pig, pepsin, which has been the subject of an article in this series[1] and of reviews elsewhere[2,3] is referred to by some authors and by the Enzyme Commission[4] as pepsin (EC 3.4.4.1) without an attached letter although other authors refer to the same enzyme as pepsin A. Pepsin A has also been used[5] to designate an unusually highly active enzyme obtained by fractional crystallization. A proposal has been made[6] for the designation of pepsins by letters to indicate their origin and numbers for their electrophoretic mobility, but this article will follow as closely as possible the recommendations of the Enzyme Commission. Thus the main enzyme is simply pepsin, the minor enzymes originally[7] called parapepsins I and II are pepsins B and C, respectively, and the most recently characterized minor enzyme[8] is pepsin D. The zymogens are named correspondingly.

Assay Methods—General Considerations

Two substrates, hemoglobin and acetyl-L-phenylalanyl-L-diiodotyrosine[9] have been used for assay of the minor pepsins and of pepsin. Pepsin and pepsin D are active in both assays, pepsin B only with APD and pepsin C only with Hb. Several other substrates of low molecular weight,[2] some having groups giving enhanced solubility[10-13] have been described for pepsin, and poly-L-glutamate[14] and aromatic sulfite esters[15] are also rapidly hydrolyzed. A very sensitive assay based on peptic inactivation of ribonuclease has been described.[16] These substrates have not been used for routine assay purposes.

In general, assays using small synthetic substrates should be preferred since the substrates are more easily characterized, but when very crude enzyme or zymogen solutions are to be assayed, methods dependent on reaction with ninhydrin give unacceptably high blank values and a protein substrate must be used.

Assay with Hemoglobin as Substrate

Principle. After partial digestion of hemoglobin at pH 1.7, undigested protein is precipitated with trichloroacetic acid. The products soluble in this reagent are estimated spectrophotometrically or by the Folin-Ciocalteu reagent.

Various modifications of the earlier assays[1,17,18] have been described.[7,19] The modifications eliminate a tedious dilution step and require less material. The method is suitable for pepsin, pepsin C, and pepsin D. The corresponding zymogens are activated so rapidly under the conditions of the assay that the method is also suitable for them. Pepsin B and pepsinogen B have almost no activity in the assay. If the enzyme or zymogen is dissolved in a buffered solution, the 0.30 *N* HCl solution used for acidifying the hemoglobin must be replaced by HCl of such greater concentration as will give a pH of 1.7 in the final digestion mixture (i.e., the additional normality should equal the normality of the cations in the buffer whose counterions are the anions of weak acids).

Reagents

HCl, 0.30 *N* (or other suitable concentration, see above)

Trichloroacetic acid, 4.0% (w/v)

Dialyzed hemoglobin, 2.5% (w/v). Bovine hemoglobin enzyme substrate powder (Armour Pharmaceutical Co.) is dissolved in water and dialyzed against two lots of distilled water (each about 10 volumes) for 2–4 hours each or overnight in the cold room. The solution is diluted to bring the concentration to 2.5% and

[1] R. M. Herriott, Vol. II, p. 3.
[2] F. A. Bovey and S. Yanari, *in* "The Enzymes" (P. D. Boyer, H. Lardy, and K. Myrbäck, eds.) 2nd ed. Vol. 4, p. 63. Academic Press, New York, 1960.
[3] R. M. Herriott, *J. Gen. Physiol.* 45, 57 (1962).
[4] Report of the Commission on Enzymes of the International Union of Biochemistry, Macmillan (Pergamon), London, 1961.
[5] R. M. Herriott, V. Desreux, and J. H. Northrop, *J. Gen. Physiol.* 24, 213 (1941).
[6] D. J. Etherington and W. H. Taylor, *Nature* 216, 279 (1967).
[7] A. P. Ryle and R. R. Porter, *Biochem. J.* 73, 75 (1959).
[8] D. Lee and A. P. Ryle, *Biochem. J.* 104, 742 (1967).
[9] The abbreviations used will be: Hemoglobin, Hb; acetyl-L-phenylalanyl-L-diiodotyrosine, APD; diethylaminoethyl, DEAE.
[10] K. Inouye, I. M. Voynick, G. R. Delpierre, and J. S. Fruton, *Biochemistry* 5, 2473 (1966).
[11] K. Inouye and J. S. Fruton, *Biochemistry* 6, 1765 (1967).
[12] T. R. Hollands, I. M. Voynick, and J. S. Fruton, *Biochemistry* 8, 575 (1969).
[13] M. Schlamowitz and R. Trujillo, *Biochem. Biophys. Res. Commun.* 33, 156 (1968).
[14] H. Neumann, N. Sharon, and E. Katchalski, *Nature* 195, 1002 (1962).
[15] T. W. Reid and D. Fahrney, *J. Am. Chem. Soc.* 89, 3941 (1967).
[16] A. Berger, H. Neumann, and M. Sela, *Biochim. Biophys. Acta* 33, 249 (1959).
[17] M. L. Anson, *J. Gen. Physiol.* 22, 79 (1939).
[18] J. H. Northrop, M. Kunitz, and R. M. Herriott, "Crystalline Enzymes" 2nd ed., p. 305. Columbia Univ. Press, New York, 1948.
[19] T. G. Rajagopalan, S. Moore, and W. H. Stein, *J. Biol. Chem.* 241, 4940 (1966).

[illegible] through Whatman No. [illegible] paper. [illegible] (2.5 [illegible] 100 ml) is added as a preservative and the solution can be kept a few days in the cold.

Procedure. The hemoglobin solution is acidified as required each day by the addition of 0.25 volume of HCl of the appropriate concentration and brought to 37°. This substrate (1.0 ml) is added to 0.2 ml of the enzyme solution (0.05–0.2 milli $[PU]^{Hb}$ of pepsin or pepsin D; 0.05–0.12 milli $[PU]^{Hb}$ of pepsin C) at 37° and after 10 minutes the reaction is stopped by the addition of 5 ml of 4% trichloroacetic acid. Blank reactions are performed when very crude material is being assayed by adding the trichloroacetic acid to the enzyme before the hemoglobin; with even partly purified material the total possible contribution of the enzyme to the blank is negligible and a blank value is obtained by using water or buffer in place of enzyme solution. Blank and test samples are run in duplicate. After 5–10 minutes, the samples are filtered through Whatman No. 3 papers (7.0 cm circles) and the absorbances at 280 nm are read against water.

Units of Activity. The mean blank values are subtracted from the mean test values and the ΔE_{280} obtained is converted into $[PU]^{Hb}$ by reference to a standard curve prepared with pure enzyme. The curve obtained may vary slightly from one batch of hemoglobin to another.

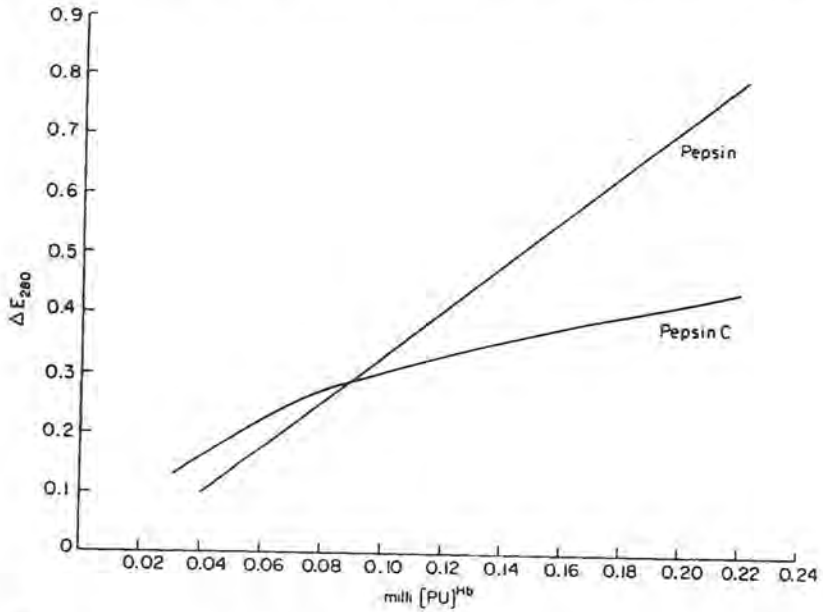


FIG. 1. Standard curves for assay of pepsin and pepsin C with hemoglobin as substrate. See text for details.

One $[PU]^{Hb}$ was defined by Anson[17] as the amount of enzyme which liberated material giving the same color as 1 mg of tyrosine with the Folin-Ciocalteu reagent under conditions slightly different from the above. The curve for pepsin (Fig. 1) was prepared by using a solution of crystalline pepsin which had been previously standardized by Anson's original method.[17] The unit for pepsin C, which gives a curve far from linear, is arbitrarily defined so that $\Delta E_{280} = 0.29$ corresponds to 0.09 milli $[PU]^{Hb}$. The curves for pepsin and pepsin C coincide at this point so that assays performed with mixtures of the enzymes near this level of total activity will not be seriously in error. For pepsin C the assay becomes unreliable when ΔE_{280} exceeds 0.35 so that assays must be restricted to not more than 0.12 milli $[PU]^{Hb}$.

Assay with *N*-Acetylphenylalanyl-L-diiodotyrosine as Substrate[20]

Principle. Hydrolysis of the substrate in 0.01 *N* HCl liberates diiodotyrosine which is estimated by its reaction with ninhydrin. The concentration of the acid used in the digestion mixture is adjusted so that the final concentration of HCl, after titration of the anions of weak acids, is 0.01 *N*. For pepsin[21] $K_m = 0.075$ m*M*, $k_0 = 0.2$ sec^{-1} at pH 2 and 37°.

Reagents

N-Acetylphenylalanyl-L-diiodotyrosine, 1 m*M* in 0.01 *N* NaOH. The acetyl peptide (which is available commercially) is dissolved in the calculated volume of 0.1 *N* NaOH and then diluted to the final concentration. The solution is stable for weeks in the cold.

HCl, 0.05 *N* (or other concentration if the enzyme or zymogen is dissolved in a buffered solution)

Ninhydrin reagent.[22] Ninhydrin (20 g) and hydrindantin (6 g) are dissolved in 750 ml of 2-methoxyethanol. To this is added 250 ml of acetate buffer (544 g $CH_3COONa \cdot 3\ H_2O$ plus 100 ml of glacial acetic acid per liter). Care is taken to avoid getting bubbles of air into the reagent during its preparation and it is stored under nitrogen in a dark bottle from which it can be dispensed.

Ethanol, 60% (v/v)

Procedure. The procedure adopted is suitable for assay of pepsin, pepsin B, pepsin D, and their zymogens. Pepsin C has very little activity

[20] A. P. Ryle and M. P. Hamilton, *Biochem. J.* **101**, 176 (1966).
[21] W. T. Jackson, M. Schlamowitz, and A. Shaw, *Biochemistry* **4**, 1537 (1965).
[22] S. Moore and W. H. Stein, *J. Biol. Chem.* **211**, 907 (1954).

in the assay. The procedure includes a 10-minute incubation of the enzyme or zymogen in acid before the substrate (or ninhydrin in the blank reactions) is added. This allows the release of peptides which occurs on activation of the zymogens to be completed in both test and blank tubes before the assay proper begins; it has no effect on the measured activity of the enzymes. Test and blank reactions are performed in duplicate.

To 0.5 ml of the enzyme or zymogen solution (0.001–0.006 APD units) at 37° is added 0.25 ml of HCl of appropriate normality. After 10 minutes 0.25 ml of APD solution is added to the test reactions and 1.0 ml of ninhydrin reagent is added to the blanks. After 20 minutes more, 1.0 ml of ninhydrin reagent is added to the test reactions and at any time 0.25 ml of APD solution is added to the blanks. All the tubes are placed in a boiling water bath for exactly 15 minutes and are then cooled in a bath of cold water. The contents of the tubes are diluted with 5 ml of 60% (v/v) ethanol and the tubes are then shaken to mix the solutions. The absorbance of the solutions at 570 nm is read against water and ΔE_{570} (mean test value minus mean blank value) calculated. If ΔE_{570} does not exceed 0.4, the reaction is linear with respect to time and the activity can be expressed in APD units.

Definition of Unit. One APD unit is the quantity of enzyme which liberates 1 micromole of diiodotyrosine per minute under the above conditions. Since E_{mmol} for the color reaction of diiodotyrosine is 22.8, the volume of the solution is 7 ml, and the time of incubation is 20 minutes. ΔE_{570} is converted to APD units by multiplying by 0.0153. Specific activity is expressed as APD units/mg N. In some early work, assays were performed at 35.5° and activities are expressed as rate of liberation of diiodotyrosine at that temperature.

Assay of Mixtures of Enzymes and Zymogens

In the assays with hemoglobin or APD as substrate described above, the zymogens are activated and assayed as the active enzyme. It is possible to assay enzymes in the presence of zymogen by using substrates whose hydrolysis can be followed at pH values of 5 or above at which the activation of the zymogens is extremely slow. Assays based on the clotting of milk for pepsin C[23] and for pepsin[24] and on the liquefaction of gelatin solutions for pepsin B[25] and for pepsin[24] have been described. These laborious assays are most unsuitable for use as routine procedures and they will not be described in detail here.

[23] A. P. Ryle, *Biochem. J.* **75**, 145 (1960).
[24] R. M. Herriott, *J. Gen. Physiol.* **21**, 501 (1938).
[25] A. P. Ryle, *Biochem. J.* **96**, 6 (1965).

Since pepsin is rapidly inactivated at pH 8, pepsin and pepsinogen in a mixture can be determined by assay of the activity at low pH directly (giving the sum of pepsin and pepsinogen) and assay at low pH after alkaline inactivation of the enzyme (giving the potential activity of pepsinogen). The activity of the enzyme is found by difference.[24] Mixtures of pepsin C and pepsinogen C may be assayed similarly but a pH of 8.5 is required for the rapid inactivation of the pepsin C.[26]

Assay of the Pepsins in the Presence of one Another

It should be possible to make use of the differences of stability and of substrate specificity to determine pepsin B, pepsin C, and pepsin plus pepsin D in a mixture. Assay with APD will give the sum of pepsin, pepsin D, and pepsin B. The same assay after treatment of the enzyme solution at pH 6.9, at which pepsin B is stable,[7] will give the activity of pepsin B alone, and assay with hemoglobin will give the sum of pepsin, pepsin D, and pepsin C. This method has not been tested beyond its use in identifying the enzymes and zymogens on preparative chromatograms.

Purification of the Zymogens

The preparation of crystalline porcine pepsinogen has been described[1,24,27] and essentially homogeneous crystalline pepsinogen from which a small contaminant can be removed by chromatography[19] is available commercially. The zymogen has also been purified by chromatography on DEAE-cellulose[28] and is obtained as a byproduct in the chromatographic purification of pepsinogens B, C, and D. This method of ion-exchange chromatography gives a demonstrable separation of pepsinogen from pepsinogens B, C, and D; it is not known whether the method of Liener[28] gives such a separation or not. A description is therefore given of the joint purification of pepsinogen and pepsinogens B, C, and D. The procedure is summarized in Table I.

Preparation of the Mucosal Extract.[25] The body (fundic) region of the stomach of pigs (recognizable by the darker brown or red of the mucosa) is collected on ice as soon as possible after slaughter and brought to the laboratory. It is convenient to work with 10–12 stomachs. All subsequent operations are performed in the cold room.

The mucosa is stripped from the muscle, scraped firmly with a microscope slide to remove as much as possible of the adhering mucus,

[26] K. K. Oduro, M.Sc. Thesis, University of Edinburgh, Edinburgh, Scotland, 1967.
[27] R. M. Herriott, *in* "Crystalline Enzymes" (J. H. Northrop, M. Kunitz, and R. M. Herriott, eds.) 2nd ed., p. 259. Columbia Univ. Press, New York, 1948.
[28] I. E. Liener, *Biochim. Biophys. Acta* **37**, 522 (1960).

TABLE I

SUMMARY OF PREPARATION OF ZYMOGENS AND ENZYMES[a]

	Ion-exchange chromatography		Exclusion chromatography	Freeze dried	
	1st	2nd		/E_{280}	/mg N
Pepsinogen B	—	0.20	0.22	0.23	2.22[b]
Pepsin B	0.11	0.21	0.22	—	1.9, 2.1[b]
Pepsinogen C	0.01–0.02	0.03–0.035	0.040	0.040	0.37
Pepsin C	0.018–0.025	0.026–0.030	0.033	0.035	0.41
Pepsinogen D	—	0.02	0.023	—	0.18
Pepsin D	0.02	0.02	0.02	—	0.20
Pepsinogen	—	—	0.020	0.022	0.21
Pepsin	0.02	0.019	0.020	0.020	0.20

[a] The figures show units of (potential) activity/ml/E_{280} and for the freeze-dried material units/mg N. For pepsin(ogen) B APD units, and for the other enzymes and zymogens $[PU]^{Hb}$ are given. For pepsin and pepsinogen 1.0 $[PU]^{Hb}$ corresponds to 6.8 APD units, for pepsin D and pepsinogen D a corresponding ratio of 7.0 has been found. Recovery of 70–80% is usually found on ion-exchange chromatography and close to 100% on exclusion chromatography.

[b] Specific activities of 1.84 and 1.86 were reported earlier for pepsinogen B and pepsin B, respectively [A. P. Ryle, *Biochem. J.* **96**, 6 (1965)]. Other preparations have given these higher figures.

and minced in a Latapie mincer. The mince is stirred for 2 hours or overnight with 4 liters/kg of 0.02 *M* phosphate buffer, pH 6.9 (0.01 *M* NaH_2PO_4, 0.01 *M* Na_2HPO_4), containing 5000 units of penicillin and 50 mg of streptomycin per liter. Filter-Cel (100 g/liter of buffer) and Hyflo Super-Cel (50 g/liter of buffer) are then stirred in and the liquid is separated by filtration or by centrifugation (1000 *g*, 10 minutes). The extract (which may still be cloudy) contains about 60 potential $[PU]^{Hb}$/liter; assays with APD are impossible because of the high blank values.

Chromatographic Separation of the Zymogens.[25] The extract obtained above is stirred for ½ to 1 hour with 0.15 volume of a suspension of DEAE-cellulose[29] (capacity about 1 meq/g) equilibrated with 0.02 *M* phosphate buffer, pH 6.9 whose concentration is such that the volume of the settled bed of ion-exchanger is about half that of the suspension. The DEAE-cellulose with adsorbed proteins is centrifuged off (1000 *g*, 10 minute). The potential activity of the supernatant is now 5–10% of that before adsorption of the zymogens. More of the zymogens can be removed by further treatment with DEAE-cellulose, but the yields of the minor pepsinogens are not thereby noticeably increased.

The DEAE-cellulose is then washed three times (1000 *g*, 10 minutes) with 0.02 *M* phosphate buffer, pH 6.9, by centrifugation in order to remove mucus which otherwise makes the subsequent chromatography very slow. The DEAE-cellulose with the adsorbed zymogens is resuspended and transferred to the top of a column of DEAE-cellulose (10 × 5 cm diameter) already equilibrated with 0.02 *M* phosphate buffer, pH 6.9. The total height of the packed bed is about 20 cm. Elution is performed with a rising gradient of sodium chloride concentration obtained by passing 0.02 *M* phosphate buffer containing 0.45 *M* NaCl into a closed mixing vessel containing 2 liters of 0.02 *M* phosphate buffer, pH 6.9, the buffers being saturated with toluene, and 50-ml fractions are collected. A typical elution profile is shown in Fig. 2.

The pepsinogen C is found in the peak emerging after 4 liters. The fractions from about 2.5–3 liters contain pepsinogen B, detected by the

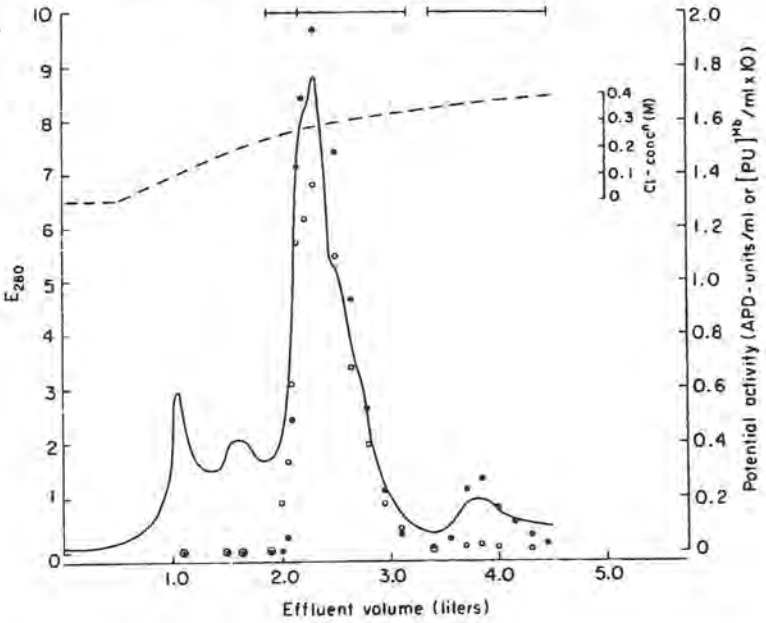


FIG. 2. Chromatography on DEAE-cellulose of zymogens adsorbed from a neutral extract of pig fundic mucosa. The column (20 × 5 cm) was packed in 0.02 *M* phosphate buffer, pH 6.9 and elution was performed at approx. 50 ml/hour with a rising gradient of sodium chloride concentration in the same buffer. ——, E_{280}; ----, chloride concentration; ●, activity against Hb; ○, activity against APD (curves for the activities have not been drawn in). The bars at the top show, from left to right, the fractions pooled for preparation of pepsinogens B and D and pepsinogen, fractions containing mainly pepsinogen, and fractions pooled for preparation of pepsinogen C.

[29] E. A. Peterson and H. A. Sober, *J. Am. Chem. Soc.* 78, 751 (1956).

high ratio of potential activity against APD to that against Hb (APD units/$[PU]^{Hb} > 7.0$); they also contain pepsinogen D and pepsinogen. The remainder of the large peak contains almost exclusively pepsinogen. In one chromatogram run on Whatman DEAE-cellulose, DE-52, a similar elution profile was found, but all the peaks emerged about 0.5 liter earlier.

Appropriate fractions are pooled (the rechromatography of pepsinogen B is unsuccessful if too much pepsinogen is present so fractions beyond the top of the peak should not be included in the pool), and the zymogens are concentrated either by dialysis against polyethylene glycol or by ultrafiltration through cellulose (Visking $^{24}/_{32}$ inch tubing) or more rapidly through a Diaflo membrane UM20E (Amicon N.V., 43 Heemskerckstraat, The Hague, Holland). The recovery of potential activity against hemoglobin on such chromatograms is 70–80%. The zymogens are dialyzed against 0.02 *M* phosphate buffer, pH 6.9 in preparation for rechromatography.

Repeated Ion-Exchange Chromatography of the Zymogens.[20,25,30] Rechromatography is performed by applying the dialyzed solutions to columns of DEAE-cellulose (30 × 2.9 cm) equilibrated with 0.02 *M* phosphate buffer, pH 6.9. Elution is achieved as above with the same (2 liter) size of mixing vessel so that the effective sodium chloride gradient is less steep. For the rechromatography of pepsinogen C, the mixing vessel may initially contain buffer with 0.1 *M* NaCl without loss of resolution. A typical elution pattern for the fraction containing pepsinogen B, pepsinogen D, and pepsinogen is shown in Fig. 3. Successful separation into three peaks is not achieved if the quantity of pepsinogen greatly exceeds that of the other zymogens.

The chromatograms of pepsinogen C show a single peak of potential activity emerging at 3 liters (or earlier if the initial NaCl concentration is 0.1 *M*) with small peaks of inactive material.

Suitable fractions are pooled and concentrated by ultrafiltration. Polyethylene glycol cannot be used at this stage because it penetrates the dialysis sac[31] and is not completely removed by the subsequent exclusion chromatography. The recovery of potential activity on rechromatography as above is 70–80%.

Further Purification of Pepsinogen B.[25] In the original report of the isolation of pepsinogen B, material obtained as above was chromatographed on Sephadex G-100 (43.5 × 2.2 cm) in 0.02 *M* phosphate buffer, pH 6.9, containing 0.45 *M* NaCl. This failed to give complete separation from inactive material, and the material from the active fractions was

[30] D. Lee and A. P. Ryle, *Biochem. J.* **104**, 735 (1967).
[31] A. P. Ryle, *Nature* **206**, 1256 (1965).

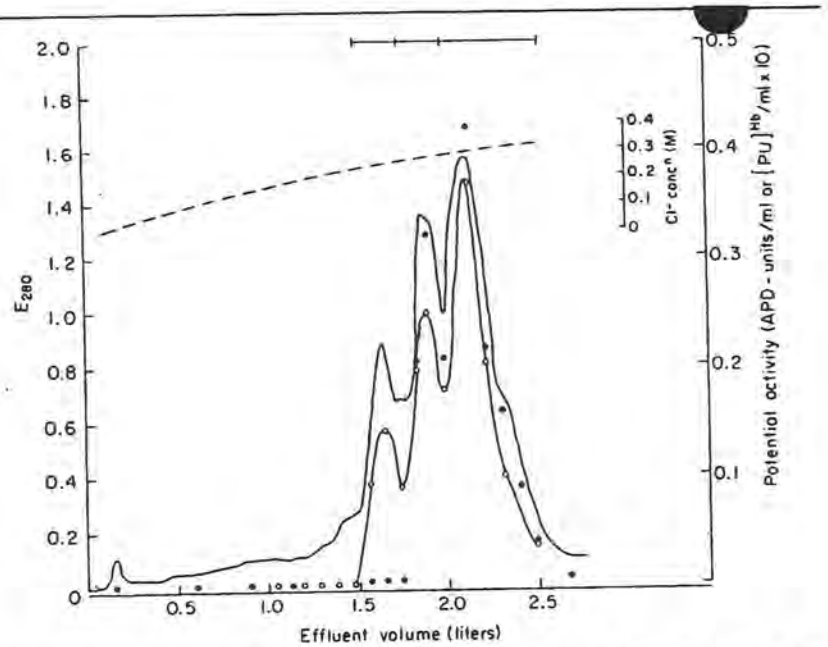


FIG. 3. Rechromatography of pepsinogen and pepsinogens B and D on DEAE-cellulose. The conditions were as in Fig. 2 except that the column was 21 × 2.4 cm and the flow rate 22 ml/hour. Symbols as in Fig. 2 except that the bars at the top show, from left to right, the fractions pooled as pepsinogen B, pepsinogen D, and pepsinogen.

chromatographed by ion-exchange again as described above. The material from this chromatogram gave a single symmetrical peak of absorbance at 280 nm and of potential activity against APD when chromatographed on Sephadex G-200. The low peak of potential activity against hemoglobin which was found very slightly preceded the peaks of absorbance and of potential activity against APD, indicating some residual contamination with pepsinogen or with pepsinogen D. The recovery of potential activity against APD was 100%. This method of purification has not been investigated much further. In some experiments the chromatography on Sephadex G-100 successfully separated pepsinogen B from inactive material. The specific potential activity of pepsinogen B obtained as described above was 1.84 micromoles of diiodotyrosine liberated per minute per mg N at 35.5°.

Exclusion Chromatography of Pepsinogen C.[20] The concentrated material obtained by ion-exchange chromatography as above is chromatographed on Sephadex G-100 which provides a better separation from inactive, ultraviolet-absorbing material than does Sephadex G-75 used

originally. The column used is 15 × 2.2 cm, the buffer is 0.02 *M* phosphate, pH 6.9, containing 0.1 *M* NaCl, and 10 ml fractions are collected. The active material (recovery is over 90%) emerges between two peaks of inactive material about midway between the void volume and the elution volume of small molecules. The zymogen is dialyzed against distilled water and freeze dried. The specific potential activity of zymogen so obtained is fairly consistently about 0.37 $[PU]^{Hb}$/mg N.

Exclusion Chromatography of Pepsinogen D.[30] Chromatography on Sephadex G-100 is carried out exactly as for pepsinogen C above with a similar separation from inactive material and recovery over 90%. The pooled material is dialyzed against 0.5 m*M* NH_4OH (exhaustive dialysis against water results in some activation of the zymogen) and freeze dried. The specific potential activity of the zymogen is 0.18 $[PU]^{Hb}$/mg N and 1.75 micromole of diiodotyrosine liberated per minute per milligram N at 35.5°.

Further Purification of Pepsinogen. Unpublished experiments show that pepsinogen can be further purified by repeated ion-exchange chromatography and exclusion chromatography on Sephadex G-100 in the same way as pepsinogens C and D. The material is dialyzed against 0.5 m*M* NH_4OH and freeze dried. The specific potential activity of the zymogen is then 0.21 $[PU]^{Hb}$/mg N (identical with that of crystalline pepsinogen[1]) and 1.4 APD unit/mg N.

Preparation of the Enzymes

The preparation of crystalline pepsin was reported long ago.[1,32,33] Pepsin B was first found in commercial crystalline pepsin, and pepsins B and C were purified from commercial crude pepsin. The heterogeneity of pepsin preparations has been observed many times, most recently by chromatography on hydroxylapatite.[19] The authors of the latter paper describe a procedure for preparing homogeneous pepsin from crystalline pepsinogen; this is given below after the preparations of the minor enzymes which are summarized in Table I (see p. 322).

Pepsins B and C.[7] The operations are all conducted in the cold room. Pepsin B.P. (500 g) is dissolved in 1 liter of 0.5 *N* H_2SO_4, the solution is filtered, and the enzymes are precipitated by addition of 1 liter of saturated $MgSO_4$, suspended in 0.1 *M* acetate buffer, pH 5.6, and dialyzed against the same buffer. (This procedure avoids the use of NaOH[7,32] for re-solution of the enzymes.) Any precipitate is removed, and the supernatant containing 6–8 g of pepsin is run onto a column (35 × 5 cm) of DEAE-cellulose equilibrated with 0.1 *M* acetate buffer, pH 4.0. The column is washed with the same buffer until E_{280} of the effluent is less than 1.0 when a rising gradient of NaCl concentration (obtained by passing 0.1 *M* acetate buffer pH 4.0 containing 0.2 *M* NaCl into a closed mixing vessel containing 2.6 liters of buffer free of NaCl) is applied. Fractions of 50 ml are collected and pepsin B, pepsin C, and pepsin emerge with maxima at about 3, 4, and 7.5 liters of effluent. Pepsins B and C are recognized by their characteristic differential activities with the substrates hemoglobin and APD.

The pepsin B and pepsin C are pooled separately, concentrated by ultrafiltration, dialyzed against 0.1 *M* acetate buffer, pH 5.6 (enzyme which may have precipitated redissolves) and further separated from one another by rechromatography at pH 4.0 on DEAE-cellulose columns (40 × 2.4 cm) with a gradient obtained as above, except that the volume of the mixing vessel is 560 ml. Pepsin B emerges at 0.75 liter and pepsin C at 1.1 liters. The pooled fractions (recovery of enzymatic activity 70–80%) are concentrated by ultrafiltration to about 5 ml and dialyzed against 0.1 *M* acetate buffer, pH 5.6, in readiness for purification by exclusion chromatography.

Exclusion Chromatography of Pepsin B. In unpublished experiments pepsin B, obtained as above, has been further purified by exclusion chromatography on columns of Sephadex G-75 in 0.1 *M* acetate buffer, pH 5.6 or in 0.02 *M* phosphate buffer, pH 6.9, containing 0.45 *M* sodium chloride. (Any contaminating pepsin or pepsin C was inactivated at the pH of the second experiment.) In both cases the enzyme emerged as a symmetrical peak shortly after a small peak of inactive material, the recovery of activity being close to 100%. The specific activities of the freeze-dried, dialyzed enzymes obtained were 1.86 and 2.11 micromoles of diiodotyrosine liberated per minute per milligram N at 35.5°, respectively.

Exclusion Chromatography of Pepsin C.[20] The material (10–20 $[PU]^{Hb}$) obtained by ion-exchange chromatography is applied to a column (40 × 2.2 cm) of Sephadex G-100 equilibrated with 0.1 *M* acetate buffer, pH 5.6, and eluted with the same buffer, 10-ml fractions being collected. The enzyme emerges at about 100 ml, before a peak of ultraviolet-absorbing, inactive material. The fractions containing enzymatic activity are pooled, dialyzed against distilled water, and freeze dried. The specific activity of the enzyme obtained is 0.41 $[PU]^{Hb}$/mg N.

Pepsin D.[8] The enzyme is not detected in chromatograms of pepsin run at pH 4.0 but has been prepared from commercial crystalline pepsin by ion-exchange chromatography at pH 3.2.

[32] J. H. Northrop, *J. Gen. Physiol.* **13**, 739 (1930).
[33] J. H. Northrop, *J. Gen. Physiol.* **30**, 177 (1947).

A column (16 × 5 cm) of DEAE-cellulose is prepared by equilibrating the ion-exchanger with 0.1 *M* citrate buffer, pH 3.2, containing 0.1 *M* NaCl and then washing it with 0.1 *M* citrate buffer, pH 3.2, until chloride cannot be detected in the effluent. Successful separation of pepsin D from pepsin is not obtained unless the DEAE-cellulose is pretreated in this way. The enzyme (1.5 g), dissolved in 0.1 *M* citrate buffer, pH 3.2, is applied to the column and elution is performed with a gradient of sodium chloride in the same buffer obtained by passing buffer containing 0.1 *M* NaCl into a closed mixing vessel containing 2 liters of buffer free of sodium chloride. Pepsin D emerges at 1 liter and pepsin at 1.5 liters. The fractions containing pepsin D are concentrated to about 5 ml by ultrafiltration and dialyzed against 0.1 *M* acetate buffer, pH 4.0.

Exclusion Chromatography of Pepsin D.[8] Filtration of the enzyme obtained as above through a column of Sephadex G-75 (19.5 × 4.8 cm), equilibrated with 0.1 *M* acetate buffer, pH 4.0, gives a peak of enzymatic activity emerging just after the void volume and separated from a peak of inactive material emerging later. The enzyme can be dialyzed against distilled water and freeze dried. The specific activity found is 0.20 [PU][11]/mg N and 1.61 micromoles of diiodotyrosine liberated/minute/mg N at 35.5°.

Preparation of Pepsin from Pepsinogen.[19] The method affords pepsin which is homogeneous as tested by a number of methods. A solution of 75 mg of crystalline pepsinogen in glass-distilled water (the solution has a pH of 5–6) is brought to 14° and acidified to pH 2.0 (glass electrode) while being stirred, by the addition of 0.75 ml of a chloroacetic acid–hydrochloric acid mixture (9 ml of 2 *N* chloroacetic acid and 1 ml of 1 *N* HCl). The acidification must be done rapidly, in 20–25 seconds. If the preparation of pepsinogen contains much salt, more acid is required and the concentration of the HCl should be increased; alternatively the pepsinogen may be freed of salt by previous passage through a Sephadex G-25 column in water.[34] The solution is maintained at 14° and pH 2.0 for exactly 20 minutes and the pH is then raised to 4.40 by the addition of 0.75 ml of 4 *N* sodium acetate buffer, pH 5.0 (544 g of $CH_3COONa \cdot 3\ H_2O$ + 100 ml of acetic acid made up to 1 liter). The activation mixture is applied within 5 minutes to a column (30 × 2.5 cm) of sulfoethyl Sephadex C-25 equilibrated with 0.40 *N* sodium acetate buffer pH 4.38 ± 0.03 (40.29 g of $CH_3COONa \cdot 3\ H_2O$ plus 29 ml of glacial acetic acid made up to 2 liters). The column is operated at 4° under gravity at about 60 ml/hour. The effluent is collected in 4-ml fractions and those emerging at the void volume and having absorbance in the ultraviolet are pooled. (The authors monitored the column effluent at 254 nm; measurements at 280 nm would also be suitable). The pooled fractions are brought to pH 5 by addition of 2 *N* sodium acetate, divided into small lots, frozen, and stored at —20°. The enzyme can be transferred to water or to buffer of any desired composition by passage through a Sephadex G-25 column.

[11] S. Moore, personal communication, 1968.

Properties

Stability. Pepsin is unstable at pH values above 6 but is relatively stable at pH values even as low as pH 1. Pepsinogen is unstable at pH values below 5, being activated to pepsin. The zymogen undergoes reversible denaturation at elevated temperatures at pH 7 or at pH values in the range 8.5–11 at room temperature.[1]

Pepsin B is stable at pH 6.9 and 25° (footnote 7) and pepsin C is much more stable than pepsin under these conditions, having a half-life of 70 minutes.[7] At pH 8.5 pepsin C is rapidly inactivated.[26] Pepsin D[8] is very unstable at pH values above 6. All the zymogens are converted to enzyme at pH values below 5; enzymes and zymogens are stable at pH 5.6.

No study of the alkali-lability of pepsinogen B has been made; pepsinogen C is stable at pH 8.5 and 37° (footnote 26) and pepsinogen D starts to lose activity at pH values above 7.5 at 35.5°, this inactivation being at least partly reversible.[30]

Purity and Physical Properties. Tables II and III list some of the properties of the enzymes and zymogens.

The preparations described above yield enzymes and zymogens which appear chromatographically homogeneous. Both pepsinogen B[25] and pepsin B[7] also appear homogeneous in the ultracentrifuge but on electrophoresis in starch gel at pH 6.9 both are found to contain two components with activity or potential activity against APD, and analysis suggests that the two components of pepsinogen B have different amino-terminal residues.[25]

Pepsin C[7] and pepsinogen C,[23] even without purification by exclusion chromatography, appeared homogeneous in the ultracentrifuge. Staining for protein after electrophoresis in starch gel showed them to be homogeneous, but very slight contamination by pepsin (or pepsin D) and pepsinogen (or pepsinogen D) could be shown by the sensitive method of detecting hemoglobin-digesting activity in the gel. The amino acid compositions[20] (Table IV) of the enzyme and zymogen provide further evidence of essential homogeneity since nearly integral values were found for the amino acids.

Pepsinogen D and pepsin D, examined by electrophoresis in starch

TABLE II

SOME PROPERTIES OF PEPSINOGENS AND PEPSINS

	$S_{20,w}$ (sec × 10^{13})	$D_{20,w}$ (cm^2/sec × 10^7)	Molecular weight	Amino terminus
Pepsinogen B[a]	2.99	7.40	39000	0.8 Met, 0.3 His
Pepsin B[b]	3.26		38600	Ala
Pepsinogen C[c]	3.36		40500, 41400[d]	Ser
Pepsin C[b]	3.32		40700, 36000[d]	Ser and Leu or Ile
Pepsinogen D[e]			35000[f]	Leu
Pepsin D[e]			41000[f]	Ile
Pepsinogen	3.24 ± 0.15[h]		40400 ± 1600[h]	Leu
			38944[i]	
	3.2–3.35[j]		41000 ± 1000[j]	
	3.6[k]	7.54[k]	42240[k]	
Pepsin	2.88 ± 0.15[h]		32700 ± 1200[h]	Ile
			34163[i]	
	2.96–3.0[l]		35000,[l] 36212[l]	
	3.25[k]	8.7[k]	32930[k]	

[a] A. P. Ryle, *Biochem. J.* **96**, 6 (1965).
[b] A. P. Ryle and R. R. Porter, *Biochem. J.* **73**, 75 (1959).
[c] A. P. Ryle, *Biochem. J.* **75**, 145 (1960).
[d] A. P. Ryle and M. P. Hamilton, *Biochem. J.* **101**, 176 (1966). Molecular weight from amino acid composition.
[e] D. Lee and A. P. Ryle, *Biochem. J.* **104**, 735 (1967).
[f] From exclusion chromatography.
[g] D. Lee and A. P. Ryle, *Biochem. J.* **104**, 742 (1967).
[h] R. C. Williams and T. G. Rajagopalan, *J. Biol. Chem.* **421**, 4951 (1966).
[i] T. G. Rajagopalan, S. Moore, and W. H. Stein, *J. Biol. Chem.* **241**, 4940 (1966). Molecular weight from amino acid composition.
[j] R. Arnon and G. E. Perlmann, *J. Biol. Chem.* **238**, 653 (1963).
[k] V. N. Orekhovitch, V. O. Shpikiter, and V. I. Petrova, *Dokl. Akad. Nauk. S.S.S.R.* **11**, 401 (1956).
[l] O. O. Blumenfeld and G. E. Perlmann, *J. Gen. Physiol.* **42**, 533 (1959). Molecular weight 36212 from amino acid composition.

TABLE III

SOME OPTICAL PROPERTIES OF PEPSINOGEN AND PEPSIN

	λ_c(nm)	$-[\alpha]_{366}$	E^{278}_{Mol}
Pepsinogen	236[a]	212°[a]	51300 (M = 41000)[b]
Pepsin	216[c]	232°[c]	50990 (M = 35000)[b]

[a] S. Rimon and G. E. Perlmann, *J. Biol. Chem.* **243**, 3566 (1968).
[b] R. Arnon and G. E. Perlmann, *J. Biol. Chem.* **238**, 963 (1963).
[c] G. E. Perlmann, *Proc. Natl. Acad. Sci. U.S.* **45**, 915 (1959).

TABLE IV

AMINO ACID COMPOSITION OF PEPSINOGENS AND PEPSINS

	Moles of amino acid per mole of protein			
	Pepsinogen C[a]	Pepsin C[a]	Pepsinogen[b,c]	Pepsin[b,d]
Lys	12	4	10	1
His	2	1	3	1
Arg	7	4	4	2
Cys	6	6	6	6
Asp	30	28	44	40
Thr	25	25	26	25
Ser	35	35	46	43
Glu	46	41	28	26
Pro	20	18	19	16
Gly	35	32	35	34
Ala	23	21	19	16
Val	22	20	23	20
Met	5	4	4	4
Ile	16	14	25	23
Leu	40	34	33	28
Tyr	22	18	17	16
Phe	24	21	15	44
Trp	6	6	6	6
NH_3	34	32	27	27

[a] A. P. Ryle and M. P. Hamilton, *Biochem. J.* **101**, 176 (1966).
[b] T. G. Rajagopalan, S. Moore, and W. H. Stein, *J. Biol. Chem.* **241**, 4940 (1966).
[c] Slightly different analyses are given by R. Arnon and G. E. Perlmann, *J. Biol. Chem.* **238**, 653 (1963) and H. van Vunakis and R. M. Herriott, *Biochim. Biophys. Acta* **32**, 600 (1957).
[d] A slightly different analysis is given by O. O. Blumenfeld and G. E. Perlmann, *J. Gen. Physiol.* **42**, 553 (1959).

gel, give rather broad zones of hemoglobin-digesting activity but show essential freedom from contamination by pepsinogen and pepsin. Analysis shows about 0.3 g atom of phosphorus per mole in both zymogen and enzyme. Treatment with a phosphatase fails to alter the electrophoretic mobility of the enzyme or activated zymogen and the residual phosphate may represent contamination with inorganic phosphate or nucleic acid.

Crystalline pepsinogen and pepsin prepared from it have been shown to be homogeneous by the criteria of chromatography, amino acid analysis and carboxyl terminal analysis,[19] and by their behavior in the ultracentrifuge.[35]

Amino Acid Composition and Sequences. Table IV shows the compositions of pepsin, pepsin C, and their zymogens. No analyses have been

[35] R. C. Williams and T. G. Rajagopalan, *J. Biol. Chem.* **241**, 4951 (1966).

Amino terminus of pepsinogen[a]

Leu-Val-Lys-Val-Pro-Leu-Val-Arg-Lys-Lys-Ser-Leu-Arg-Glu-Asn-Leu-Ile-Lys-Asp-Gly-Lys-Leu-Lys-Asp-Phe-Leu-Lys-Thr-His-Lys-His-Asn-Pro-Ala-Ser-Lys-Tyr-Phe-Pro-Ala-Glu

Amino terminus of pepsin[b]

Ile-Gly-Asp-Glu-Pro

Tryptophan peptides of pepsin[c]

Val-Phe-Asp-Asn-Leu-Trp-Asp-Glu-Gly
Leu-Trp-Val-Pro-Ser
Val-Glu-Gly(Trp,Glu)
Leu-Asn-Trp-Val-Pro

Cystine peptides of pepsin[d]

Cys-Ser-Ser-Ile-Asp-Glu Glx-Asx(Asx,Ser)Cys(Thr,Ser,Asp,Glu)
Cys-Ser-Gly-Gly-Cys-Glu
Cys-Ser-Ser-Leu-Ala-Cys-Ser-Asx-His(Glx,Asx)

Cyanogen bromide clearage peptides of pepsin[e]

P
|
Ile-Gly-Asp-Glu. . .Glu-Ala-Thr-Ser-Glu-Glu-Leu. . .(Ile,Thr)Met (Peptide B-2, amino terminus of pepsin, 130–140 residues)
Val-Ile. . .Met (32 residues)
Asp-Gly-Glu-Thr-Ile. . .Met (38–39 residues)
Asp-Val-Pro-Thr-Ser- (continues to C-terminus.)

Peptides from peptide B-2 above[f]

Asp-Gly-Ile-Gly-Leu-Leu
P
|
Glu-Ala-Thr-Ser-Glu-Glu-Leu-Ser(Thr,Ile)Tyr
Ser-Val-Tyr
Cys-Ser-Ser-Leu(Cys,Ser_2,Ala,Asx_3,His,Thr,Glx,Pro)Phe-Phe
Gly-Leu-Ser-Glu-Thr-Glu-Pro-Gly-Ser-Phe
Leu-Tyr
Tyr-Ala-Pro-Phe
Leu-Gly-Gly-Ile(Ser,Asp)Ser-Tyr-Tyr

Carboxyl terminus of pepsin[g]

Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu-Leu-Trp-Ile(Asp,Thr,Ser_2,Glu,Pro,Gly,Val,Leu,Tyr,Phe)Ile[h]-Leu-Gly-Asp-Val-Phe-Ile[i]-Arg[j]-Glu-Tyr-Tyr-Thr-Val-Phe-Asp-Arg-Ala-Asn-Asn-Lys-Val-Gly-Leu-Ala-Pro-Val-Ala

Peptide sequence round an active aspartate residue[k]

Ile-Val-Asp-Thr-Gly-Thr-Ser

[a] E. B. Ong and G. E. Perlmann, *J. Biol. Chem.* **243**, 6104 (1968).
[b] V. M. Stepanov, *Abstr. FEBS Meeting 5th, Prague*, no. 1089, (1968).
[c] T. A. A. Dopheide and W. M. Jones, *J. Biol. Chem.* **243**, 3906 (1968).
[d] J. Tang and B. S. Hartley, *Biochem. J.*, in press.

reported for pepsinogen B or pepsin B. The amino acid compositions of pepsinogen D and pepsin D are not distinguishable from those of pepsinogen and pepsin, respectively.[35a] Table V lists the amino acid sequences identified in pepsin and pepsinogen. It has recently been reported that pepsinogen contains about 3 moles of sugar per mole and that the sugar is lost from the protein on activation to pepsin.[36]

Activators and Inactivators. No activators of pepsin have been reported. The protease activity is destroyed to the extent of 70–80%, and the activity with benzyloxycarbonyl-L-glutamyl-L-tyrosine is destroyed completely, by reaction with *p*-bromophenacyl bromide.[37] The analogous α-diazo-*p*-bromoacetophenone causes complete inactivation, and is also capable of reacting with, and completely inactivating, pepsin previously treated with *p*-bromophenacyl bromide.[38] With both reagents 1 mole of inhibitor is bound per mole of enzyme. The reagents apparently react with different amino acid residues in the enzyme.

With diphenyldiazomethane[39] 50% loss of activity against hemoglobin is achieved with reaction of about 2.4 moles of reagent per mole of enzyme. The reagent also reacts with, and prevents the activation of, pepsinogen, and the inactivation of pepsin is enhanced by the presence of peptide substrates.[10] These facts suggest that inactivation does not occur solely by reaction at the active center of the enzyme.

[35a] A. P. Ryle and F. Falla, unpublished observations, 1970.
[36] H. Neumann, U. Zehavi, and T. D. Tanksley, *Biochem. Biophys. Res. Commun.* **36**, 151 (1969).
[37] B. F. Erlanger, S. M. Vratsanos, N. Wasserman, and A. G. Cooper, *J. Biol. Chem.* **240**, PC 3447 (1965).
[38] B. F. Erlanger, S. M. Vratsanos, N. Wasserman, and A. G. Cooper, *Biochem. Biophys. Res. Commun.* **28**, 203 (1967).
[39] G. R. Delpierre and J. S. Fruton, *Proc. Natl. Acad. Sci. U.S.* **54**, 1161 (1965).

[e] V. I. Ostoslavskaya, I. B. Pugacheva, E. A. Vakhitova, V. F. Krivtsov, G. L. Muratova, E. D. Levin, and V. M. Stepanov, *Biochemistry (U.S.S.R.)* **33**, 276 (1968),
[f] E. A. Vakhitova, I. B. Pugacheva, M. M. Amirkhayan, R. Valiulis, L. G. Senyutenkova, B. G. Belen'kii, and V. M. Stepanov, *Izv. Akad. Nauk S.S.S.R.; Ser. Khim.* p. 2415 (1968); *Chem. Abstr.* **70**, 25942s (1969).
[g] V. Kostka, L. Morávek, I. Kluh, and B. Keil, *Biochim. Biophys. Acta* **175**, 459 (1969).
[h] From this residue to terminus; see also T. A. A. Dopheide, S. Moore, and W. H. Stein, *J. Biol. Chem.* **242**, 1833 (1967).
[i] From this residue to the terminus; see also R. A. Matveeva, V. F. Krivtsov, and V. M. Stepanov, *Biochemistry (U.S.S.R.)* **33**, 142 (1968).
[j] From this residue to the terminus; see also R. N. Perham and G. M. T. Jones, *European J. Biochem.* **2**, 84 (1967).
[k] R. S. Bayliss, J. R. Knowles, and G. B. Wybrandt, *Biochem. J.* **113**, 377 (1969).

[illegible] has also been found in presence of cupric ions with a series of aliphatic diazocarbonyl compounds,[40] some having structural resemblance to pepsin substrates. The substrate analog L-1-diazo-4-phenyl-3-tosylamidobutanone (tosyl-L-phenylalanyldiazomethane),[41] causes complete inhibition in the presence of cupric ions when 1 mole of inhibitor has reacted per mole of enzyme and does not react with pepsinogen or denatured pepsin. Benzyloxycarbonyl-L-phenylalanyldiazomethane[42] inhibits pepsin similarly, except that cupric ions are not necessary. The reaction of other chromophoric phenylalanine diazoketones has also been investigated.[43]

The diazoacetyl amino acid ester derivatives of norleucine,[44] glycine,[45] and phenylalanine[46] also cause complete inhibition at a 1:1 molar ratio of reaction and the aspartic acid at which the phenylalanine compound reacts[46] may also be the site of reaction of 1-diazo-4-phenyl-butanone-2 (footnote 47 and Table V) and of *N*-diazoacetyl-*N'*-2,4-dinitrophenylethylenediamine.[48]

Pepsin C, like pepsin, is inactivated by diazoacetyl-DL-norleucine methyl ester.[49]

The activation peptides of pepsinogen C[50,50] and those of pepsinogen D,[51] like those of pepsinogen,[52] contain material which is inhibitory to the milk-clotting activity of the enzymes.

Specificity. Pepsin has a rather broad specificity which has been examined with numerous synthetic peptides[2,11] and by a survey of the bonds it is known to hydrolyze in natural polypeptides.[53] Peptides containing *p*-nitrophenylalanine,[11] 3,5-dibromotyrosine,[54] or 3,5-dinitrotyrosine[55] as well as ones containing diiodotyrosine[2] are hydrolyzed and have been used in kinetic investigations. While the general conclusion is that activity is enhanced by the presence of bulky nonpolar sidechains in the amino acids on either side of the sensitive bond, activity is not restricted to such bonds. Poly-L-glutamic acid has been shown[14] to be very rapidly hydrolyzed to oligopeptides; the initial rate of splitting of peptide bonds at pH 2.3 and 35° corresponds to about 100 moles of peptide bond per mole of pepsin per minute when the substrate is about 0.22 millimolar. A free amino group on the acyl moiety of the peptide bond is inhibitory,[2] and a free carboxyl group on the amino moiety is inhibitory when it is dissociated.[10,56,57] Curiously, replacement of the terminal phenylalanine methyl ester in the substrate benzyloxycarbonylphenylalanylphenylalanine methyl ester by phenylalaninol (replacement of —COOMe by —CH_2OH) produces a substance resistant to hydrolysis.[11]

In addition to hydrolyzing peptide bonds, pepsin catalyzes transpeptidation reactions, shown to be of the amino-transfer type[58,59] although the product of activation of pepsinogen at pH 3 lacks such activity.[60]

Pepsin has recently been shown to act as an esterase. It catalyzes the hydrolysis of the ester bond in acetyl-L-phenylalanyl-L(β-phenyl)-lactic acid[61] and the corresponding bond in benzyloxycarbonyl-L-histidyl-*p*-nitro-L-phenylalanyl-L-(β-phenyl)-lactic acid methyl ester.[11] It also catalyzes the hydrolysis of aromatic sulfite esters.[15]

The specificity of the minor pepsins has not been extensively investigated, but their sites of attack on the B-chain of oxidized insulin have been examined.[7,8,62] Pepsin D is indistinguishable by this test from pepsin; pepsins B and C show a slightly more restricted specificity than pepsin. Pepsin B is a highly active gelatinase and a poor milk-clotting enzyme. The most marked difference between the enzymes is that shown by their relative activities against APD and Hb. Pepsin C has been shown to catalyze transpeptidation reactions of the amino-transfer type[63] and although it has little activity against *N*-acylated di- and tripeptides

[40] L. V. Kozlov, L. M. Ginodman, and V. N. Orekhovich, *Biochemistry (U.S.S.R.)* 32, 839 (1967).
[41] G. R. Delpierre and J. S. Fruton, *Proc. Natl. Acad. Sci. U.S.* 56, 1817 (1966).
[42] E. B. Ong and G. E. Perlmann, *Nature* 215, 1492 (1967).
[43] R. A. Badley and F. W. J. Teale, *Biochem. J.* 108, 15P (1968).
[44] T. G. Rajagopalan, W. H. Stein, and S. Moore, *J. Biol. Chem.* 241, 4295 (1966).
[45] R. L. Lundblad and W. H. Stein, *J. Biol. Chem.* 244, 154 (1969).
[46] R. S. Bayliss, J. R. Knowles, and G. R. Wybrandt, *Biochem. J.* 113, 377 (1969).
[47] K. T. Fry, O-K. Kim, J. Spona, and G. A. Hamilton, *Biochem. Biophys. Res. Commun.* 30, 489 (1968).
[48] V. M. Stepanov and T. I. Vaganova, *Biochem. Biophys. Res. Commun.* 31, 825 (1968).
[49] J. Kay and A. P. Ryle, unpublished observations, 1968.
[50] K. K. Oduro and A. P. Ryle, *Abstr. Intern. Congr. Biochem. 7th Tokyo*, F-306 (1967).
[51] D. Lee, *Can. J. Biochem.* 45, 1002 (1967).
[52] H. van Vunakis and R. M. Herriott, *Biochim. Biophys. Acta* 22, 537 (1960).
[53] J. Tang, *Nature* 199, 1094 (1963).
[54] E. Zeffren and E. T. Kaiser, *J. Am. Chem. Soc.* 88, 3129 (1966).
[55] J. R. Knowles, H. Sharp, and P. Greenwell, *Biochem. J.* 113, 343 (1969).
[56] J. L. Denburg, R. Nelson, and M. S. Silver, *J. Am. Chem. Soc.* 90, 479 (1968).
[57] A. J. Cornish-Bowden and J. R. Knowles, *Biochem. J.* 113, 353 (1969).
[58] H. Neumann, Y. Levin, A. Berger, and E. Katchalski, *Biochem. J.* 73, 33 (1959).
[59] J. S. Fruton, S. Fujii, and M. H. Knappenberger, *Proc. Natl. Acad. Sci. U.S.* 47, 759 (1961).
[60] H. Neumann and N. Sharon, *Biochim. Biophys. Acta* 41, 370 (1960).
[61] L. A. Lokshina, V. N. Orekhovich, and V. A. Sklyankina, *Nature* 204, 580 (1964).
[62] A. P. Ryle, J. Leclerc, and F. Falla, *Biochem. J.* 110, 4P (1968).
[63] A. P. Ryle, *Bull. Soc. Chim. Biol.* 42, 1223 (1960).

containing peptide bonds which are sensitive in the B-chain of oxidized insulin, longer peptides containing these bonds are hydrolyzed.[62]

Distribution. Examination of the zymogens found in the gastric mucosa of individual pigs[64] has shown that some animals make all four zymogens but some (at the time of slaughter) apparently lacked one or more of the zymogens. Pepsin C is relatively more abundant than the other zymogens in the pyloric mucosa, but the absolute amounts found in the fundus are much greater. All three minor pepsins as well as pepsin have been found in the stomach contents of four pigs examined.[65]

Status of the Minor Pepsins and Pepsinogens. Since pepsins are acid proteases, and since commercial pepsin is obtained after extensive acid treatment of mucosal extracts and is heterogeneous,[19] one must consider the possibility that the minor pepsins are degradation products of pepsin (or vice versa) or are partial activation products of pepsinogen. The finding of a zymogen in neutral mucosal extracts for each of the enzymes found in crude or crystalline pepsin lends considerable weight to the belief that the enzymes and zymogens are not such degradation products. There is little evidence beyond this to establish pepsin B and pepsinogen B as proteins in their own right except that they do not contain phosphate and so cannot be autolytic precursors of pepsin or pepsinogen. If one is to maintain that pepsin B is a degradation product of pepsin (and similarly for the zymogens) one must postulate either loss of the phosphate residue or its excision as a phosphoserine peptide during the isolation of both zymogen and enzyme under their different conditions of pH.

Pepsinogen D and pepsin D have not been found to differ from pepsinogen and pepsin in any property except chromatographic behavior, electrophoretic mobility, a small difference in specific activity, and phosphate content. Pepsin and pepsinogen become electrophoretically indistinguishable from pepsin D and pepsinogen D after treatment with phosphatase. It is therefore suggested[5,30] that the minor forms may merely be the major ones without the phosphate. It is an open question whether the phosphate is lost during isolation or was never attached during the biosynthesis of the zymogen.

The amino acid analyses of pepsinogen C and pepsin C and those of pepsinogen and pepsin (Table IV) make it impossible that either enzyme should be a degradation product of the other or either zymogen a degradation product of the other (see, for example, the contents of aspartic and glutamic acids).

[64] A. P. Ryle, *in* "The Council for International Organizations of Medical Sciences Symposium The Role of the Gastrointestinal Tract in Protein Metabolism" (H. Munro, ed.), p. 25. Blackwell, Oxford, 1964.
[65] A. P. Ryle, *Biochem. J.* 98, 485 (1966).

[21] Bovine Pepsinogen and Pepsin

By Beatrice Kassell and Patricia A. Meitner

Introduction

Bovine pepsin was isolated from gastric juice and crystallized by Northrop[1] in 1933. Bovine pepsinogen was isolated by Chow and Kassell[2] and shown to differ considerably in amino acid composition from porcine pepsinogen [20].

Assay Methods

The Hemoglobin Method for Total Potential Peptic Activity[2,3]

Principle. The extent of digestion of denatured hemoglobin substrate is determined by measuring the formation of trichloroacetic acid[4] soluble material as increase in absorbance at 280 nm.

Reagents

Buffer 1. HCl, 0.01 *M*, containing 0.1 *M* NaCl (pH 2.0)
Buffer 2. Sodium acetate, 0.1 *M*, pH 5.3
HCl, 0.3 *M*
Substrate. Dissolve 1 g of crystalline hemoglobin (Pentex Biochemicals, Kankakee, Ill.) in 32 ml of water; add 8 ml of 0.3 *N* HCl. Filter through a coarse sintered glass filter. The solution may be kept at 4° for 1 week.
TCA, 5% aqueous solution
Standard Enzyme Solution. Dissolve a few milligrams of crystalline porcine pepsin (Worthington Biochemical Corp., Freehold, N.J.) in 5 ml of the acetate buffer. Measure the absorbance at 280 nm ($A_{280} \times 676 = \mu g/ml$). Prepare several dilutions containing 30–100 μg/ml in the same buffer.
Unknown Pepsinogen or Pepsin Solution. Dilute to about 50 μg/ml with Tris phosphate buffer, 0.05 *M*, pH 7.

Procedure. Predetermine the amount of 0.3 *N* HCl required to bring 0.1 ml of each assay solution to pH 2.0. Place polycarbonate centrifuge tubes (16 × 100 mm) in a 37.6° water bath, add the required amount of 0.3 *N* HCl plus sufficient buffer 1 to bring the volume to 0.9 ml. Add 0.1

[1] J. H. Northrop, *J. Gen. Physiol.* 16, 615 (1933).
[2] R. B. Chow and B. Kassell, *J. Biol. Chem.* 243, 1718 (1968).
[3] M. L. Anson, *J. Gen. Physiol.* 22, 79 (1938).
[4] Abbreviations used are: TCA, trichloroacetic acid; DEAE, diethylaminoethyl.

ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS 169, 627–637 (1975)

The Capacity of α-Amylases to Catalyze the Nonhydrolytic Degradation of Starch and Glycogen with Formation of Novel Glycosylation Products[1]

MASARU TAKESHITA AND EDWARD J. HEHRE

Department of Microbiology and Immunology, Albert Einstein College of Medicine, New York 10461

Received January 1, 1975

α-Amylases have been found to convert starch and glycogen, in part, to products other than hemiacetal-bearing entities (maltose, maltodextrins, etc.)—hitherto, the only products obtained from natural α-glucans by α-amylolysis. Glycosides of maltosaccharides were synthesized by purified α-amylases acting on starch or bacterial glycogen in the presence of *p*-nitrophenyl α- or β-D-glucoside. From a digest with crystallized *B. subtilis* var. *amyloliquefaciens* α-amylase, containing 4 mg/ml of [^{14}C]glycogen and 40 mM *p*-NP β-D-glucoside, three pairs of correspondingly labeled glycosides and sugars were recovered: *p*-NP α-D-[^{14}C]glucopyranosyl (1 → 4) β-D-glucopyranoside, and [^{14}C]glucose; *p*-NP α-[^{14}C]maltosyl (1 → 4) β-D-glucopyranoside, and [^{14}C]maltose; *p*-NP α-[^{14}C]maltotriosyl (1 → 4) β-D-glucopyranoside, and [^{14}C]maltotriose. The three glycosides accounted for 11.4% of the [^{14}C]glycogen donor substrate; the three comparable sugars, for 30.4%; higher maltodextrins, for 58.2%. Calculations based on the molar yields of all reaction products show that [^{14}C]glycosyl moieties were transferred from donor to *p*-NP β-D-glucoside with a frequency of 0.234 relative to all transfers to water. This is a very high value considering the minute molar ratio (0.0007) of β-D-glucoside-to-water concentration. Less striking but similar findings were obtained with cryst. hog pancreatic and *Aspergillus oryzae* α-amylases. The results extend earlier findings (Hehre *et al.*, *Advan. Chem. Ser.* (1973) **117**, 309) in showing that α-amylases have a substantial capacity to utilize the C_4-carbinols of certain D-glucosyl compounds as acceptor sites.

Recent studies from this laboratory (1,2) have shown that α-amylases as a group are glycosylases with a wider range of catalytic capabilities than can be accounted for by the usual representation of these enzymes as hydrolases. Highly purified α-amylases of bacterial, animal, and fungal origins have been found not only to synthesize maltosaccharides and maltosaccharide 1-fluorides from α-D-glucosyl fluoride (1) and α-maltosyl fluoride (2), but to do so in preference to hydrolyzing these substrates. The finding that α-amylases have a higher affinity for at least some glucosyl compounds (as acceptors) than for water has led us to investigate whether, and to what extent, degradation of starch or glycogen by α-amylases might be diverted from hydrolysis by the presence of moderate concentrations of suitable acceptors. In general, biochemists and physiologists have assumed that the α-amylolysis of starch-glycogen class polysaccharides is a purely hydrolytic process, although, in 1968, Greenwood and his associates (3) reported finding a small redistribution of radioactivity among the maltosaccharides formed when purified α-amylases of different sources were allowed to act on potato amylose in the presence of [^{14}C]glucose or [^{14}C]maltose. The authors were unable to distinguish whether this effect was the result of "transferase activity" or "condensation." The main conclusion was, in any case, that the observed redistribution was

[1] Aided by a grant from the Corn Refiners Association, Inc.

too small to have an appreciable effect on the action pattern, i.e., on the relative yields of specific maltosaccharide products, reported earlier to be characteristic of the individual α-amylases studied.

We now report what appears to be the first evidence that the digestion of starch or glycogen by α-amylases can lead, under certain circumstances, to products differing structurally from those obtainable by hydrolysis. Digests of α-glucans, carried out with certain aryl D-glucosides serving as glycosyl acceptors, have been found to yield novel glycosylation products in addition to the hemiacetal-bearing sugars and dextrins which, until now, have been the only products obtained from natural α-glucans by α-amylolysis. Such digests, moreover, have permitted the application of a direct method to quantitatively assess the capacity of α-amylases to utilize these glucosides as acceptors in comparison to water.

MATERIALS AND METHODS

Reaction components. Corn amylose, prepared by the pentasol method and crystallized four times, was a gift of the late Dr. Thomas J. Schoch; solutions were made in cold 0.5 N NaOH and neutralized immediately before use. Soluble potato starch of dextrose equivalent 0.5 was a commercial (Mallinckrodt) product; solutions were freshly prepared by briefly heating 1% suspensions in 0.05 M pH 5.6 acetate buffer in boiling water. Glycogenlike α-glucan, uniformly ^{14}C-labeled (0.76 μCi/mg), was synthesized *in vitro* by the action of purified *Neisseria perflava* amylosucrase on [^{14}C]sucrose; its preparation and properties have been described by Okada and Hehre (4). Uniformly labeled [^{14}C]maltose (ICN Corp.) was purified by paper chromatography and diluted to a specific activity of 0.33 μCi/mg with highly purified β-maltose (5). *p*-Nitrophenyl α-D-glucopyranoside was purchased from Sigma Chemical Co., St. Louis, MO; *p*-nitrophenyl β-D-glucopyranoside, $[\alpha]_D$ = 100°, was the gift of Dr. Gentaro Okada.

Liquifying α-amylase of *Bacillus subtilis* var. *amyloliquefaciens* Fukumoto was the three-times crystallized product (lot OP 23) from Seikagaku Kogyo Co., Tokyo; stock solutions made in cold distilled water were diluted with appropriate buffer before use. *Aspergillus oryzae* α-amylase, three-times crystallized and homogeneous on electrophoresis and ultracentrifugation, (lot DEO2-73145) was from Sankyo Co., Ltd., Tokyo; it was dissolved in 0.06 M pH 5.6 acetate buffer. Twice crystallized hog pancreatic α-amylase (lot AA9DC), from Worthington Biochemicals Corp., Freehold, NJ, was separated from the suspending medium by centrifugation and dissolved in 0.06 M pH 5.6 acetate. All three α-amylase preparations were free from contamination with α- or β-D-glucosidase as judged by their failure to cause detectable cleavage of maltose, *p*-NP α-D-glucoside, or *p*-NP β-D-glucoside on prolonged incubation.

Techniques. Chromatograms were made with Whatman No. 1 paper that had been specially washed to reduce the content of impurities that would interfere with spectrophotometric determinations of glycosidically linked *p*-nitrophenol in eluted products. Papers were individually irrigated with 2 *N* acetic acid for 2 days as in descending chromatography, dried in air at 25°C, then briefly soaked in water and redried. All chromatograms were developed for 18 h (descending) with *n*-butanol:pyridine:water (6:4:3). Saccharides were stained with $AgNO_3$ by a dipping technique (6). Paranitrophenyl glycosides were detected by the yellow staining afforded by the *p*-nitrophenyl group released by alkaline hydrolysis; i.e., dried chromatograms were dipped in a 2% solution of NaOH in ethanol, Placed (in a frame) in a preheated autoclave, and exposed to steam at atmospheric pressure (100°C) for 3 min. (Free *p*-nitrophenol can be distinguished by its yellow staining on simple alkalinization, without steaming.) Free *p*-nitrophenol in solution was determined using fluids free of particles, by measuring the absorption at 400 nm and pH 7.5 (ε = 12.1 × 10^3 cm^2/mmol (7)) in a Gilford Model 2000 spectrophotometer.

Density gradient centrifugation of *B. subtillis* α-amylase was carried out in a SW 50.1 swinging bucket rotor and Beckman-Spinco Model L2-65B centrifuge operated at 100,000 *g* and 4°C for 24 h. Gradient fractions (0.25 ml) were collected by extrusion with the aid of a micrometric syringe-buret (Model SB2, Micro-Metric Instrument Co., Cleveland, OH). Relative protein content of the fractions was determined by measuring optical density at 280 nm. Specific hydrolytic activity was assayed by mixing 0.50 ml of unfractionated enzyme or gradient fraction (diluted to contain 1.6 μg protein/ml) with 0.50 ml of 1% soluble starch in pH 5.6 acetate buffer (both attempered at 37°C); incubating, with appropriate controls, at 37°C for 3.0 min; determining as glucose the reducing sugars released (8, 9).

Radioactive compounds on chromatograms were located and measured with the aid of a 4π windowless radiochromatogram scanner (Baird-Atomic; Model 363) equipped with a digital integrator. For elution of individual radioactive compounds from chromatograms, the paper section (5–7 × 4 cm) containing the compound was placed on a piece of aluminum foil and moistened with ca. 0.2 ml water; paper and foil were then rolled together and centrifuged in a small graduated tube. This process was repeated three times, and the eluate volume adjusted to 1.0 ml. The content of glycosidically bound *p*-nitrophenol in each eluate was

determined from the absorbance at 3
5.6, with a small correction for the
material eluted from a control of blan
tions (pH 5.6) of pure *p*-NP β-D-glucos
provide the specific absorption. T
[^{14}C]glucose in each eluate was deter
suring the radioactivity of an aliquot in
in a scintillation counter (Beckman
cific activity of 2.31 × 10^6 cpm/μm
found by counting solutions contai
weight of the radioactive α-glucan, wa

EXPERIMENTAL

The Nonhydrolytic Degradatio by α-Amylase in the Pres p-Nitrophenyl α-D-Gluc

In designing a method suitab
mining the extent to which the
of starch by α-amylase might
from hydrolysis by relatively
trations of an added acceptor,
tention was given to the choic
and acceptor co-substrate. M
ments were carried out with
lized, α- and β-D-glucosid
subtillis liquifying α-amylase
other highly purified α-amyla
haps to a greater degree than
been shown (1–3) to be capable
lating the C_4-carbinol site of g
tose, α-D-glucosyl fluoride, an
fluoride. The latter acceptor
were not regarded as suitable
ing the type of quantitative
being sought. The need was fo
which, when present in an
amylopolysaccharide digest, w
transfer products that could ea
rated from and measured indep
the sugars and maltodextrins
produced by α-amylolysis;
acceptor that would not serv
substrate for the α-amylase.

The first attempt to find

[2] Crude preparations of *B.* s
amyloliquefaciens α-amylase have be
Pasur and Okada (10, 11) to contain a
4-α-glucanosyltransferase. Catalytic
utable to this transferase was not dete
in a *B. subtilis* α-amylase preparation
and crystallized by the method of
Okada (12). The present crystalline en
a commercial product, had been pr
Professor Fukumoto's supervision.

determined from the absorbance at 300 nm and pH 5.6, with a small correction for the absorbance of material eluted from a control of blank paper. Solutions (pH 5.6) of pure *p*-NP β-D-glucoside were used to provide the specific absorption. The content of [^{14}C]glucose in each eluate was determined by measuring the radioactivity of an aliquot in Bray's solution in a scintillation counter (Beckman LS-230); a specific activity of 2.31×10^5 cpm/μmol [^{14}C]glucose, found by counting solutions containing a known weight of the radioactive α-glucan, was used.

EXPERIMENTAL

The Nonhydrolytic Degradation of Starch by α-Amylase in the Presence of p-Nitrophenyl α-D-Glucoside

In designing a method suitable for determining the extent to which the breakdown of starch by α-amylase might be diverted from hydrolysis by relatively low concentrations of an added acceptor, special attention was given to the choice of enzyme and acceptor co-substrate. Most experiments were carried out with a recrystallized, α- and β-D-glucosidase-free *B. subtillis* liquifying α-amylase[2] which, like other highly purified α-amylases (and perhaps to a greater degree than most), had been shown (1–3) to be capable of glycosylating the C_4-carbinol site of glucose, maltose, α-D-glucosyl fluoride, and α-maltosyl fluoride. The latter acceptors, however, were not regarded as suitable for generating the type of quantitative information being sought. The need was for an acceptor which, when present in an α-amylase-amylopolysaccharide digest, would provide transfer products that could easily be separated from and measured independently of the sugars and maltodextrins ordinarily produced by α-amylolysis; also, for an acceptor that would not serve as a donor substrate for the α-amylase.

The first attempt to find an acceptor meeting these criteria was made with *p*-NP α-D-glucoside, a compound not previously known to be an α-amylase co-substrate. Phenyl α-D-glucoside, however, had been shown by Yoshida *et al.* (13) to serve as an acceptor for the unusual saccharifying α-amylase of *B. subtilis* var. *amylosaccharíticus* which caused glucosyl transfer from maltose to the phenyl α-D-glucoside, yielding phenyl α-maltoside; then hydrolysis of the latter to maltose and phenol (13).

When mixtures of corn amylose (4 mg/ml), *p*-NP α-D-glucoside (2.5, 5, and 10 mM), and *B. subtilis* var. *amyloliquefaciens* α-amylase (40 μg/ml) in pH 7.5 phosphate buffer, were incubated at 25°C, reactions analogous to that observed by Yoshida *et al.* (13) were obtained. Chromatograms of the mixtures (7-h digests) showed, in addition to the very rapidly migrating *p*-NP α-D-glucoside, the presence of two new *p*-NP-containing products migrating faster than glucose and other "normal" α-amylolytic products, plus a third that co-migrated with glucose. Each of the test mixtures also showed (Fig. 1) a progressive accumulation of free *p*-nitro-

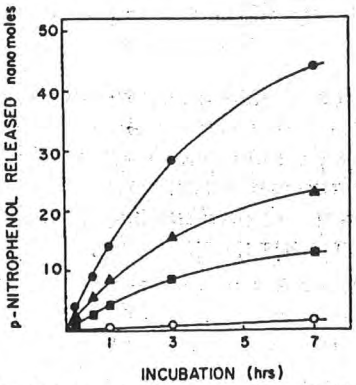


FIG. 1. Release of *p*-nitrophenol in test mixtures (1 ml, pH 7.5) containing corn amylose (4 mg), *B. subtilis* liquefying α-amylase (40 μg, 26 units), *p*-NP α-D-glucoside (10 μmol (●), 5 μmol (▲), or 2.5 μmol (■)), and phosphate (16 μmol). Incubation was at 25°C, concurrently with a control mixture (○) containing α-amylase, *p*-NP α-D-glucoside (10 μmol) and pH 7.5 buffer, but no amylose. (Results obtained with the latter were indistinguishable from those of a 10 mM solution of the glucoside in the pH 7.5 buffer.) Free *p*-nitrophenol was determined from the optical absorbancy at 400 nm (corrected for absorbancy at zero time) using $\epsilon = 12.1 \times 10^3$ cm^2/mmol (7).

[2] Crude preparations of *B. subtilis* var. *amyloliquefaciens* α-amylase have been found by Pasur and Okada (10, 11) to contain an accompanying 4-α-glucanosyltransferase. Catalytic activity attributable to this transferase was not detected, however, in a *B. subtilis* α-amylase preparation (10) purified and crystallized by the method of Fukumoto and Okada (12). The present crystalline enzyme, although a commercial product, had been prepared under Professor Fukumoto's supervision.

phenol with time, corresponding at 7 h to ca. 0.5% of the glucoside originally present. The release of *p*-nitrophenol was related to the concentration of the aryl glucoside in the digest, but it was in addition completely dependent on the presence of the amylose substrate. Control mixtures of α-amylase and *p*-NP α-D-glucoside, without amylose, yielded no free *p*-nitrophenol beyond the faint trace released from the glucoside held at 25°C in pH 7.5 buffer.

To clarify the relationship of α-amylase *per se* to these effects, a 2 mg sample of the crystalline enzyme (in 0.2 ml water) was layered on a gradient of 5–40% glycerol and centrifuged in a swinging bucket rotor at 100,000*g* for 24 h, as described by Pazur and Okada (11) for separating 4-α-glucanosyltransferase from α-amylase. Fractions of 0.25 ml were collected, diluted 5-fold, and examined for protein content (as optical density at 280 nm) and for catalytic activities.

As shown in Fig. 2, the ultracentrifuged α-amylase produced a single protein peak

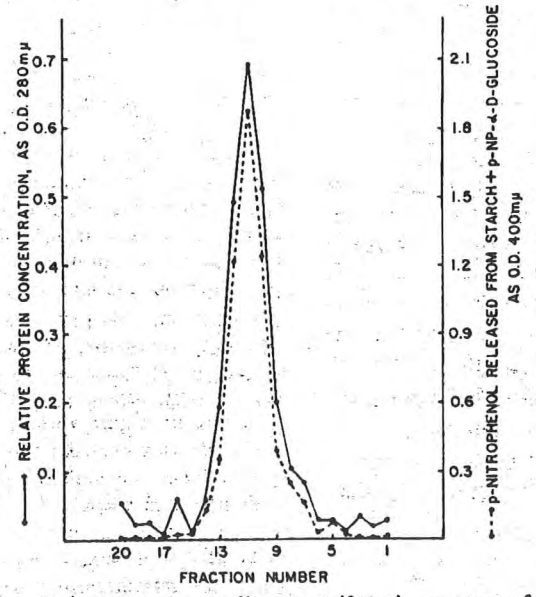


FIG. 2. Density gradient centrifugation pattern of crystalline *B. subtilis* var. *amyloliquefaciens* α-amylase. Relative protein concentration as optical density at 280 nm of fractions diluted 1:5, ——. *p*-Nitrophenol release by an aliquot of each fraction incubated (5 h, 25°C) with 4.5 mg/ml soluble starch and 10 mM *p*-NP α-D-glucoside, expressed as increase in optical density at 400 nm (centrifuged digests, pH 7.5), – – –.

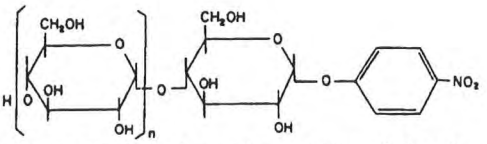


FIG. 3. Presumed generic structure of the glycosides found in the α-amylase-amylose digests containing *p*-NP α-D-glucoside. It appears likely from their migration rates that the products observed in chromatograms were *p*-NP α-maltoside, α-maltotrioside, and α-maltotetraoside, respectively, i.e., n = 1–3.

that was correlated with the capacity of fraction aliquots to cause *p*-nitrophenol release from a mixture of soluble starch and *p*-NP α-D-glucoside. The peak represented α-amylase of high specific activity, i.e., fractions no. 9, 11, and 13 were found to be 1.1, 1.3, and 1.1 times more active, respectively, than the original crystalline α-amylase (assayed concurrently), in terms of μmol of reducing sugar (as glucose) formed from soluble starch/min/mg protein (estimated from optical density). Moreover, the peak fractions behaved like the crystalline α-amylase preparation from which they were derived in catalyzing the formation of chromatographically demonstrable *p*-NP-containing products when tested in a system containing 40 μg/ml protein, 5 mg/ml soluble starch, and 40 mM *p*-NP α-D-glucoside (incubation: 5 h at 37°C, pH 5.6).

All of the observed results are consistent with the occurrence of α-amylase-catalyzed transfer of glycosyl groups from amylose or starch (and/or their hydrolytic products)[3] to the C_4-carbinol site of *p*-NP α-D-glucoside, forming a set of *p*-NP α-maltosaccharide glycosides of the type illustrated in Fig. 3. A low degree of susceptibility of the aryl glucosidic linkage of at least some members of such a series to

[3] The terms "hydrolytic products" and "sugars formed by hydrolysis" have been used to avoid circumlocution. Although the sugars and maltodextrins found in α-amylase digests of α-glucans are formed predominantly by hydrolysis, they do not arise exclusively in that way or necessarily represent segments as they preexisted in the substrate molecules. To some extent (a minor extent, according to Greenwood *et al.* (3)), their formation also involves a redistribution of glycosyl units from one saccharide molecule to another.

α-amylolytic cleavage is as responsible for the observed nitrophenol. α-Amylases fr sources other than *B. s am lolyticus* have been repor lyz *p*-NP α-maltoside (n = and *p*-nitrophenol (14–17).

Although it is evident th digests of amylose or soluble presence of *p*-NP α-D-glucos several higher glycosides by gl fer, such digests were not cor for determining the extent of tive to hydrolysis. The obse *p*-NP-containing glycoside r not fully separated from the se formed by hydrolysis; furth release of *p*-nitrophenol durin indicated that some degrad glycosides had occurred with less of these compounds and free sugars than had original duced.

p-Nitrophenyl β-D-Glucosi Acceptor for B. subtilis amyloliquefaciens α-Ar

Attempts to find a system shortcomings noted above with glucoside as acceptor led to the that *p*-NP β-D-glucoside is also

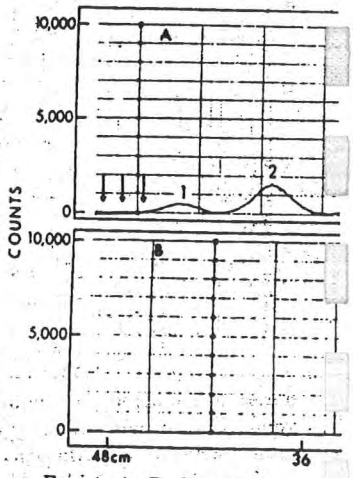


FIG. 4. A: Radioactive comp incubated mixture of [^{14}C] glyc The position of the latter (unlabele or compounds recovered from pe (0.135 ml) containing [^{14}C]glyc and incubated under conditions i

α-amylolytic cleavage is assumed to be responsible for the observed release of *p*-nitrophenol. α-Amylases from various sources other than *B. subtilis* var. *amylolyticus* have been reported to hydrolyze *p*-NP α-maltoside (n = 1) to maltose and *p*-nitrophenol (14–17).

Although it is evident that α-amylase digests of amylose or soluble starch, in the presence of *p*-NP α-D-glucoside, yielded several higher glycosides by glycosyl transfer, such digests were not considered ideal for determining the extent of transfer relative to hydrolysis. The observed set of *p*-NP-containing glycoside products was not fully separated from the series of sugars formed by hydrolysis; furthermore, the release of *p*-nitrophenol during incubation indicated that some degradation of the glycosides had occurred with time, leaving less of these compounds and more of the free sugars than had originally been produced.

p-Nitrophenyl β-D-Glucoside as an Acceptor for *B. subtilis* var. *amyloliquefaciens* α-Amylase

Attempts to find a system without the shortcomings noted above with *p*-NP α-D-glucoside as acceptor led to the observation that *p*-NP β-D-glucoside is also an efficient acceptor for *B. subtilis* α-amylase in the presence of amylopolysaccharide donors. A system of analysis was then devised to identify and measure the individual glycosides and sugars formed in such digests, so as to determine the capacity of the α-amylase to glycosylate the β-D-glucoside relative to its capacity to glycosylate water.

In a typical experiment, a test mixture (0.60 ml, pH 5.6) was prepared containing a ^{14}C-labeled glycogen-type polysaccharide (4) as donor substrate (2.43 mg; 15 μmol as glucose; 1.85 μCi); *p*-NP β-D-glucoside (24 μmol); *B. subtilis* var. *amyloliquefaciens* α-amylase (20 μg, 12.9 units); acetate (10 μmol). After incubation (37°C, 6 h) and inactivation (100°C, 15 min), the digest was subdivided and chromatographed: 0.15 ml samples were banded on each of three 4-cm-wide strips of washed Whatman no. 1 paper, and 30-μl samples were spotted (with glucose and *p*-NP β-D-glucoside standards) on each of two similar strips. As a first step, all chromatograms were analyzed for radioactive compounds using a radiochromatogram scanner and integrator.

Figure 4 illustrates the appearance of a radiochromatogram scan from one of the 0.15-ml samples, compared with the scan of a similarly prepared and treated control

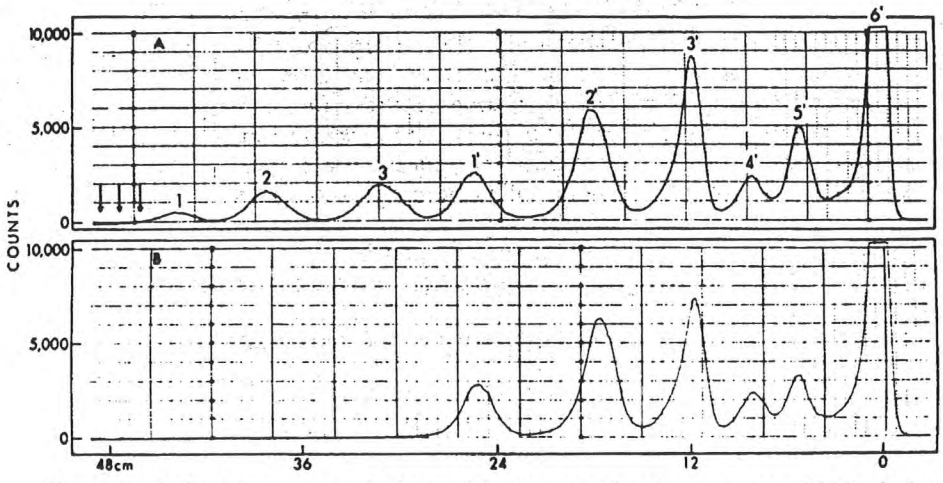


FIG. 4. A: Radioactive compounds detected in the scan of a chromatogram of 0.15 ml of an incubated mixture of [^{14}C] glycogen, *B. subtilis* liquefying α-amylase, and *p*-NP β-D-glucoside. The position of the latter (unlabeled) acceptor is indicated by arrows. See Table II for the identity or compounds recovered from peaks 1–3 and 1′–3′. B: Scan of a chromatogram of a control digest (0.135 ml) containing [^{14}C]glycogen and α-amylase, but without *p*-NP β-D-glucoside, prepared and incubated under conditions identical with the foregoing.

digest of α-amylase and α[^{14}C]glucan without *p*-NP β-D-glucoside. The amylase-glucan-glucoside mixture shows three well separated and fast-moving radioactive products (nos. 1–3) not present in the control, plus three other peaks (nos. 1′–3′) corresponding to the most rapidly migrating products (glucose, maltose, and maltotriose) in the control "ordinary hydrolysis" digest. Less well separated radioactive peaks, (nos. 4′–6′), near and at the origin, correspond in size and location to the tetra-, penta-, and various higher saccharides formed in the control digest of the glycogen by this α-amylase (18).

By staining the two chromatograms prepared with 30 μl of test digest, (a) for alkali-sensitive *p*-NP glycosides and (b) for reducing sugars with $AgNO_3$ (in each case after the radioactive peaks had been located by scanning), compounds 1–3 were recognized as transfer products containing *p*-NP β-D-glucoside as the acceptor; products 1′–6′, as reducing sugars and saccharides (Table I). It was inferred from the chromatographic mobilities (and later confirmed by analysis of the isolated compounds, described below) that the most rapidly migrating of the glycosides and sugars, 1 and 1′, contain a D-[^{14}C]glucosyl residue; 2 and 2′, a [^{14}C]maltosyl residue; 3 and 3′, a [^{14}C]maltotriosyl residue. The correspondence in the glycogen-derived (radioactive) part of each of the three pairs of products provides an unusual opportunity to directly compare the efficiencies of *p*-NP β-D-glucoside and water as acceptors for the α-amylase.

As shown in Table I, 11.4% of the [^{14}C]glucose originally present as glycogen in the chromatographed 0.15-ml digest sample was found associated with *p*-NP β-D-glucoside, as products 1–3; 30.4%, with sugars 1′–3′; 58.6% with higher saccharides. The relative yield of products containing a particular [^{14}C]glycosyl residue,

TABLE I

YIELDS OF GLYCOSIDES AND HYDROLYTIC PRODUCTS IN A *B. subtilis* var. *amyloliquefaciens* α-AMYLASE DIGEST OF [^{14}C]GLYCOGEN PLUS *p*-NITROPHENYL β-D-GLUCOSIDE

Product	Chromatographic mobility ($R_{glucose}$)	Assumed components of the product		Radioactivity[c] counts	Glucose from glycogen[d] (%)	Product formed; acceptor used[e] (μmol)
		Acceptor[a]	[^{14}C]Glycosyl[b]			
1	1.7	*p*-NP β-D-glucoside	Glucosyl	1,240	1.8	0.067
2	1.5	*p*-NP β-D-glucoside	Maltosyl	3,260	4.7	0.088
3	1.2	*p*-NP β-D-glucoside	Maltotriosyl	3,420	4.9	0.062
1′	1.0	H_2O	Glucosyl	3,540	5.1	0.191
2′	0.7	H_2O	Maltosyl	8,630	12.5	0.233
3′	0.5	H_2O	Maltotriosyl	8,870	12.8	0.160
4′	0.33	H_2O	Maltotetraosyl	2,610	3.8	0.035
5′	0.2	H_2O	Maltopentaosyl	4,770	6.9	0.052
6′	0.0	H_2O	Maltohexaosyl, heptaosyl, etc.	32,950	47.5	0.255
1–3		*p*-NP β-D-glucoside			11.4[f]	0.22[g]
1′–6′		H_2O			88.6	0.93

[a] *p*-NP β-D-glucoside is listed as acceptor for products giving a staining reaction for bound *p*-nitrophenol (yellow color on alkaline hydrolysis); water, for products staining rapidly with $AgNO_3$.

[b] Inferred from the chromatographic mobility; confirmed for products 1–3 and 1′–3′ by analysis of the isolated products (See Table II).

[c] Chromatogram of 0.15 ml digest (cf. Fig. 4A) containing 0.61 mg of [^{14}C]glycogen (3.75 μmol, as glucose).

[d] Counts × 100 ÷ total counts (69,290).

[e] Proportion of [^{14}C]glucose × 3.75 μmoles ÷ number of [^{14}C]glucose units in the product (assumed to average 7 in product 6′, located at the origin (18)).

[f] Corresponding sugars, 1′–3′, accounted for 30.4% of the glycogen.

[g] Molar ratio of *p*-NP β-D-glucoside/H_2O available in digest, 0.0007; ratio of use as acceptor by α-amylase, 0.24.

joined to *p*-NP β-D-glucoside
was 0.35 for 1/1′, involving D-[
0.38 for 2/2′, with [^{14}C]malto
3/3′ with [^{14}C]maltotriosyl. Th
high proportions considerin
molar concentration of water
1400 times that of the (40 mM
glucoside available as alternat

A further comparison of th
and water as acceptors for the
taking account of all digest
(and assuming the complete
p-NP-containing glycosides fr
1′–6′), is as follows. Under th
tions, approximately 30.5% of
lic bonds of the [^{14}C]glycoge
were cleaved, providing 1.15 μ
ous products. The glycosyl g
available by these cleavages w
cally linked with 0.22 μmol of
glucoside (forming compoun
with 0.93 μmol of water (form
1′–6′). Thus, the proportion of
side/water molecules utilized
was 0.24. Since the relative c
of glucoside/water molecules
the enzyme was only 0.0007
appear that *p*-NP β-D-glucosi
by the *B. subtilis* var. *amyl*
α-amylase several hundred tim
ciently than water.

It also appears likely that
1–3 owe their formation to th
glycosyl groups from the [
substrate and/or maltodextr
than to a process of condens
anomeric-OH of sugar prod
carbinol of the *p*-NP β-D-glu
radioactive glucoside was de
the [^{14}C]glycogen was replace
[^{14}C]maltose in a *p*-NP β-D-glu
taining control digest otherwis
incubated, and analyzed in t
as the test mixture.

Structures of the isolated pro
having localized and measure
activity of all components of
lase-glycogen-glucoside diges
matogram scanning, products 1
were individually eluted and c
(Table II). Compounds 1–3
rechromatography) were found
1.0, 2.0, and 3.0 mol of do

joined to *p*-NP β-D-glucoside or to water, was 0.35 for 1/1′, involving D-[^{14}C]glucosyl; 0.38 for 2/2′, with [^{14}C]maltosyl; 0.39 for 3/3′, with [^{14}C]maltotriosyl. These are very high proportions considering that the molar concentration of water was nearly 1400 times that of the (40 mM) *p*-NP β-D-glucoside available as alternate acceptor.

A further comparison of the glucoside and water as acceptors for the α-amylase, taking account of all digest components (and assuming the complete absence of *p*-NP-containing glycosides from products 1′–6′), is as follows. Under the test conditions, approximately 30.5% of the glycosylic bonds of the [^{14}C]glycogen substrate were cleaved, providing 1.15 μmol of various products. The glycosyl groups made available by these cleavages were enzymically linked with 0.22 μmol of *p*-NP β-D-glucoside (forming compounds 1–3) and with 0.93 μmol of water (forming products 1′–6′). Thus, the proportion of *p*-NP glucoside/water molecules utilized as acceptor was 0.24. Since the relative concentration of glucoside/water molecules surrounding the enzyme was only 0.0007, it would appear that *p*-NP β-D-glucoside was used by the *B. subtilis* var. *amyloliquefaciens* α-amylase several hundred times more efficiently than water.

It also appears likely that compounds 1–3 owe their formation to the transfer of glycosyl groups from the [^{14}C]glycogen substrate and/or maltodextrins, rather than to a process of condensation of the anomeric-OH of sugar products with a carbinol of the *p*-NP β-D-glucoside. No radioactive glucoside was detected when the [^{14}C]glycogen was replaced by (25 mM) [^{14}C]maltose in a *p*-NP β-D-glucoside-containing control digest otherwise prepared, incubated, and analyzed in the same way as the test mixture.

Structures of the isolated products. After having localized and measured the radioactivity of all components of the α-amylase-glycogen-glucoside digest by chromatogram scanning, products 1–3 and 1′–3′ were individually eluted and characterized (Table II). Compounds 1–3 (purified by rechromatography) were found to contain 1.0, 2.0, and 3.0 mol of donor-derived [^{14}C]glucose, respectively, per mol of *p*-NP β-D-glucoside. Moreover, the [^{14}C]glucose was present exclusively in α-1, 4-linked, pyranoid form. In each case, all of the radioactivity was released as glucose on digestion of the isolated glycoside with crystalline *Rhizopus niveus* glucoamylase. Also released by the treatment of each was a nonradioactive substance, identified as *p*-NP β-D-glucoside by its chromatographic mobility and its hydrolysis to glucose and *p*-nitrophenol by β-D-glucosidase. Products 1 to 3 are thus homologous *p*-NP β-glycosides containing two, three, and four glucose units, respectively; their structures, including the location of the [^{14}C]glucosyl, [^{14}C]maltosyl, and [^{14}C]maltotriosyl moieties derived from the [^{14}C]glycogen donor, are illustrated in Fig. 5. Sugars 1′–3′ were identified as [^{14}C]glucose, [^{14}C]maltose, and [^{14}C]maltotriose.

p-Nitrophenyl β-D-Glucoside as an Acceptor for Hog Pancreatic and Aspergillus oryzae α-Amylases

Results similar to those found with the *B. subtilis* enzyme, though quantitatively much less striking, were obtained with crystalline hog pancreatic and *A. oryzae* α-amylases (Table III). Mixtures of these α-amylases with α-[^{14}C]glucan and *p*-NP β-D-glucoside were prepared to contain the same concentrations of donor and acceptor as in the experiment with the bacterial enzyme; they were also incubated and examined under the same conditions.

The digest with pancreatic α-amylase produced two of the transfer products, nos. 1 and 2, which together accounted for 3.6% of the [^{14}C]glucose derived from the α-glucan donor. Formation of these glycosides involved the utilization of 0.10 μmol of *p*-NP β-D-glucoside as acceptor, compared with 1.10 μmol of water used in forming all hydrolytic products 1′–6′,[3] i.e., representing a ratio of 1/11. In the case of the digest with *A. orzae* α-amylase, only transfer product no. 1 [*p*-NP-α-D-[^{14}C]glucopyranosy (1 → 4) β-D-glucopyranoside] was found. Its synthesis represented the use of 0.043 μmol of the aryl glucoside acceptor compared with 1.26 μmol of water used in forming products 1′–6′, or a ratio of 1/30.

TABLE II

IDENTIFICATION OF GLYCOSIDES AND CORRESPONDING SUGARS FORMED BY α-AMYLASE FROM [^{14}C]GLYCOGEN IN THE PRESENCE OF *p*-NITROPHENYL β-D-GLUCOSIDE[a]

Product	Identity	[^{14}C] Glucose[b]	*p*-NP β-D-glucoside[c]	[^{14}C]Glucose / *p*-NP β-D glucoside[d]	Compounds freed on digestion with glucoamylase[e]
		nanomoles	*nanomoles*	*moles/mole*	
1	*p*-NP α-D-[^{14}C]glucopyranosyl-(1 → 4) β-D-glucopyranoside	70.5	67.8	1.04	[^{14}C]Glucose; *p*-NP β-D-glucoside
2	*p*-NP α-[^{14}C]maltosyl-(1 → 4) β-D-glucopyranoside	238	121.6	1.96	[^{14}C]Glucose; *p*-NP β-D-glucoside
3	*p*-NP α-[^{14}C]maltotriosyl-(1 → 4) β-D-glucopyranoside	317	106.4	2.98	Glucose-^{14}C; *p*-NP β-D-glucoside
1′	[^{14}C]Glucose	193	5.5[f]		
2′	[^{14}C]Maltose	416	<2.0[f]		
3′	[^{14}C]Maltotriose	492	<2.0[f]		

[a] Products 1–3 were recovered from three chromatograms made with 0.15 ml digest. The eluates of the individual products were combined and rechromatographed; tabulated data refer to the purified compounds. Products 1′–3′ were eluted from a single chromatogram and analyzed without purification.

[b] An aliquot of eluate was examined in a scintillation counter and the [^{14}C]glucose content calculated from the specific activity (2.31×10^8 cpm per μmol) of a sample of the [^{14}C]glucan substrate.

[c] Optical density of the eluate at 300 nm was determined and the *p*-NP β-D-glucoside content calculated from the absorbance of the authentic compound at pH 5.6.

[d] Product 1, as eluted from chromatograms of the digest, had ratios of <0.19, 0.70, and 0.91 due to contamination with the abundant *p*-NP β-D-glucoside acceptor which migrated slightly faster. Ratios for products 2 (1.97, 1.94, 1.94) and 3 (2.95, 3.04, 3.01) were not significantly changed by rechromatography.

[e] Each transfer product, in 0.1 ml of 20 mM pH 5.0 acetate buffer, was digested (2 h, 40°C) with 5 μg of crystallized *Rhizopus niveus* glucoamylase. Chromatograms in each showed one radioactive peak, migrating as glucose, plus unlabeled material migrating as *P*-NP β-D-glucoside (R_{glue} 1.9). The latter, on treatment with β-glucosidase, yielded compounds migrating and staining as glucose and *p*-nitrophenol.

[f] Products 1′–3′, indistinguishable from authentic glucose, maltose, and maltotriose in chromatographic mobility and $AgNO_3$ staining, appear contaminated with minute amounts of higher transfer products. That accompanying glucose (no. 1′) is presumably *p*-NP α-[^{14}C]maltotetraosyl-(1 → 4) β-D-glucopyranoside.

Though small, these ratios of organic acceptor/water usage by the pancreatic and mold α-amylase digests are vastly larger than the 1/1400 ratio of concentrations of glucoside/water molecules available to the enzymes. It would appear that these α-amylases, like that from *B. subtilis*, may utilize *p*-NP β-D-glucoside disproportionately (vs water) as an acceptor.

DISCUSSION

The present experiments reveal the capacity of α-amylases to partially convert starch and glycogen polysaccharides, in the presence of appropriate acceptor molecules, to products other than maltose, maltosaccharides, dextrins, and glucose. This previously unrecognized capacity, demonstrated with the aid of synthetic aryl D-glucosides as acceptors, raises the question of whether unusual and possibly metabolically significant products are formed as part of the natural *in vivo* α-amylolysis of starch or glycogen. It would appear worthwhile to search among cellular constituents for acceptors able to divert the α-amylolytic degradation of α-glucans in part to products other than hemiacetalbearing entities. The natural functions so far considered for α-amylases, *viz.*, release of maltosaccharides from starch and glycogen, and the possible origination of primer for starch and glycogen synthesis from α-maltose (19, 20), do not exhaust the glycosylative potentialities of these enzymes.

There is no reason to do
α-amylases themselves were re
formation of the present tran
The highly purified enzym
widely different biological sou
been prepared and crystallized
methods. Despite large diff
extent to which they yiel
based on *p*-NP β-D-glucoside
of conditions used, the thre
had in common the capacity
the C_4-carbinol of the accep
pletely negative results obt
[^{14}C]maltose was substitute
cogen as donor rule out the
contaminating α-D-glucosidas
enzymes catalyzing α-D-gluc
from maltose, were respon
observed glycosylations. M
age formation has been ob

YIELDS OF GLYCOSIDES AND HYDROLYT
[^{14}C](

Product	Chromatographic mobility ($R_{glucose}$)	Presun
1	1.7	*p*-NP α-D-[^{14}C]g β-D-glucosi
2	1.5	*p*-NP α-[^{14}C] β-D-glucosi
3	1.2	*p*-NP α-[^{14}C]ma β-D-glucosi
1′	1.0	[^{14}C]Glucose
2′	0.7	[^{14}C]Maltose
3′	0.5	[^{14}C]Maltotrios
4′	0.33	[^{14}C]Maltotet
5′	0.2	[^{14}C]Maltope
6′	0.0	[^{14}C]Maltohe
1–3		Glycosides
1′–6′		Hydrolytic pr

[a] Test mixtures (0.15 ml, pH 5.6 μmol); crystallized hog pancreatic β-a acetate (2.5 μmol). After incubatio chromatographed as a 4-cm band *p*-NP-containing glycosides.

[b] Control digests, in which the [^{14}C each enzyme and analyzed in the sam either control digest.

[c] Precisely localized bands staine

[d] Molar ratio of *p*-NP β-D-glucoside α-amylase, 0.094, and by mold α-amy

There is no reason to doubt that the α-amylases themselves were responsible for formation of the present transfer products. The highly purified enzymes were from widely different biological sources, and had been prepared and crystallized by different methods. Despite large differences in the extent to which they yielded products based on *p*-NP β-D-glucoside under the set of conditions used, the three α-amylases had in common the capacity to glycosylate the C_4-carbinol of the acceptor. The completely negative results obtained when [^{14}C]maltose was substituted for [^{14}C]glycogen as donor rule out the possibility that contaminating α-D-glucosidases, or other enzymes catalyzing α-D-glucosyl transfer from maltose, were responsible for the observed glycosylations. Maltosidic linkage formation has been observed with highly purified α-amylases from various other biological sources acting on substrates in low concentration (1, 3, 13) and, in our opinion, is an inherent capacity of this group of enzymes. On the other hand, several enzymes classed as 4-α-glucanotransferases (E.C. 2.4.1.25) (21) are also known to catalyze maltosidic linkage

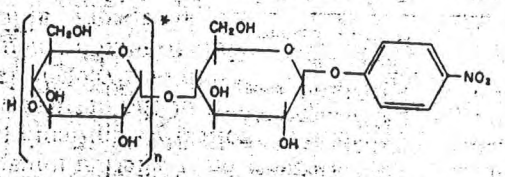


FIG. 5. Generic structure of the glycosides recovered from the α-amylase-[^{14}C]glycogen digest containing *p*-NP β-D-glucosides. Products no. 1-3 (n = 1-3) represent *p*-NP β-maltoside, β-maltotrioside, and β-maltotetraoside, respectively. * = [^{14}C]-glucose units derived from the glycogen.

TABLE III

YIELDS OF GLYCOSIDES AND HYDROLYTIC PRODUCTS IN HOG PANCREATIC AND *A. oryzae* α-AMYLASE DIGESTS OF [^{14}C]GLYCOGEN PLUS *p*-NP β-D-GLUCOSIDE[a, b]

Product	Chromatographic mobility ($R_{glucose}$)	Presumed structure	Pancreatic α-amylase		*A. oryzae* α-amylase	
			Radioactivity (counts)	Product formed; acceptor used (μmol)	Radioactivity (counts)	Product formed; acceptor used (μmol)
1	1.7	*p*-NP α-D-[^{14}C]glucosyl (1 → 4) β-D-glucoside	1,360[c]	0.073	800[c]	0.043
2	1.5	*p*-NP α-[^{14}C]maltosyl (1 → 4) β-D-glucoside	1,130[c]	0.030	0	0.000
3	1.2	*p*-NP α-[^{14}C]maltotriosyl (1 → 4) β-D-glucoside	0	0.000	0	0.000
1′	1.0	[^{14}C]Glucose	2,720	0.146	4,500	0.243
2′	0.7	[^{14}C]Maltose	12,400	0.335	25,400	0.685
3′	0.5	[^{14}C]Maltotriose	21,200	0.382	2,500	0.045
4′	0.33	[^{14}C]Maltotetraose	0	0.000	0	0.000
5′	0.2	[^{14}C]Maltopentaose	1,500	0.016	1,600	0.017
6′	0.0	[^{14}C]Maltohexaose, etc.	29,200	0.225	34,600	0.267
1-3		Glycosides		0.10[d]		0.043[d]
1′-6′		Hydrolytic products		1.10		1.26

[a] Test mixtures (0.15 ml, pH 5.6) contained [^{14}C]glycogen (3.75 μmol, as glucose) [b]; *p*-NP β-D-glucoside (6 μmol); crystallized hog pancreatic β-amylase (8.5 μg, 5 units) or crystallized *A. oryzae* α-amylase (8 μg, 1 unit); acetate (2.5 μmol). After incubation (37°C, 6 h) and inactivation (100°C, 15 min), each digest was chromatographed as a 4-cm band. The chromatograms were scanned for radioactivity, then stained for *p*-NP-containing glycosides.

[b] Control digests, in which the [^{14}C]glycogen was replaced by [^{14}C]maltose (3.75 μmol), were prepared for each enzyme and analyzed in the same way as the test mixture. No radioactive *p*-NP glycoside was detected in either control digest.

[c] Precisely localized bands stained positively for bound *p*-nitrophenol.

[d] Molar ratio of *p*-NP β-D-glucoside/H_2O available in digests, 0.0007; ratio of use as acceptor by pancreatic α-amylase, 0.094, and by mold α-amylase, 0.034.

synthesis from amylopolysaccharides or from maltosaccharides other than maltose. These enzymes are physically separable from the α-amylases with which they are associated in extracts from various plants (22–25), rabbit muscle (26), or *B. subtilis* (*B. amyloliquefaciens*) (11). As crude extracts may contain such enzymes, it is clear that the capacity of α-amylases to form maltosidic linkages can only be established through studies made with highly purified enzyme preparations. It is equally clear that such capacity may escape recognition unless appropriately specific conditions are applied. For example, the liquefying α-amylase ("endo-α-1,4-glucan hydrolase") of *B. subtilis* has been reported not to effect transfer reactions (11), based on a negative result with a crystalline preparation tested under conditions (10) allowing the observation of maltosyl transfer, from one maltopentaose molecule to another, by crude *B. subtilis* α-amylase preparations containing glucanosyltransferase. Present as well as earlier (1–3) findings show that, with other substrates and under other reaction conditions, maltosidic linkage formation is catalyzed by *B. subtilis* var. *amyloliquefaciens* α-amylase, as well as by hog pancreatic and *A. oryzae* α-amylases.

The limited range of experimental conditions and analyses employed in the present study precludes identification of the exact reaction steps leading to the observed products, especially as all digests were examined at a late stage of α-amylolysis. The first step in the formation of the glycosides of Tables I–III, for example, must have involved glycosylation of the C_4-carbinol of the *p*-NP β-D-glucoside acceptor. However, it is not clear what the size or sizes of these initially added [^{14}C]glycosyl groups might have been, or to what extent they were derived from the radioactive glycogen vs its cleavage products. Studies of the initial stages of conversion would be required to settle these questions. Also unknown is the extent to which the [^{14}C]glucose content of some or all of the initially formed glycosides may have been increased by additional transfers from maltodextrins; redistributed by transfers among the glycosides; diminished (or lost) by transfers of some or all of the α-1, 4-linked [^{14}C]glucose to water or to sugar carbinol sites. The negative results with the [^{14}C]maltose control digests suggest that condensations between hemiacetal-OH groups and C_4-carbinols had little, if any, part in the glycoside formation. The final result observed was that one or more radioactive glycosides (depending on the enzyme used) persisted into the late stage of amylolysis, side by side with the usual sugar and maltodextrin products. Of interest is that the pattern of yields of the latter sugars and dextrins relative to each other, found in digests of α-[^{14}C]glucan plus *p*-NP β-D-glucoside, was similar to the pattern of yields of these products in digests of the α-glucan alone. This is illustrated in the radiochromatogram scans (Fig. 4) of test and control digests made with *B. subtilis* α-amylase, but was also observed for test and control digests with hog pancreatic and *A. oryzae* α-amylases (scans not illustrated).

The use of low molecular weight compounds as substrates in the study of α-amylases has been criticized in the past (27) on the basis that such compounds do not completely occupy all of the binding sites available for large molecules and, hence, give only an incomplete picture of the enzyme-substrate interaction. Actually, the traditional type of study using high polymer α-glucan alone as "the substrate" has been a major factor in perpetuating the view that α-amylases are hydrolases when, in fact, a number of reports (1–3, 13, 28–30) have shown that these enzymes are catalysts of glycosyl-mobilization reactions in which water is only one of the possible co-substrates. The capacity to catalyze a range of glycosylation reactions does not deny that hydrolysis may be (and has usually been observed to be) predominant even in the presence of organic acceptors, as in the current experiments. The present work indicates that studies of α-glucan utilization, made in the presence of low molecular weight acceptor co-substrates other than hydrolytic products[3], can yield information on the catalytic capacities of α-amylases obtainable in no other way.

REFERENCES

1. HEHRE, E. J., GENGHOF, D. S., A (1971) *Arch. Biochem. Biophy*
2. HEHRE, E. J., OKADA, G., AND (1973) *Advan. Chem. Series* 117
3. GREENWOOD, C. T., MILNE, E. A., A (1968) *Arch. Biochem. Biophy*
4. OKADA, G., AND HEHRE, E. J. (197 249, 126–135.
5. HEHRE, E. J., OKADA, G., AND G (1969) *Arch. Biochem. Biophy*
6. TREVELYAN, W. E., PROCTER, D. P J. S. (1950) *Nature (London)* 16
7. WALLENFELS, K., MALHOTRA, O. P D., (1960) *Biochem. Z.* 333, 3
8. SOMOGYI, M. (1952) *J. Biol. Che*
9. NELSON, N. (1944) *J. Biol. Chem.*
10. PAZUR, J. H., AND OKADA, S. (1966) 241, 4146–4151.
11. PAZUR, J. H., AND OKADA, S. (196 243, 4732–4738.
12. FUKUMOTO, J., AND OKADA, S. (19 *Technol.* 41, 427.
13. YOSHIDA, H., HIROMI, K., AND *Biochem.* 66, 183–190.
14. JANSEN, A. P., AND WYDEVELD, P. *Nature (London)* 182, 525–52
15. MATSUBARA, S., IKENAKA, T., (1959) *J. Biochem.* 46, 425–431.
16. SUETSUGU, N., HIROMI, K., TAKAGI S. (1968) *J. Biochem.* 64, 619
17. SHIBAOKA, T., SUETSUGU, N.,

a...fers of some or all of the [^{14}C]glucose to water or to l ...tes. The negative results]n ...ltose control digests sug- ...sations between hemiace- ...nd C_4-carbinols had little, t... glycoside formation. The ...s ...ved was that one or more ycosides (depending on the ...persisted into the late stage ... le by side with the usual t... extrin products. Of inter- pattern of yields of the latter x... ...ns relative to each other, s... of α-[^{14}C]glucan plus *p*- ...side, was similar to the pat- ...f these products in digests of a ...ne. This is illustrated in m ...togram scans (Fig. 4) of trol digests made with *B.* ...ylase, but was also observed o... rol digests with hog pan- ...*oryzae* α-amylases (scans l).

...l ...v molecular weight com- b ...rates in the study of α- ...been criticized in the past ...sis that such compounds do y... ...ccupy all of the binding e... ...or large molecules and, ...ly an incomplete picture of ...ubstrate interaction. Ac- ...ional type of study using a... glucan alone as "the sub- ...een a major factor in per- ... ...iew that α-amylases are h... , in fact, a number of 13, 28–30) have shown that ...s are catalysts of glycosyl r... ...ctions in which water is e... ...ossible co-substrates. The ...talyze a range of glycosyla- ...does not deny that hydrol- n ...has usually been observed i...ant even in the presence of ...tors, as in the current experi- ...ent work indicates that u... ...n utilization, made in the w molecular weight a...eptor other than hydrolytic n...yield information on the c...es of α-amylases obtaina- r way.

REFERENCES

1. Hehre, E. J., Genghof, D. S., and Okada, G. (1971) *Arch. Biochem. Biophys.* **142**, 382–393.
2. Hehre, E. J., Okada, G., and Genghof, D. S. (1973) *Advan. Chem. Series* **117**, 309–333.
3. Greenwood, C. T., Milne, E. A., and Ross, G. R. (1968) *Arch. Biochem. Biophys.* **126**, 244–248.
4. Okada, G., and Hehre, E. J. (1974) *J. Biol. Chem.* **249**, 126–135.
5. Hehre, E. J., Okada, G., and Genghof, D. S. (1969) *Arch. Biochem. Biophys.* **135**, 75–89.
6. Trevelyan, W. E., Procter, D. P., and Harrison, J. S. (1950) *Nature (London)* **166**, 444–445.
7. Wallenfels, K., Malhotra, O. P., and Dabich, D. (1960) *Biochem. Z.* **333**, 377–394.
8. Somogyi, M. (1952) *J. Biol. Chem.* **195**, 19–23.
9. Nelson, N. (1944) *J. Biol. Chem.* **153**, 375–380.
10. Pazur, J. H., and Okada, S. (1966) *J. Biol. Chem.* **241**, 4146–4151.
11. Pazur, J. H., and Okada, S. (1968) *J. Biol. Chem.* **243**, 4732–4738.
12. Fukumoto, J., and Okada, S. (1963) *J. Ferment. Technol.* **41**, 427.
13. Yoshida, H., Hiromi, K., and Ono, S. (1969) *J. Biochem.* **66**, 183–190.
14. Jansen, A. P., and Wydeveld, P. G. A. B. (1958) *Nature (London)* **182**, 525–526.
15. Matsubara, S., Ikenaka, T., and Akabori, S. (1959) *J. Biochem.* **46**, 425–431.
16. Suetsugu, N., Hiromi, K., Takagi, M., and Ono, S. (1968) *J. Biochem.* **64**, 619–624.
17. Shibaoka, T., Suetsugu, N., Hiromi, K., and Ono, S. (1971) *Fed. Eur. Biochem. Soc. Lett.* **16**, 33–36.
18. Robyt, J., and French, D. (1963) *Arch. Biochem. Biophys.* **100**, 451–467.
19. Whelan, W. J. (1971) *Biochem. J.* **122**, 609–622.
20. Ryman, B. E., and Whelan, W. J. (1971) *Advan. Enzymol.* **34**, 285–443.
21. Commission on Biochemical Nomenclature (1972) Enzyme Nomenclature, pp. 142–143, Elsevier, Amsterdam.
22. Peat, S., Whelan, W. J., and Rees, W. R. (1956) *J. Chem. Soc.* 44–53.
23. Peat, S., Turvey, J. R., and Jones, G. (1959) *J. Chem. Soc.* 1540–1544.
24. Manners, D. J., and Rowe, K. L. (1969) *Carbohyd. Res.* **9**, 441–450.
25. Jones, G., and Whelan, W. J. (1969) *Carbohyd. Res.* **9**, 483–490.
26. Brown, D. H., and Illingworth, B. (1964) *in* Control of Glycogen Metabolism (W. J. Whelan and M. P. Cameron, eds.), p. 139–150, Churchill, London.
27. Fisher, E. H., and Stein, E. A. (1960) *in* The Enzymes (Boyer, P. D., Lardy, H., and Myrbäck, K., eds.), 2nd ed., Vol. 4, p. 334, Academic Press, New York.
28. Matsubara, S. (1961) *J. Biochem.* **49**, 226–231.
29. Okada, S., Kitahara, S., Higashihara, M., and Fukumoto, J. (1969) *Agr. Biol. Chem.* **33**, 900–906.
30. Robyt, J. F., and French, D. (1970) *J. Biol. Chem.* **245**, 3917–3927.

Proc. Natl. Acad. Sci. USA
Vol. 76, No. 9, pp. 4350–4354, September 1979
Biochemistry

Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications

(ribosomal proteins/radioimmunoassay/fluorescent antibody assay/peroxidase-conjugated antibody/autoradiography)

HARRY TOWBIN*, THEOPHIL STAEHELIN†, AND JULIAN GORDON*‡

*Friedrich Miescher-Institut, P. O. Box 273, CH-4002 Basel, Switzerland; and †Pharmaceutical Research Department, Hoffman-La Roche, CH-4002 Basel, Switzerland

Communicated by V. Prelog, June 12, 1979

ABSTRACT A method has been devised for the electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets. The method results in quantitative transfer of ribosomal proteins from gels containing urea. For sodium dodecyl sulfate gels, the original band pattern was obtained with no loss of resolution, but the transfer was not quantitative. The method allows detection of proteins by autoradiography and is simpler than conventional procedures. The immobilized proteins were detectable by immunological procedures. All additional binding capacity on the nitrocellulose was blocked with excess protein; then a specific antibody was bound and, finally, a second antibody directed against the first antibody. The second antibody was either radioactively labeled or conjugated to fluorescein or to peroxidase. The specific protein was then detected by either autoradiography, under UV light, or by the peroxidase reaction product, respectively. In the latter case, as little as 100 pg of protein was clearly detectable. It is anticipated that the procedure will be applicable to analysis of a wide variety of proteins with specific reactions or ligands.

Polyacrylamide gel electrophoresis has become a standard tool in every laboratory in which proteins are analyzed and purified. Most frequently, the amount and location of the protein are of interest and staining is then sufficient. However, it may also be important to correlate an activity of a protein with a particular band on the gel. Enzymatic and binding activities can sometimes be detected *in situ* by letting substrates or ligands diffuse into the gel (1, 2). In immunoelectrophoresis, the antigen is allowed to diffuse (3) or electrophoretically move (4) against antibody. A precipitate is then formed where the antigen and antibody interact. Modifications have been described in which the antigen is precipitated by directly soaking the separation matrix in antiserum (5, 6). The range of gel electrophoretic separation systems is limited by the pore size of the gels and diffusion of the antibody. The systems are also dependent on concentration and type of antigen or antibody to give a physically immobile aggregate.

Analysis of cloned DNA has been revolutionized (7) by the ability to fractionate the DNA electrophoretically in polyacrylamide/agarose gels first and then to obtain a faithful replica of the original gel pattern by blotting the DNA onto a sheet of nitrocellulose on which it is immobilized. The immobilized DNA can then be analyzed by *in situ* hybridization. The power of immobilized two-dimensional arrays has been extended to the analysis of proteins by use of antibody-coated plastic sheets to pick up the corresponding antigen from colonies on agar plates (8). Sharon *et al.* (9) have used antigen-coated nitrocellulose sheets to pick up antibodies secreted by hybridoma clones growing in agar.

In this report we describe a procedure for the transfer of proteins from a polyacrylamide gel to a sheet of nitrocellulose in such a way that a faithful replica of the original gel pattern is obtained. A wide variety of analytical procedures can be applied to the immobilized protein. Thus, the extreme versatility of nitrocellulose binding assays can be combined with high-resolution polyacrylamide gel electrophoresis. The procedure brings to the analysis of proteins the power that the Southern (7) technique has brought to the analysis of DNA.

MATERIALS AND METHODS

Immunogens and Immunization Procedures. *Escherichia coli* ribosomal proteins L7 and L12 were extracted (10) from 50S subunits and purified as described (11) by ion-exchange chromatography on carboxymethyl- and DEAE-cellulose. Antibodies were raised in a goat by injecting 250 μg of protein emulsified with complete Freund's adjuvant intracutaneously distributed over several sites. *Bacillus pertussis* vaccine (1.5 ml of Bordet–Gengou vaccine, Schweizerisches Serum- und Impfinstitut, Bern, Switzerland) was given subcutaneously with every antigen injection. Booster injections of the same formulation were given on days 38, 79, and 110. The animal was bled on day 117.

Subunits from chicken liver ribosomes (12) were combined in equimolar amounts, and 200-μg aliquots were emulsified with 125 μl of complete Freund's adjuvant injected at one intraperitoneal and four subcutaneous sites into BALB/c mice. Booster injections of 400 μg of ribosomes in saline were given intraperitoneally on days 33, 57, 58, and 59. The animals were bled on day 71.

Electrophoretic Blotting Procedures. Proteins were first subjected to electrophoresis in the presence of urea either in two dimensions (12) or in one-dimensional slab gels corresponding to the second dimension of the same two-dimensional system. The proteins were then transferred to nitrocellulose sheets as follows. The physical assembly used is shown diagrammatically in Fig. 1. A sheet of nitrocellulose (0.45 μm pore size in roll form, Millipore) was briefly wetted with water and laid on a scouring pad (Scotch-Brite) which was supported by a stiff plastic grid (disposable micropipette tray, Medical Laboratory Automation, Inc., New York). The gel to be blotted was put on the nitrocellulose sheet and care was taken to remove all air bubbles. A second pad and plastic grid were added and rubber bands were strung around all layers. The gel was thus firmly and evenly pressed against the nitrocellulose sheet. The assembly was put into an electrophoretic destaining chamber with the nitrocellulose sheet facing the cathode. The chamber contained 0.7% acetic acid. A voltage gradient of 6 V/cm was applied for 1 hr.

For polyacrylamide electrophoresis in the presence of sodium dodecyl sulfate (13) instead of urea, the procedure was as de-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

‡ To whom reprint requests should be addressed.

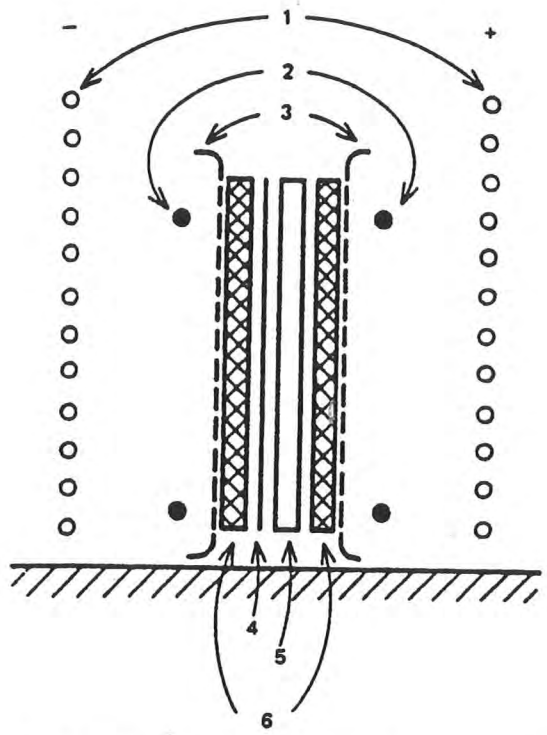


FIG. 1. Assembly for electrophoretic blotting procedure. 1, Electrodes of destainer; 2, elastic bands; 3, disposable pipette-tip tray; 4, nitrocellulose sheets; 5, polyacrylamide gel; 6, Scotch-Brite pads. Assembly parts are shown separated for visualization only.

scribed above except that the polarity of the electrodes was reversed and the electrode buffer was 25 mM Tris·192 mM glycine/20% (vol/vol) methanol at pH 8.3.

Staining for Protein. The blot may be stained with amido black (0.1% in 45% methanol/10% acetic acid) and destained with 90% methanol/2% acetic acid (see ref. 14).

Immunological Detection of Proteins on Nitrocellulose. The electrophoretic blots (usually not stained with amido black) were soaked in 3% bovine serum albumin in saline (0.9% NaCl/10 mM Tris·HCl, pH 7.4) for 1 hr at 40°C to saturate additional protein binding sites. They were rinsed in saline and incubated with antiserum appropriately diluted into 3% bovine serum albumin in saline also containing carrier serum with concentration and species as indicated in the legends. The sheets were washed in saline (about five changes during 30 min, total) and incubated with the second (indicator) antibody directed against the immunoglobulins of the first antiserum. As indicator antibodies we used ^{125}I-labeled sheep anti-mouse IgG. This had been purified with affinity chromatography on Sepharose-immobilized myeloma proteins and labeled by a modified version of the chloramine T method in 0.5 ml with 0.5 mg of IgG and 1 mCi of $Na^{125}I$ (1 Ci = 3.7×10^{10} becquerels) for 60 sec at room temperature. The specific activity was approximately 1.5 μCi/μg of IgG. ^{125}I-Labeled IgG was diluted to 10^6 cpm/ml in saline containing 3% bovine serum albumin and 10% goat serum, and 3 ml of this solution was used for a nitrocellulose sheet of 100 cm^2. Incubation was in the presence of 0.01% NaN_3 for 6 hr at room temperature. The electrophoretic blots were washed in saline (five changes during 30 min, total) and thoroughly dried with a hair dryer. The blots were exposed to Kodak X-Omat R film for 6 days.

Fluorescein- and horseradish peroxidase-conjugated rabbit anti-goat IgG (Nordic Laboratories, Tilburg, Netherlands) were reconstituted before use according to the manufacturer's instructions. Fluorescein-conjugated antibodies were used at 1:50 dilution in saline containing 3% bovine serum albumin and 10% rabbit serum. After incubation for 30 min at room temperature, the blots were washed as above and inspected or photographed with a Polaroid camera under long-wave UV light through a yellow filter.

Horseradish peroxidase-conjugated IgG preparations were used at 1:2000 dilution in saline containing 3% bovine serum albumin and 10% rabbit serum. The blots were incubated for 2 hr at room temperature and washed as described above. For the color reaction (15), the blots were soaked in a solution of 25 μg of *o*-dianisidine per ml/0.01% H_2O_2/10 mM Tris·HCl, pH 7.4. This was prepared freshly from stock solutions of 1% *o*-dianisidine (Fluka) in methanol and 0.30% H_2O_2. The reaction was terminated after 20–30 min by washing with water. The blots were dried between filter paper. Drying considerably reduced the background staining. The blots were stored protected from light.

RESULTS

Electrophoretic Transfer of Ribosomal Proteins from Polyacrylamide Gels to Nitrocellulose Sheets. Most proteins or complexes containing protein adsorb readily to nitrocellulose filters (16), whereas salts, many small molecules, and RNA are usually not retained. These binding properties are widely used for binding assays with nitrocellulose filters. We found that proteins were retained on these filters equally well when carried towards the filter in an electric field. If the electric field was perpendicular to a slab gel containing separated proteins (see Fig. 1), we obtained a replica of the protein pattern on the nitrocellulose sheet. This is demonstrated with ribosomal proteins from *E. coli*; a conventionally stained gel (Fig. 2*A*) and a stained electrophoretic blot of an identical gel (Fig. 2*B*) are shown. All ribosomal proteins from chicken liver and *E. coli* ribosomes detectable on two-dimensional gels could be seen on the electrophoretic blots produced from them. An example of a blot from a two-dimensional gel is given in Fig. 3. When the original polyacrylamide gel was stained after blotting, no protein could be detected. Thus, the blotting procedure removed all protein from the gel.

To establish whether the proteins removed from the gels were quantitatively deposited on the nitrocellulose sheet, we separated ^{3}H-labeled proteins from chicken liver 60S ribosomal subunit by two-dimensional electrophoresis and compared the radioactivity that could be recovered from the blot with that recovered directly from the gel (Table 1). Single proteins or groups of poorly separated proteins were cut out and radioactivity was measured after combustion of the samples. The results were within the variability inherent to two-dimensional analyses. Variations could be accounted for by variable transfer of proteins into the second dimension gel and the acuity with which spots can be cut out.

At loads exceeding the capacity of nitrocellulose, losses of protein occurred. Titration with radioactive ribosomal proteins under blotting conditions showed that at concentrations below 0.15 μg/mm^2 all protein was adsorbed. Overloading became apparent when a second sheet of nitrocellulose directly underneath the first one took up protein or when protein became visible on the cathodal surface of amido black-stained blots.

The conservation of resolution together with the high recovery of ribosomal proteins simplifies the procedure for autoradiography. The common procedure involving drying of polyacrylamide gels under heat and reduced pressure (19), which is tedious and time consuming, may be eliminated. Because the proteins become concentrated on a very thin layer, autoradiography from ^{14}C- and ^{35}S-labeled proteins should be highly efficient even without 2,5-diphenyloxazole impregnation

FIG. 2. Electrophoretic blotting of ribosomal proteins from one-dimensional gels. Total ribosomal proteins from *E. coli* were separated on an 18% polyacrylamide slab gel containing 8 M urea. (*A*) A section of the gel was stained with Coomassie blue; (*B*) another section was electrophoretically blotted and the blot was stained as described in *Materials and Methods*. Electrophoresis was from left to right.

(19). We have successfully obtained such autoradiograms from gels of ^{35}S-labeled proteins (not shown). Further, preliminary experiments with tritiated proteins have shown that dried blots may be processed for fluorography by brief soaking in 10% diphenyloxazole in ether (20).

The above experiments were done with ribosomal proteins separated on polyacrylamide gels containing urea. We have electrophoretically blotted proteins from sodium dodecyl sulfate by the modified procedure also described in *Materials and Methods*. Again, there was no loss of resolution. However, differences of staining intensities between proteins on the gel and the blot were apparent. In spite of the apparently incomplete recovery, blots from polyacrylamide gels containing sodium dodecyl sulfate may be used for detection of antigen in the same way as described below for ribosomal proteins (unpublished experiments).

Detection of Antigen by Antibody Binding on Blots *In Situ*. We found that proteins transferred to nitrocellulose sheets remained there without being exchanged over several days. Because a blot could be saturated with bovine serum albumin to block the residual binding capacity of the sheet, it can be treated as a solid-phase immunoassay. In the following immunological applications, we used indirect techniques throughout. Thus, antibody bound by the immobilized antigen was detected by a second, labeled antibody directed against the first antibody, and in each case excess unbound antibody was washed out.

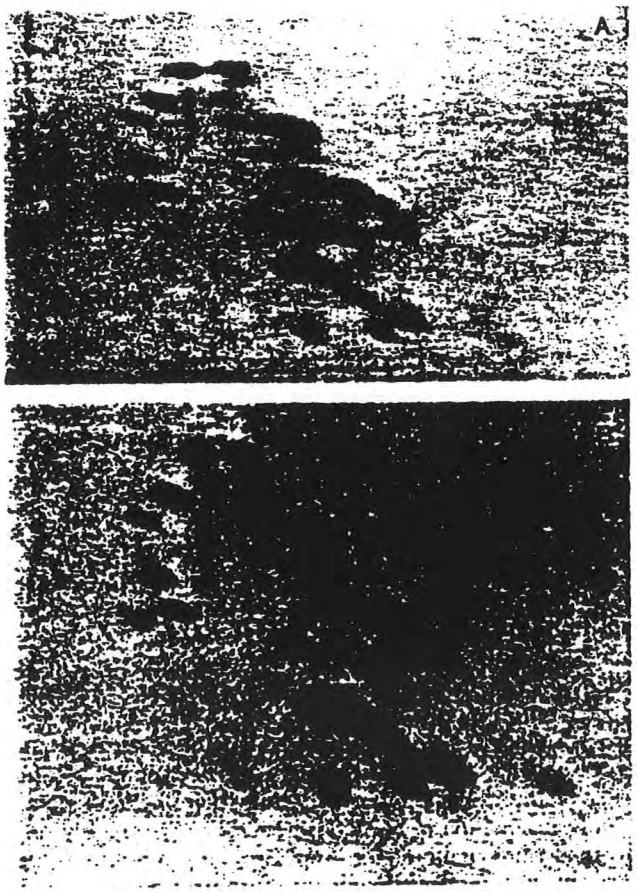


FIG. 3. Electrophoretic blotting of ribosomal proteins from two-dimensional gels. Proteins (35 μg) extracted from the 60S ribosomal subunit of chicken liver (12) were separated by two-dimensional gel electrophoresis. (*A*) Gel stained with Coomassie blue; (*B*) blot of an identical gel. Electrophoresis: 1st dimension, from left to right (towards cathode); 2nd dimension, top to bottom.

Table 1. Efficiency of transfer of ribosomal proteins to nitrocellulose sheets

Protein or group of proteins analyzed	Recovery on blot, %
3	123
4, 4A	104
5	111
6	107
7, 8	86
9	80
10	112
11	79
12, 16	93
13	95
15, 15A, 18	125
17	115
19	139
21, 23	118
26	114
27	143
28, 29	69
31	117
33	131

Ribosomal large-subunit proteins from chicken liver were tritiated by reductive methylation (17) and separated by two-dimensional electrophoresis (12) in the presence of 35 μg of carrier protein. Two identical gels were run. One was stained; the other was electrophoretically blotted on a nitrocellulose sheet. Spots were identified according to our nomenclature for chicken ribosomes (12), which differs only in minor respects from that established for rat ribosomes (18). Corresponding spots or groups of spots were cut from the gel and the blot. Their radioactivity was determined after conversion to tritiated water in a sample oxidizer (Oxymat).

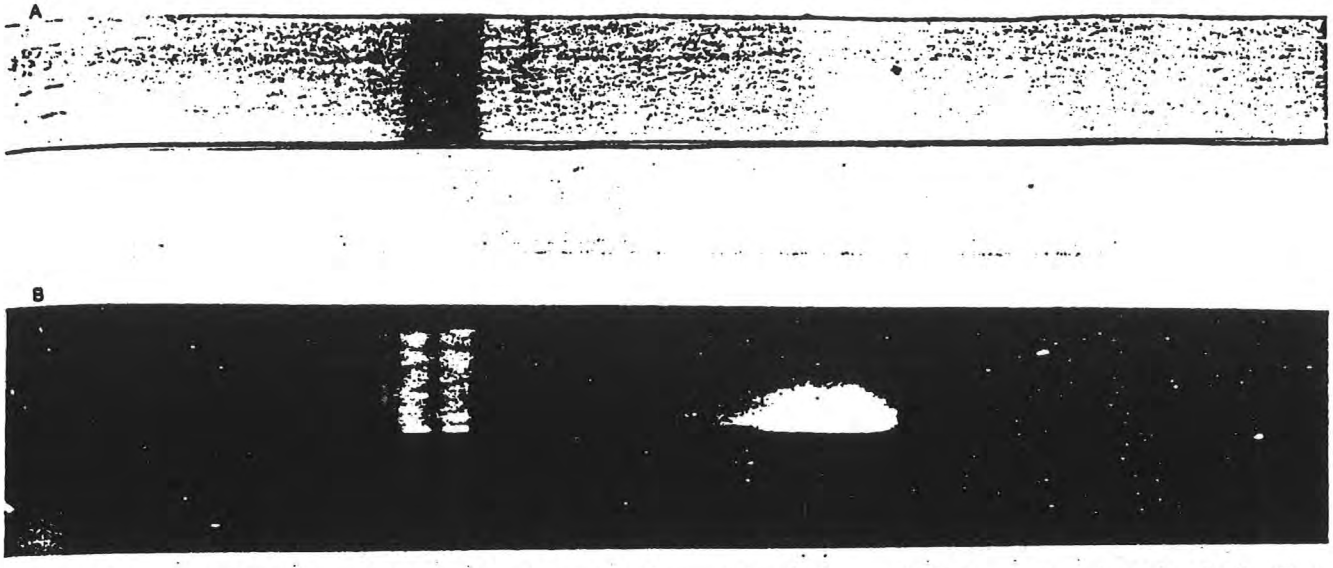


FIG. 4. Detection of *E. coli* ribosomal proteins L7 and L12 by (*A*) horseradish peroxidase- and (*B*) fluorescein-conjugated antibodies. Total ribosomal proteins from *E. coli* were separated and blotted as in Fig. 2. The anti-L7/L12 serum had a titer of 340 pmol of 70S ribosomes per ml of serum as determined by turbidity formation (20). Incubation was for 2 hr at room temperature in goat antiserum diluted 1:10 in saline containing 3% bovine serum albumin and 10% rabbit carrier serum and then with conjugated anti-goat IgG. In each case the lower strip is a control with preimmune antiserum. Electrophoresis was from left to right.

In Fig. 4 the detection of *E. coli* ribosomal proteins L7 and L12 with a goat serum specific for proteins L7 and L12 is shown. L7 is identical to L12, except for its *N*-acetylated NH_2-terminal amino acid (21). L7 and L12 fully crossreact immunologically (22) and are separated on acidic polyacrylamide gels (21). Both peroxidase- (Fig. 4*A*) and fluorescein-conjugated (Fig. 4*B*) antibodies were able to reveal immunoglobulin that was specifically retained by proteins L7 and L12. In each case, the lower gel is a control with preimmune serum. Peroxidase-conjugated antibodies were far more sensitive than fluorescein-conjugated ones. They could therefore be used at much higher dilution. This also permitted the detection of very small amounts of antigen. With a rabbit serum (23) we could detect 100 pg of L7 and L12 with serum and incubation conditions similar to those of the experiment described in Fig. 4 (not shown).

Because we can use the procedure to detect a specific antibody reacting with a specific protein after electrophoresis in polyacrylamide, we should also be able to determine which proteins have elicited antibodies in a complex mixture of immunogens. In the experiment of Fig. 5, individual sera of five mice immunized with chicken liver ribosomes were tested. We used ^{125}I-labeled sheep anti-mouse immunoglobulins to detect the presence of mouse immunoglobulins. In all mice, antibodies were preferentially produced against slowly moving proteins, presumably of high molecular weight. The procedure can thus characterize the antigen population against which specific antibodies have been raised in a mixture of immunogens.

DISCUSSION

The electrophoretic blotting technique described here produces replicas of proteins separated on polyacrylamide gels with high fidelity. We obtained quantitative transfer with proteins from gels containing urea. This was established here with ribosomal proteins. More generally, nitrocellulose membranes have been used to retain proteins from dilute solutions for their subsequent quantitative determination (16). Still, there remains the possibility that certain classes of protein do not bind to nitrocellulose. In this case absorbent sheets other than nitrocellulose or different blotting conditions may be helpful.

We have demonstrated that proteins immobilized on nitrocellulose sheets can be used to detect their respective antibodies.

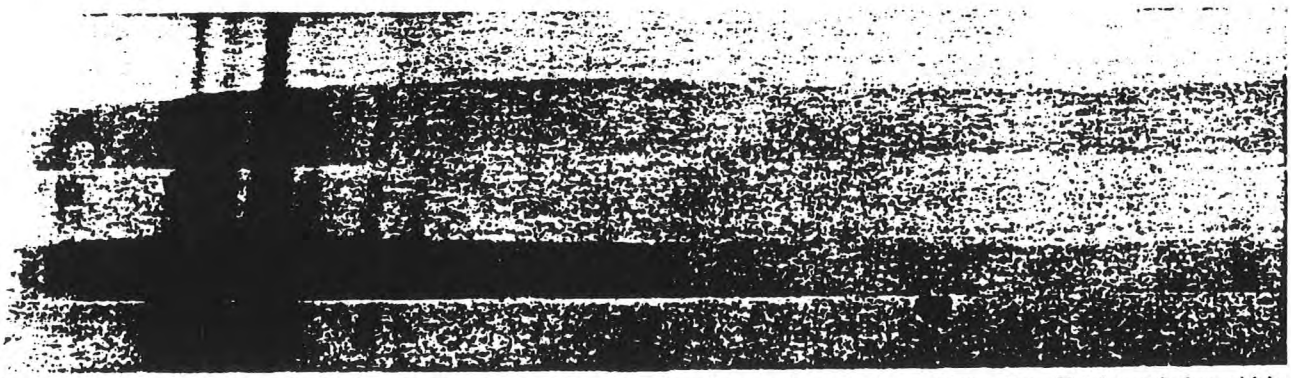


FIG. 5. Detection of immunoglobulin from individual mice directed against ribosomal proteins from chicken liver. Total protein from chicken liver ribosomes (12) was electrophoretically separated and blotted as in Fig. 2. Sera were obtained from five individual mice immunized against combined 40S and 60S subunits. The antisera were diluted 1:50 in saline containing 3% bovine serum albumin and 10% goat carrier serum. The blots were incubated in 250 μl of the diluted antiserum for 6 hr at room temperature. The blots were combined and treated with ^{125}I-labeled sheep anti-mouse IgG and autoradiographed. Electrophoresis was from left to right.

With radioactively labeled or peroxidase-conjugated antibodies the method is sensitive enough to detect small amounts of electrophoretically separated antigen, and this simple procedure can also be used to show the presence of small amounts of antibody in a serum of low titer. Because the antigen is immobilized on a sheet, the antibody is not required to form a precipitate with the antigen. The blotting technique therefore has the potential for immunoelectrophoretic analysis of proteins by using binding of Fab fragments or binding of antibodies against a single determinant, such as monoclonal antibodies produced by hybridomas (24). This could not be done by current immunoelectrophoretic techniques. If hybridoma clones are obtained from a mouse immunized with impure immunogen, it will be possible to use the technique to screen for clones making antibody directed against a desired antigen. Provided the desired antigen has a characteristic mobility in polyacrylamide gel electrophoresis, the appropriate clone can be selected without ever having pure antigen.

The procedure described here also has potential as a tool for screening pathological sera containing auto-antibodies—e.g., those against ribosomes (25–27). The precise identification of the immunogenic components may be a useful diagnostic tool for various pathological conditions.

A further advantage of immobilization of proteins on nitrocellulose is the ease of processing for autoradiography. Conventional staining, destaining, and drying of polyacrylamide gels takes many hours, and the exact drying conditions are extremely critical, especially for 18% gels as used in the second dimension for ribosomal proteins (12). When the proteins are transferred to a nitrocellulose support, as described here, the electrophoretic blotting takes 1 hr, staining and destaining less than 10 min, and drying an additional 5 min. This is thus both faster and simpler than conventional procedures, and it eliminates the tedious and hazardous procedure of soaking the gels in diphenyloxazole (19).

The technique has been developed to detect specific antisera against ribosomal proteins. However, it is applicable to any analytical procedure depending on formation of a protein–ligand complex. With the blotting technique, the usual procedure of forming a complex in solution and retaining it on a membrane would have to be reversed: the protein, already adsorbed to the membrane, would have to retain the ligand from a solution into which the membrane is immersed. Interactions that can possibly be analyzed in this way include hormone–receptor, cyclic AMP–receptor, and protein–nucleic acid interactions. The ligand may also be a protein. Enzymes separated on polyacrylamide gels could also be conveniently localized on blots by *in situ* assays. A critical requirement for these applications is that the protein is not damaged by the adsorption process and that binding sites remain accessible to ligands and substrates. In this respect, considerations similar to those in affinity chromatography and insoluble enzyme techniques pertain.

The method could also be adapted to the procedure of Cleveland *et al.* (28) for the analysis of proteins eluted from bands in polyacrylamide gels by one-dimensional fingerprints: one could label by iodination *in situ* on the nitrocellulose and then carry out the proteolytic digestion.

In preliminary experiments we have attempted to identify ribosomal RNA binding proteins by binding RNA to ribosomal proteins immobilized on nitrocellulose by the procedure of this paper, followed by staining for RNA (unpublished data), and have found a tendency for nonspecific binding. However, J. Steinberg, H. Weintraub, and U. K. Laemmli (personal communication) have independently developed a similar procedure for identifying DNA binding proteins.

We thank Drs. J. Schmidt and F. Dietrich for advice and help with immunization procedures and Mrs. M. Towbin for advice on setting up the peroxidase assay.

1. Gordon, A. H. (1971) in *Laboratory Techniques in Biochemistry and Molecular Biology*, eds. Work, T. S. & Work, E. (North-Holland, Amsterdam), p. 62.
2. Williamson, A. R. (1971) *Eur. J. Immunol.* 1, 390–394.
3. Grabar, P. & Williams, C. A. (1955) *Biochim. Biophys. Acta* 17, 67–74.
4. Laurell, C.-B. (1965) *Anal. Biochem.* 10, 358–361.
5. Zubke, W., Stadler, H., Ehrlich, R., Stöffler, G., Wittmann, H. G. & Apirion, D. (1977) *Mol. Gen. Genet.* 158, 129–139.
6. Showe, M. K., Isobe, E. & Onorato, L. (1970) *J. Mol. Biol.* 107, 55–69.
7. Southern, E. M. (1975) *J. Mol. Biol.* 98, 503–517.
8. Broome, S. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* 75, 2746–2749.
9. Sharon, J., Morrison, S. L. & Kabat, E. A. (1979) *Proc. Natl. Acad. Sci. USA* 76, 1420–1424.
10. Hamel, E., Koka, M. & Nakamoto, T. (1972) *J. Biol. Chem.* 247, 805–814.
11. Möller, W., Groene, A., Terhorst, C. & Amons, R. (1972) *Eur. J. Biochem.* 25, 5–12.
12. Ramjoué, H.-P. R. & Gordon, J. (1977) *J. Biol. Chem.* 252, 9065–9070.
13. Laemmli, U. K. (1970) *Nature (London)* 227, 680–685.
14. Schaffner, W. & Weissmann, C. (1973) *Anal. Biochem.* 56, 502–514.
15. Avrameas, S. & Guilbert, B. (1971 *Eur. J. Immunol.* 1, 394–396.
16. Kuno, H. & Kihara, H. K. (1967) *Nature (London)* 215, 974–975.
17. Moore, G. & Crichton, R. R. (1974) *Biochem. J.* 143, 604–612.
18. McConkey, E. H., Bielka, H., Gordon, J., Lastick, S. M., Lin, A., Ogata, K., Reboud, J.-P., Traugh, J. A., Traut, R. R., Warner, J. R., Welfle, H. & Wool, I. G. (1979) *Mol. Gen. Genet.* 169, 1–6.
19. Bonner, W. M. & Laskey, R. L. (1974) *Eur. J. Biochem.* 46, 83–88.
20. Randerath, K. (1970) *Anal. Biochem.* 34, 188–205.
21. Terhorst, C., Wittmann-Liebold, B. & Möller, W. (1972) *Eur. J. Biochem.* 25, 13–19.
22. Stöffler, G. & Wittmann, H. G. (1971) *J. Mol. Biol.* 62, 407–409.
23. Howard, G., Smith, R. L. & Gordon, J. (1976) *J. Mol. Biol.* 106, 623–637.
24. Köhler, G. & Milstein, C. (1976) *Eur. J. Immunol.* 6, 511–519.
25. Schur, P. H., Moroz, L. A. & Kunkel, H. G. (1967) *Immunochemistry* 4, 447–453.
26. Miyachi, K. & Tan, E. M. (1979) *Arthritis Rheum.* 22, 87–93.
27. Gerber, M. A., Shapiro, J. M., Smith, H., Jr., Lebewohl, O. & Schaffner, F. (1979) *Gastroenterology* 76, 139–143.
28. Cleveland, D. W., Fischer, S. G., Kirschner, M. W. & Laemmli, U. K. (1977) *J. Biol. Chem.* 252, 1102–1106.

With radioactively labeled or peroxidase-conjugated antibodies the method is sensitive enough to detect small amounts of electrophoretically separated antigen, and this simple procedure can also be used to show the presence of small amounts of antibody in a serum of low titer. Because the antigen is immobilized on a sheet, the antibody is not required to form a precipitate with the antigen. The blotting technique therefore has the potential for immunoelectrophoretic analysis of proteins by using binding of Fab fragments or binding of antibodies against a single determinant, such as monoclonal antibodies produced by hybridomas (24). This could not be done by current immunoelectrophoretic techniques. If hybridoma clones are obtained from a mouse immunized with impure immunogen, it will be possible to use the technique to screen for clones making antibody directed against a desired antigen. Provided the desired antigen has a characteristic mobility in polyacrylamide gel electrophoresis, the appropriate clone can be selected without ever having pure antigen.

The procedure described here also has potential as a tool for screening pathological sera containing auto-antibodies—e.g., those against ribosomes (25–27). The precise identification of the immunogenic components may be a useful diagnostic tool for various pathological conditions.

A further advantage of immobilization of proteins on nitrocellulose is the ease of processing for autoradiography. Conventional staining, destaining, and drying of polyacrylamide gels takes many hours, and the exact drying conditions are extremely critical, especially for 18% gels as used in the second dimension for ribosomal proteins (12). When the proteins are transferred to a nitrocellulose support, as described here, the electrophoretic blotting takes 1 hr, staining and destaining less than 10 min, and drying an additional 5 min. This is thus both faster and simpler than conventional procedures, and it eliminates the tedious and hazardous procedure of soaking the gels in diphenyloxazole (19).

The technique has been developed to detect specific antisera against ribosomal proteins. However, it is applicable to any analytical procedure depending on formation of a protein–ligand complex. With the blotting technique, the usual procedure of forming a complex in solution and retaining it on a membrane would have to be reversed: the protein, already adsorbed to the membrane, would have to retain the ligand from a solution into which the membrane is immersed. Interactions that can possibly be analyzed in this way include hormone–receptor, cyclic AMP–receptor, and protein–nucleic acid interactions. The ligand may also be a protein. Enzymes separated on polyacrylamide gels could also be conveniently localized on blots by *in situ* assays. A critical requirement for these applications is that the protein is not damaged by the adsorption process and that binding sites remain accessible to ligands and substrates. In this respect, considerations similar to those in affinity chromatography and insoluble enzyme techniques pertain.

The method could also be adapted to the procedure of Cleveland *et al.* (28) for the analysis of proteins eluted from bands in polyacrylamide gels by one-dimensional fingerprints: one could label by iodination *in situ* on the nitrocellulose and then carry out the proteolytic digestion.

In preliminary experiments we have attempted to identify ribosomal RNA binding proteins by binding RNA to ribosomal proteins immobilized on nitrocellulose by the procedure of this paper, followed by staining for RNA (unpublished data), and have found a tendency for nonspecific binding. However, J. Steinberg, H. Weintraub, and U. K. Laemmli (personal communication) have independently developed a similar procedure for identifying DNA binding proteins.

We thank Drs. J. Schmidt and F. Dietrich for advice and help with immunization procedures and Mrs. M. Towbin for advice on setting up the peroxidase assay.

1. Gordon, A. H. (1971) in *Laboratory Techniques in Biochemistry and Molecular Biology*, eds. Work, T. S. & Work, E. (North-Holland, Amsterdam), p. 62.
2. Williamson, A. R. (1971) *Eur. J. Immunol.* 1, 390–394.
3. Grabar, P. & Williams, C. A. (1955) *Biochim. Biophys. Acta* 17, 67–74.
4. Laurell, C.-B. (1965) *Anal. Biochem.* 10, 358–361.
5. Zubke, W., Stadler, H., Ehrlich, R., Stöffler, G., Wittmann, H. G. & Apirion, D. (1977) *Mol. Gen. Genet.* 158, 129–139.
6. Showe, M. K., Isobe, E. & Onorato, L. (1970) *J. Mol. Biol.* 107, 55–69.
7. Southern, E. M. (1975) *J. Mol. Biol.* 98, 503–517.
8. Broome, S. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* 75, 2746–2749.
9. Sharon, J., Morrison, S. L. & Kabat, E. A. (1979) *Proc. Natl. Acad. Sci. USA* 76, 1420–1424.
10. Hamel, E., Koka, M. & Nakamoto, T. (1972) *J. Biol. Chem.* 247, 805–814.
11. Möller, W., Groene, A., Terhorst, C. & Amons, R. (1972) *Eur. J. Biochem.* 25, 5–12.
12. Ramjoué, H.-P. R. & Gordon, J. (1977) *J. Biol. Chem.* 252, 9065–9070.
13. Laemmli, U. K. (1970) *Nature (London)* 227, 680–685.
14. Schaffner, W. & Weissmann, C. (1973) *Anal. Biochem.* 56, 502–514.
15. Avrameas, S. & Guilbert, B. (1971 *Eur. J. Immunol.* 1, 394–396.
16. Kuno, H. & Kihara, H. K. (1967) *Nature (London)* 215, 974–975.
17. Moore, G. & Crichton, R. R. (1974) *Biochem. J.* 143, 604–612.
18. McConkey, E. H., Bielka, H., Gordon, J., Lastick, S. M., Lin, A., Ogata, K., Reboud, J.-P., Traugh, J. A., Traut, R. R., Warner, J. R., Welfle, H. & Wool, I. G. (1979) *Mol. Gen. Genet.* 169, 1–6.
19. Bonner, W. M. & Laskey, R. L. (1974) *Eur. J. Biochem.* 46, 83–88.
20. Randerath, K. (1970) *Anal. Biochem.* 34, 188–205.
21. Terhorst, C., Wittmann-Liebold, B. & Möller, W. (1972) *Eur. J. Biochem.* 25, 13–19.
22. Stöffler, G. & Wittmann, H. G. (1971) *J. Mol. Biol.* 62, 407–409.
23. Howard, G., Smith, R. L. & Gordon, J. (1976) *J. Mol. Biol.* 106, 623–637.
24. Köhler, G. & Milstein, C. (1976) *Eur. J. Immunol.* 6, 511–519.
25. Schur, P. H., Moroz, L. A. & Kunkel, H. G. (1967) *Immunochemistry* 4, 447–453.
26. Miyachi, K. & Tan, E. M. (1979) *Arthritis Rheum.* 22, 87–93.
27. Gerber, M. A., Shapiro, J. M., Smith, H., Jr., Lebewohl, O. & Schaffner, F. (1979) *Gastroenterology* 76, 139–143.
28. Cleveland, D. W., Fischer, S. G., Kirschner, M. W. & Laemmli, U. K. (1977) *J. Biol. Chem.* 252, 1102–1106.

Appendix J: Report of [redacted], Pioneer, ELISA Lot 1568-022A and 1568-022B

Study No: MYCO98-001

Principal Investigator: [redacted]

Signature

11-2-98

Date

Extraction of Proteins from Plant Extract

A portion of each Cry1F Plant extract sample was weighed and then extracted as described in SOPs #PHIAR-SM-0001/01 with the exception that a 50mM carbonate/bicarbonate buffer was used. Total soluble protein in the corn tissue supernatants was measured by the method of Bradford (1976) using the microtiter plate application of the Bio-Rad Protein Assay according to SOP PHIAR-SP-0001/01. Bovine serum albumin (Sigma, St. Louis, MO) was used as the protein standard.

Analytical Methods

Cry1F ELISA. A direct double antibody sandwich enzyme-linked immunosorbent assay (ELISA) was developed and validated to quantitate levels of Cry1F protein in genetically modified maize. The method uses a polyclonal rabbit antibody specific to Cry1F protein to capture the protein in the microtiter well. The captured protein is detected by the same polyclonal antibody conjugated to biotin. The binding of the biotinylated antibody to the captured protein is detected by a conjugate of streptavidin-alkaline phosphatase (SA/AP). The enzyme substrate, pNPP, was added for the color development. Qunatitation of Cry1F protein was accomplished by extrapolation (based on sample absorbance (OD) value) from a Cry1F standard protein concentration curve. The Cry1F ELISA SOP # PHIAR-SM-0001/01, measures the level of protein, in pg/ug total protein in maize extracts.

Data Reduction

Cry1F protein concentrations expressed in pg/ug dry weight were calculated using JMP software (SAS Institute). Absorbance readings from the ELISAs and total soluble protein determinations were recorded using the Bio-Rad Model 3550 plate reader. The data was transferred to a JMP file where mean pg/ug and standard deviation calculations were made for each sample based on the concentration values from each sample analysis. Each plant extract sample was analyzed a minimum of three times over a three day time period. On the ELISA reaction plate, duplicate wells were used for each sample and for each sample total protein concentration assayed. If the duplicate wells/sample possessed a percent coefficient of variation (%CV) of greater than 15%, the sample analysis for that day was not used in the final concentration calculation/sample. If the %CV of duplicate wells was equal to or less than 15%, the mean Cry1F protein concentration was calculated for the ug total protein/well. The final mean concentration/sample was the mean determined from the 2-3 means (from 2-3 analyses). The associated standard deviation was also calculated for this mean.

$$\frac{\text{Cry1F pg}}{\text{Ug TP}} = \frac{\text{pg}}{\text{50ul}} \times \frac{\text{x}}{\text{1000ul}} \qquad \text{x} = \text{pg/ml}$$

To correct for concentration used in assay: $\frac{\text{pg/ml}}{\text{ug/ml (use equation above to convert ug/well to ug/ml)}} = \frac{\text{Cry1F pg}}{\text{ug TP}}$

$$\frac{\text{Bt pg}}{\text{Ug dwt}} = \frac{\text{Cry1F pg/ug}}{\text{ug TP}} \times \frac{\text{Bradford mg TP}}{\text{ml}} \times \frac{10^3 \text{ ug (units)}}{\text{mg}} \times \frac{\text{Weight \& Dilution of sample 0.6ml}}{\text{0.0155g}} \times 10 \text{ (correct for ml)} \times \frac{\text{g (units)}}{10^6 \text{ ug}}$$

Corn Samples Received 8/26/98

Mycogen

REVISED 9/14/98 to reflect correct pg/ug concentrations

Date Run Analyst	Extract #	Extract ID	Weight mg	Solutions mg/ml	Bradford mg/ml	Cry1F pg/ug
8/27/98	1	Bottle #1 Lot 1568-022A		5	4.498	neg
JCR	2	Bottle #2 Lot 1568-022B		5	4.11	1595
8/31/98	1	Bottle #1 Lot 1568-022A		5	4.24	neg
VAK	2	Bottle #2 Lot 1568-022B		5	4.1	2279.9
9/1/98	1	Bottle #1 Lot 1568-022A		5	4.37	neg
VAK	2	Bottle #2 Lot 1568-022B		5	4.43	1770.6

Samples run at 0.1ug/well (2ug/ml).

Cry1F results based on Bradford protein results.

Neg: Those samples which are designated as negative (0) refer to those samples yielding an interpolated concentration of <10pg/well for the Cry1F assay. The 10pg/well standard concentration is the limit of detection for the Cry1F ELISA.

Initial dilutions to make 10ug/well (200ug/ml) solution

then 1:10 dilutions made to run 1, 0.1, 0.01 ug/well

Date Run Analyst	Extract #	Extract ID	Solution ul	PBST ul
8/27/98	1	Bottle #1 Lot 1568-022A	22.2	477.8
JCR	2	Bottle #2 Lot 1568-022B	24.3	475.7
8/31/98	1	Bottle #1 Lot 1568-022A	24	476
VAK	2	Bottle #2 Lot 1568-022B	24	476
9/1/98	1	Bottle #1 Lot 1568-022A	23	477
VAK	2	Bottle #2 Lot 1568-022B	23	477

Appendix K: Report of [redacted] Mycogen, SDS PAGE of Lot 1599-39 and 1599-45

Quantitation

The Cry1F in the lyophilized protein powder was quantified by SDS-PAGE followed by scanning laser densitometry (1599-044). Lyophilized powder (10.0 mg) was weighed using a Mettler AM100 Microbalance and transferred to a glass 25 ml beaker. Distilled water (1.6 ml) was added and the powder allowed to dissolve completely. This solution was transferred quantitatively to a 2.0 ml volumetric tube and adjusted to volume. An aliquot (25 µl) of the protein solution was diluted 1:1 with Laemelli buffer (Bio-Rad Laboratories, Lot#82240,Product #161-0737) and boiled for 4 minutes. A protein standard was prepared from Bovine Serum Albumin (BSA) (Pierce Chemical Co., Lot #98011372, Product #23208) by diluting standard solutions (2000 µg/ml, 1500 µg/ml, 1000 µg/ml, 750 µg/ml, 500 µg/ml, 250 µg/ml, and 125 µg/ml) 1:10 in Laemelli buffer and boiling for 4 minutes. Each lane was loaded with a Hamilton 25 µl syringe. Samples (5-15µl) were run on 10% acrylamide pre-poured gels (Bio-Rad Laboratories, Lot#L100998B1, Product#161-1101) in Tris/glycine/SDS running buffer (BioRad Laboratories, Lot#61184A, Product #1610732). Gels were developed at a constant 200 volts for 1 hr without cooling. The gels were stained with 100 ml of a solution of Coomassie blue R-250 (BioRad, Lot#60851A, Product #1610436) and destained in Coomassie Destain Solution (Bio-Rad Laboratories, Lot#61057A, Product #1610438). The stained protein bands were digitized using a laser scanning densitometer (Molecular Dynamics) and quantified using ImageQuant software (ver 3.3, Molecular Dynamics). Sample bands were correlated to the BSA standards using standard linear regression analysis.

The results indicated that 10mg of lyophilized powder of batch 1599-39 yields 1.14mg of CRY1F toxin and 10mg of lyophilized powder of batch 1599-45 also yields 1.14 mg if CRY1F toxin. These results are posted in NB1599 p.47.

[redacted]

10·14·98

Appendix L: Report of [redacted], UCSD, N-Terminal

[redacted]
Phone:
Fax: [redacted]

UCSD Biology Department Protein Sequencer Lab

Fax

To: [redacted]	**From:** [redacted]
Fax: 455-7936	**Pages:** 48
Phone:	**Date:** 10/17/98
Re: Sequence Analyses	**CC:**

☐ **Urgent** ☐ **For Review** ☐ **Please Comment** ☐ **Please Reply** ☐ **Please Recycle**

● **Comments**

[redacted]

I am terribly sorry that I did not get the chromatograms to you in a more timely manner. Please accept my apologies.

I have reprinted the chromatograms this morning and have included them. Please let me know if there is any more that I can do.

	FC	PHB	P1
1	— — — — —	— —	— — a
2	— — — — —	— —	— — b
3	— — — — —		

"b" samples Not yet sequenced.

[redacted]

PHB - band a

	Major	Secondary
1	Xaa / (Arg, Phe, Ser, etc)	
2	Leu	Thr, Ala
3	Leu	Asp, Asn, Gly
4	Arg	Pro, Val
5	Leu	Glu, Phe, Ile
6	(Pro, Phe, Xaa)	⟷
7	Val/Leu	
8	Asp/Pro	
9	Gly/Ile	

For residues 2 - 5, there was a clear strongest signal. From residue 6 on, I could no longer tell which was the strongest signal.

P1 - band a

1 Ser, Phe
2 Thr, Leu
3 Gly, Leu
4 Ser, Arg
5 Glu, Leu
6 Pro, Phe
7 Val, Leu

Two samples sequencing at roughly equivalent amounts, so that it was not possible to identify one from the other

Band FC-1

Some background levels of amino acids are present. Slight rise in a couple of amino acids at first couple of steps, but they do not fall.

Conclude sample N-terminally blocked.

Band FC 2

Background levels of amino acids are present.
But no indications of sequencing.

Sample is N-Terminally blocked.

Band FC 3

Background levels of amino acids are lowers than previous samples

Suspect that sample is also N-terminally blocked. However because of lowers background, cannot be sure that there is too little sample present.

Appendix M: Report of [redacted], Ricerca, Biological Equivalency

Copy 2

Ricerca, Inc.

REPORT

Study Title

Bioequivalency of Bacterially Expressed Cry1F Protein with Maize Expressed Cry1F Protein

Ricerca, Inc. Document No. 7442-98-0079-AC-001

Author

[redacted]

Testing Facility
Ricerca, Inc.
7528 Auburn Road
P.O. Box 1000
Painesville, OH 44077

Project Identification
7442

Study Sponsor
Mycogen Corporation
5501 Oberlin Dr.
San Diego, CA 92121

Sponsor's Representative
Diane Shanahan

COMPLIANCE STATEMENT

The work reported in this report "Bioequivalency of Bacterially Expressed Cry1F Protein with Maize Expressed Cry1F Protein", Ricerca Document No. 7442-98-0079-AC-001, was conducted and reported in compliance with the Good Laboratory Practice Standards set forth in Title 40, part 160 of the Code of Federal Regulations of the United States of America.

The stability, homogeneity and chemical characterization of the test and control substances, per 40CFR160.105, were not determined by Ricerca, Inc. and were the responsibility of the Sponsor.

[redacted]

10/12/98
Date

Principal Investigator
Ricerca, Inc.

DISTRIBUTION

Document Number: 7442-98-0079-AC-001

Original
Ricerca Archivist

Copies
Mycogen /(1) bound, (2) unbound
Analytical and Biological Services/(3)
[redacted]

APPROVALS

Study Title:	**Bioequivalency of Bacterially Expressed Cry1F Protein with Maize Expressed Cry1F Protein**
Document Number:	**7442-98-0079-AC-001**
Performing Laboratory:	Ricerca, Inc. 7528 Auburn Road P.O. Box 1000 Painesville, OH 44077-1000

[redacted] 10/12/98
Date

Senior Research Entomologist
Analytical and Biological Services

[redacted] Oct. 12, 1998
Date

Section Head
Analytical and Biological Services

[redacted] Oct. 12, 1998
Date

Director
Analytical and Biological Services

QUALITY ASSURANCE STATEMENT

The Ricerca, Inc. Quality Assurance Unit has performed inspections on the study, " Bioequivalency of Bacterially Expressed Cry1F Protein with Maize Expressed Cry1F Protein". Ricerca Document Number 7442-98-0079-AC-001. Listed below are the dates inspections were conducted by Quality Assurance and the dates those findings were reported to the Study Director and to management.

Phase Inspected	Date(s) of Inspection	Dates Reported to Principal Investigator	Date(s) Reported to Ricerca Management	Date(s) Reported to Sponsor Study Director/Management
Bioassay Dilution Prep	9/4/98	9/9/98	10/9/98	9/18/98
Bioassay Diet Addition, Infestation	10/2/98	10/5/98	10/9/98	10/9/98
Data, Report	10/11&12/98	10/12/98	10/12/98	10/12/98

[redacted]
Quality Assurance Auditor

Oct. 12, 1998
Date

DEVIATIONS TO THE PROTOCOL:

None

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

INTRODUCTION

STUDY TITLE

Bioequivalency of Bacterially Expressed Cry1F Protein with Maize Expressed Cry1F Protein

SPONSOR

Mycogen
5501 Oberlin Drive
San Diego, CA. 92121

SPONSOR REPRESENTATIVE

[redacted]
Product Registration Manager
Mycogen
5501 Oberlin Drive
San Diego, CA. 92121

TESTING FACILITY

Ricerca, Inc.
Agricultural Research/Chemistry/Microbiology
7528 Auburn Road
P.O. Box 1000
Painesville OH 44077-1000

PRINCIPAL BIOLOGICAL INVESTIGATOR

[redacted]
Senior Research Entomologist
[redacted]

RICERCA STUDY NUMBER

7442-98-0079-AC

PURPOSE/OBJECTIVE

This study was designed to determine whether the spectrum of Cry1F endotoxin activity differs depending on the tissue or expression system in which it is produced. This study looked at activity from bacterially produced Cry1F and corn leaf tissue from Cry1F expressing plants and control plants.

JUSTIFICATION OF TEST SYSTEM

Ostrinia nubilalis (European corn borer), *Agrotis ipsilon* (Black cutworm), *Heliothis virescens* (Tobacco budworm), *Spodoptera frugiperda* (Fall armyworm) and *Helicoverpa zea* (Corn earworm) differ in their sensitivity to Cry1F endotoxin based on both field studies and laboratory bioassays. This study is designed to show that the active protein, Cry1F, exerts the same relative toxicity to these species regardless of whether it is produced in a bacterial expression system or in maize plants.

SCHEDULE OF EVENTS

Date of Experimental Bioassay Initiation: September 4, 1998
Date of Experimental Bioassay Termination: October 8, 1998
Bioassay Study Completion Date: October 13, 1998

MATERIALS AND METHODS

TEST PROCEDURE

Overview of Experimental Design

The experiment followed Mycogen Protocol MYCO98-001.

Test materials were weighed out, formulated and dispensed onto the surface of artificial diet. Neonate insect larvae were placed on the treated diet and mortality was assessed six days later. Using probit analysis comparisons were made to determine if the spectrum of activity (i.e. rank order of toxicity) differed between sources of the Cry1F toxin.

TEST ORGANISMS

The following insects were tested:

Name	Test Stage	Source
Agrotis ipsilon Black cutworm	Neonate larvae	French Agricultural Research, Lamberton, MN.
Helicoverpa zea Corn earworm	Neonate larvae	French Agricultural Research, Lamberton, MN.
Ostrinia nubilalis European corn borer	Neonate larvae	French Agricultural Research, Lamberton, MN.
Spodoptera frugiperda Fall armyworm	Neonate larvae	French Agricultural Research, Lamberton, MN.
Heliothis virescens Tobacco budworm	Neonate larvae	French Agricultural Research, Lamberton, MN.

REARING CONDITIONS

Eggs were purchased from commercial sources and shipped by overnight delivery service for next day arrival. Upon arrival at Ricerca, the eggs were placed in an incubator maintained at 10 ± 5°C until needed unless they were used immediately. To prepare neonate larvae, eggs were surface sterilized with a solution consisting of 4% bleach in deionized water containing 96 ppm Triton-X-155 and allowed to incubate in the Insectary at 26 ± 3°C until hatch.

DIET PREPARATION

Insect diet was prepared and dispensed into diet trays according to the standard procedures described in the protocol.

TEST PROCEDURE

Route of Administration and Reason for Choice

Test materials were administered using diet surface overlay techniques in flat-bottomed, 96-well polystyrene tissue culture plates (Becton-Dickinson Labware, Catalog No. 3070). Dilutions were made in 0.1% bovine serum albumin (BSA) in ph 7.2 phosphate buffer. Fifteen microliter aliquots of test

solutions were pipetted onto the surface of diet in each well and allowed to dry. For each dose, 48 wells were treated although sometimes larvae escaped or not enough larvae were available to fill all the wells. Once the test solution had dried, one neonate larva was added to each well. After the entire tray had been infested and covered, the tray was placed in a Forma brand environmental chamber set at 25 ± 4°C, and approximately 60 ± 10% RH. Mortality was assessed after six days. Larvae that were clearly dead or stunted were classified as dead.

Generally, three replications of each test substance/species combination were tested over time although some exceptions were made. If negligible mortality occurred (defined as less than 15%) in either of the first two runs, a third run was not carried out (i.e. for control plant assays).

Two types of controls were initially carried out for each bioassay. These were 1) an untreated control consisting of only phosphate buffer plus 0.1% BSA and 2) tissue blanks of lyophilized plant tissue from control plants. The plant blanks were formulated to contain the same amount of test material as was tested in the Plant Cry1F bioassays.

Preparation of the BT Inoculums

Test materials were weighed out following Standard Operating Procedures and diluted using 0.1% BSA (bovine serum albumin) in phosphate buffer, pH 7.2, as the diluent.

TEST SUBSTANCES

The following samples were tested:

Identity (Textual reference name)	Source	Sponsor's Identification	Ricerca ID(s)
Bacterial Cry1F	Mycogen	1599-39	TCR 198-00090
Plant Cry1F	Mycogen	1568-022B	TCR 198-00092
Control Plant Tissue	Mycogen	1568-022A	TCR 198-00091
Dipel	Locally Purchased	-	JK-0

For clarity, several of the names of the test materials have been shortened. These textual reference names are in the above table.

Physical and chemical characterizations of the test substance to comply with 40CFR160.135 were the responsibility of the Sponsor. The stability of these test and reference substances was not within the scope of this project.

STORAGE OF TEST MATERIALS

The test materials were stored prior to use under conditions specified by the Sponsor (room temperature and -20°C). In accordance with 40CFR160.105(d), since the experimental phase of this study lasted longer than 4 weeks, retain samples of test and control materials were kept.

WATER

Water was supplied by the city of Painesville, Ohio Municipal Water Company and deionized on-site. This water is periodically monitored for the presence of pesticides, heavy metals, PCB's, and solvents.

ENVIRONMENTAL CONDITIONS

Bioassays were carried out in LS-133. Bioassay trays were kept in a Forma brand environmental chamber maintained at 25 ± 4°C and ca. 60 ± 10 % RH. Temperature/Relative Humidity chart records are included in Appendix A.

STATISTICAL ANALYSIS

The data for each species/test material combination were pooled and subjected to probit analysis using POLO-PC (LeOra Publishing, Berkeley, CA) to calculate LC_{50}'s, slopes, and other statistics.

Due to the experimental nature of this experiment, some data was excluded for various reasons. Several of the initial bioassays were compromised by high control mortality and mortality which did not increase regularly as the dose increased. When this occurred, the data for that day's treatment/insect combination were not included in the probit analysis.

The protocol further stated that doses would be chosen to produce five rates giving between 0% and 100% mortality, with at least two causing less than 50% mortality, and at least two causing greater than 50% mortality. These conditions were not met for all bioassays.

The percent mortality at each dose was calculated. Tests with greater than 15% control mortality were considered invalid and the results from those tests were not used further.

RESULTS

The dose-response data for the three test materials (Bacterial Cry1F, Plant Cry1F and Dipel®) against all five species are presented in Tables 1 and 2. Table 3 contains the data for the bioassays of the plant control tissue. Appendix B contains the original data for all bioassays.

All treatment/species bioassays produced highly significant dose-response lines with significance greater than 95%. In most cases, the criteria established in the protocol for validity (number of rates, number above and below 50%, control mortality) were met. The two relatively insensitive species, Corn earworm and Black cutworm, were exceptions. For these two species the two highest test doses of the Plant Cry1F sample failed to produce greater than 50% mortality. Despite this limitation, probit analysis did generate LC_{50} values and other statistics.

Table 1: **Results of surface treatment bioassays of Bacterial Cry1F and Plant Cry1F against European corn borer (ECB), Fall armyworm (FAW), Tobacco budworm (TBW), Black cutworm (BCW) and Corn earworm (CEW). The LC_{50} units are ng Cry1F/cm^2.**

Test Material	Species	LC_{50} (95% C.L.)[a]	Slope	n	Significance of regression line
Bacterial Cry1F	ECB	0.58 (0.44 – 0.75)	1.1	982	99%
	TBW	1.88 (1.25 – 2.66)	1.3	1120	99%
	FAW	2.49 (1.29 – 4.10)	1.1	1467	99%
	CEW	51.6 (32.8 – 88.9)	0.9	1139	99%
	BCW	69.2 (35.9 – 128.6)	1.4	1169	95%
Plant Cry1F	ECB	0.58 (0.46 – 0.72)	1.7	921	99%
	TBW	1.74 (1.43 – 2.09)	1.5	1127	99%
	FAW	2.21 (1.34 – 3.22)	1.4	1135	99%
	CEW	>> 15.8 [b]	0.5	1095	99%
	BCW	22.7 (17.7 – 38.0)	2.6	1129	95%

[a] (95% C.L.) is 95% Confidence limits.

[b] Less than 50% mortality was observed at the highest rate tested, 15.8 ng, so LC_{50} is greater than 15.8. Computer generated confidence limits are not meaningful.

European corn borer was significantly more sensitive to both Bacterial Cry1F and Plant Cry1F than any of the other species based on non-overlapping 95% confidence limits. Tobacco budworm and Fall armyworm were slightly less sensitive than European corn borer to Bacterial Cry1F and Plant Cry1F but much more sensitive than both Black cutworm and Corn earworm. Black cutworm and Corn earworm were the least sensitive species from this group.

Bioassays were carried out with Dipel® in order to show the validity of the bioassay system (Table 2). Against all five insect species, highly significant dose-response lines were generated. Black cutworm was the least sensitive species. European corn borer, Corn earworm and Tobacco budworm were all equally sensitive.

Table 2: Results of surface treatment bioassays of Dipel® against European corn borer (ECB), Fall armyworm (FAW), Tobacco budworm (TBW), Black cutworm (BCW) and Corn earworm (CEW). The LC_{50} units are ppm.

Species	LC_{50} (95% C.L.)[a]	Slope	N	Significance of regression line
TBW	0.70 (0.48 – 0.96)	1.61	837	99%
CEW	0.75 (0.45 – 1.15)	1.09	851	99%
ECB	1.03 (0.66 – 1.50)	1.53	788	99%
FAW	10.5 (7.9 – 13.9)	1.11	859	99%
BCW	189 (129 – 336)	0.99	1126	99%

Table 3 contains the results from bioassays of the control plant extract. Multiple doses of the extract were tested since the rate structure was modified throughout the experiment in response to results. The dose was equal to the amount of total plant material from the bioassays with Plant Cry1F.

The three relatively sensitive species, European corn borer, Fall armyworm and Tobacco budworm, showed very little mortality from the control plant extract. Mortality due to the control plant extract was generally less than twice the buffer control mortality. A notable exception occurred in Assay 1 for TBW. The results of this assay are suspect since control mortality was 51.1%. For Fall armyworm, mortality was generally under 10% except for the first assay in which it reached 37.5%. Mortality of the Corn earworm was high in the initial tests but low in subsequent tests. Three out of the four assays with Black cutworm showed greater than twice as much mortality as the concurrent buffer control. Although it is possible that this effect contributed to the mortality observed in the Plant Cry1F bioassays, separating the two effects is beyond the scope of this project. In any event, the control plant mortality results do not alter the basic conclusions about the relative sensitivities of these insects to these materials.

Table 3: Results of surface treatment bioassays of Control Plant extract against European corn borer (ECB), Fall armyworm (FAW), Tobacco budworm (TBW), Black cutworm (BCW) and Corn earworm (CEW) larvae. The dose units are ng plant extract powder/cm^2. Individual assay results are shown since the dose rate was changed over time.

Species	Dose (ng Control Plant extract/cm2)	Assay No. 1	Assay No. 2	Assay No. 3	Assay No. 4	Assay No. 5
BCW	0	15.2	6.4	7.4	4.2	-
	5000	29.2	-	-	-	-
	10000	-	-	25.0	4.3	-
	20000	-	44.7	-	-	-
CEW	0	1.1	7.5	0.0	7.3	-
	5000	22.2	-	-	-	-
	10000	-	39.6	4.3	8.3	-
ECB	0	2.3	6.4	4.4	7.7	-
	500	6.3	-	-	-	-
	5000	-	7.3	2.2	12.8	-
FAW	0	10.9	8.3	4.2	9.4	-
	5000	37.5	2.1	4.2	4.7	-
TBW	0	51.1	10.4	4.2	8.4	2.1
	5000	34.8	16.7	6.3	2.2	2.1

DISCUSSION

Based on LC_{50} values, the rank order of sensitivity to Bacterial Cry1F appears to be:

ECB > TBW = FAW >>CEW > BCW

Based on LC_{50} values, the rank order of sensitivity to Plant Cry1F appears to be:

ECB > TBW = FAW >>CEW > BCW

The rank order of sensitivity appears to be same regardless of the source of the Cry1F protein.

Names of Key Individuals Involved in Study

Name	Title
[redacted]	Senior Research Entomologist
	Research Entomologist I
	Research Biologist

Disposition of the Raw Data

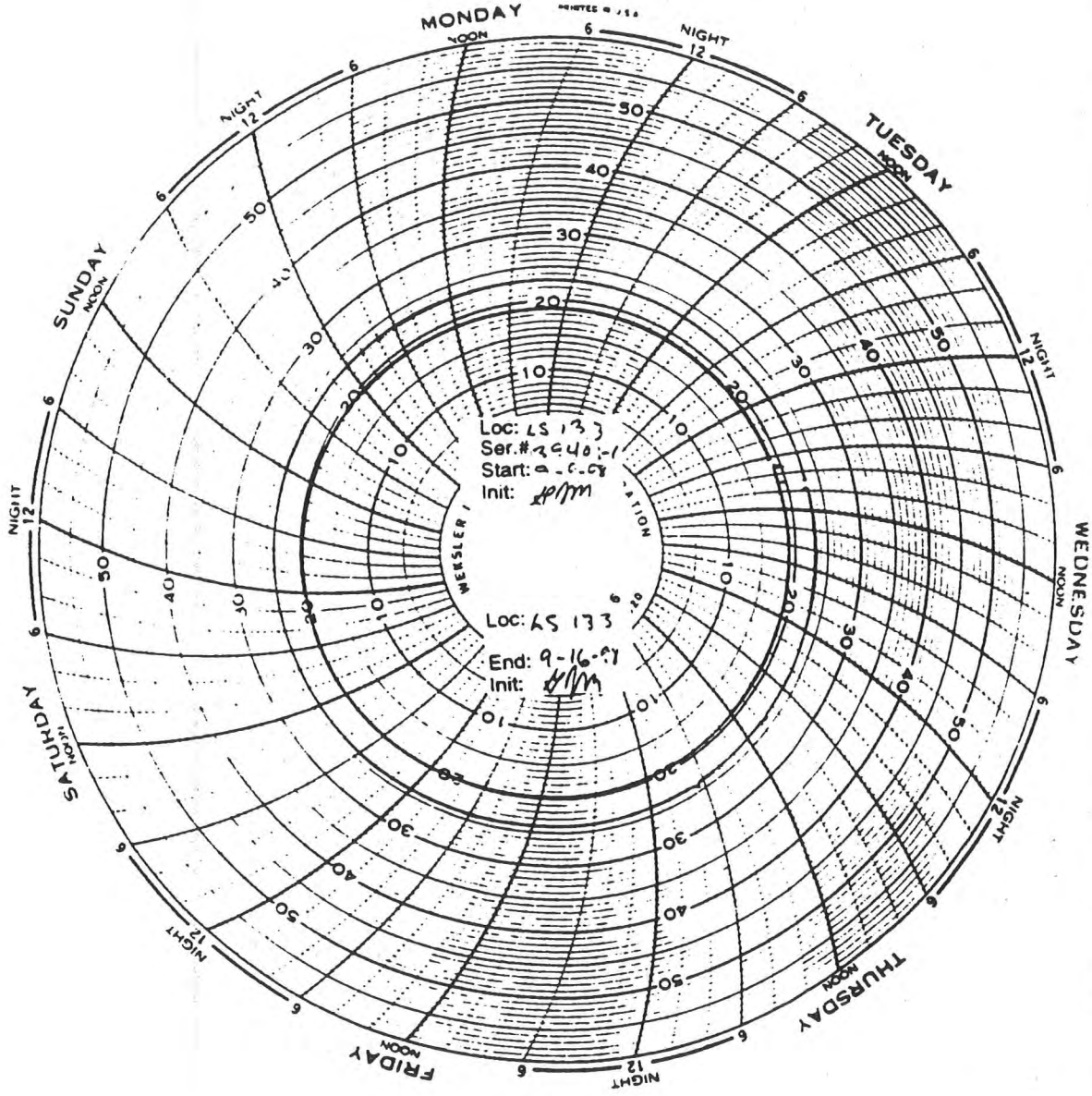
The protocol, raw data and a copy of the final report has been archived at the termination of the study. The raw data will be transferred to Mycogen Corporation. Ricerca will maintain a copy of the protocol and raw data for a minimum of three years.

APPENDIX A

Temperature/Relative Humidity Charts for Forma Incubator

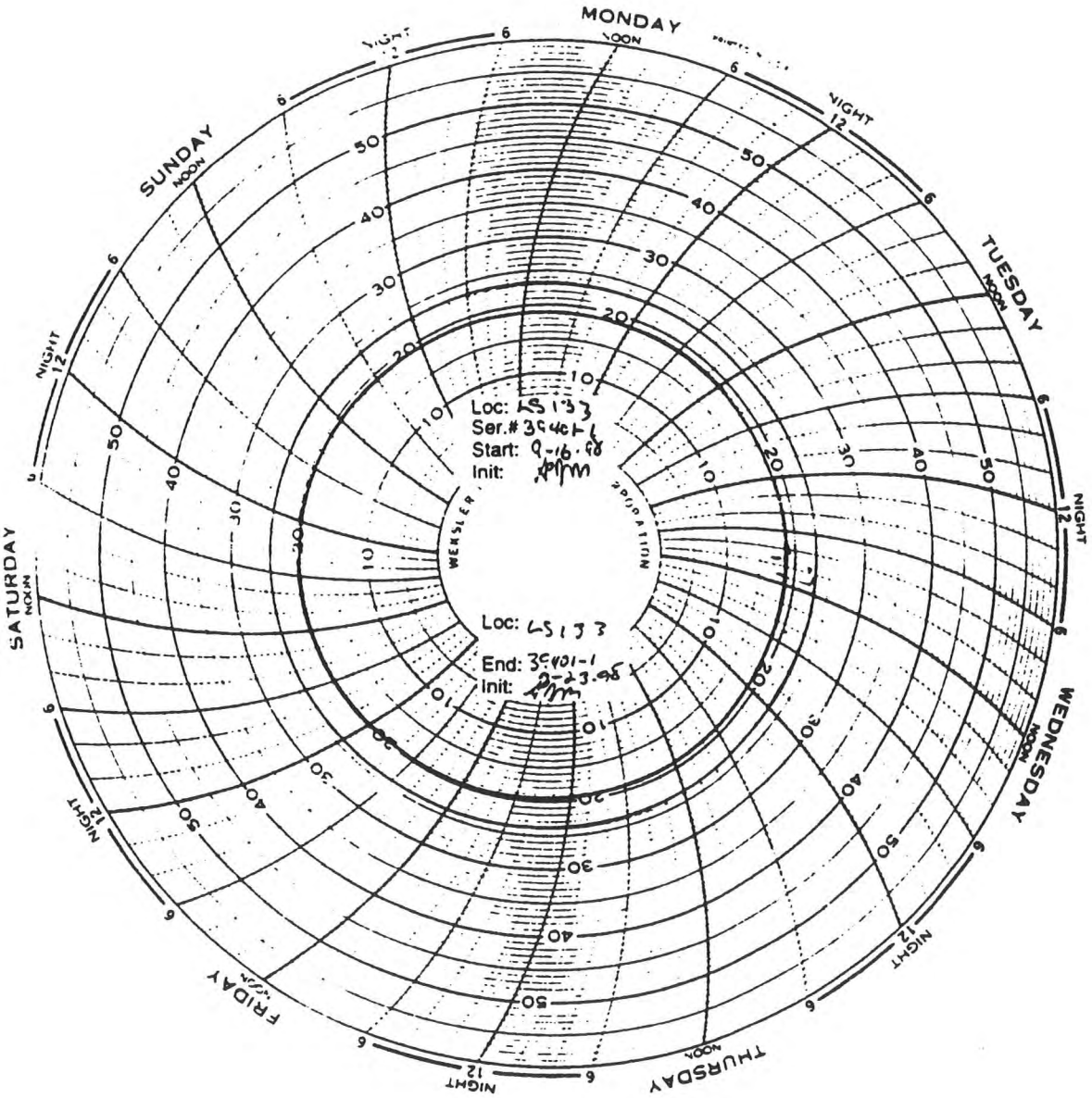
Temperature/Relative Humidity Chart Recording for Forma Incubator 9/9/98 to 9/16/98

Note: Outside tracing is dry bulb temperature. Inner tracing is wet bulb temperature. Actual relative humidity can be calculated by referring to a standard wet bulb/dry bulb temperature/relative humidity conversion chart.



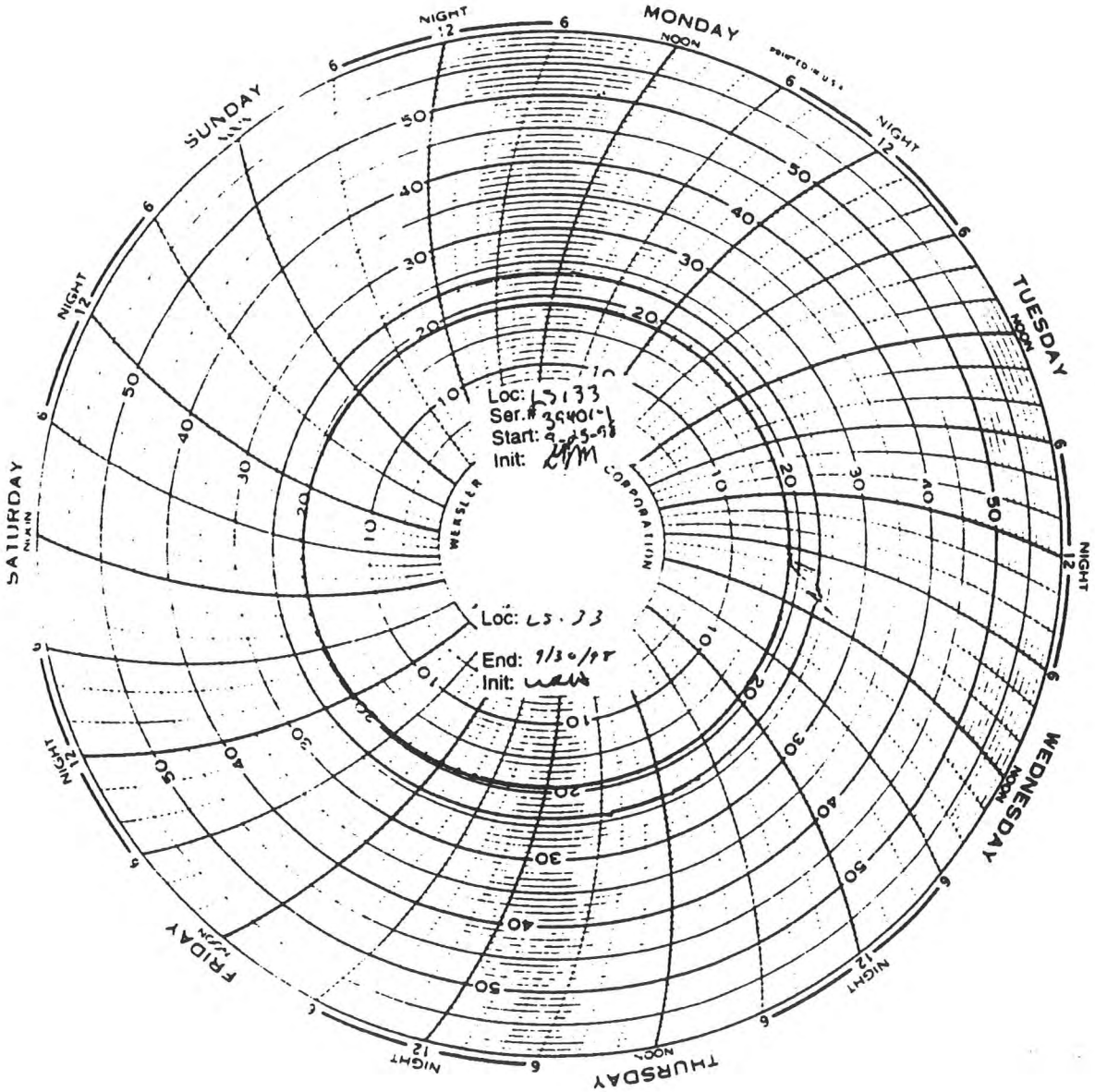
Temperature/Relative Humidity Chart Recording for Forma Incubator 9/16/98 to 9/23/98

Note: Outside tracing is dry bulb temperature. Inner tracing is wet bulb temperature. Actual relative humidity can be calculated by referring to a standard wet bulb/dry bulb temperature/relative humidity conversion chart.



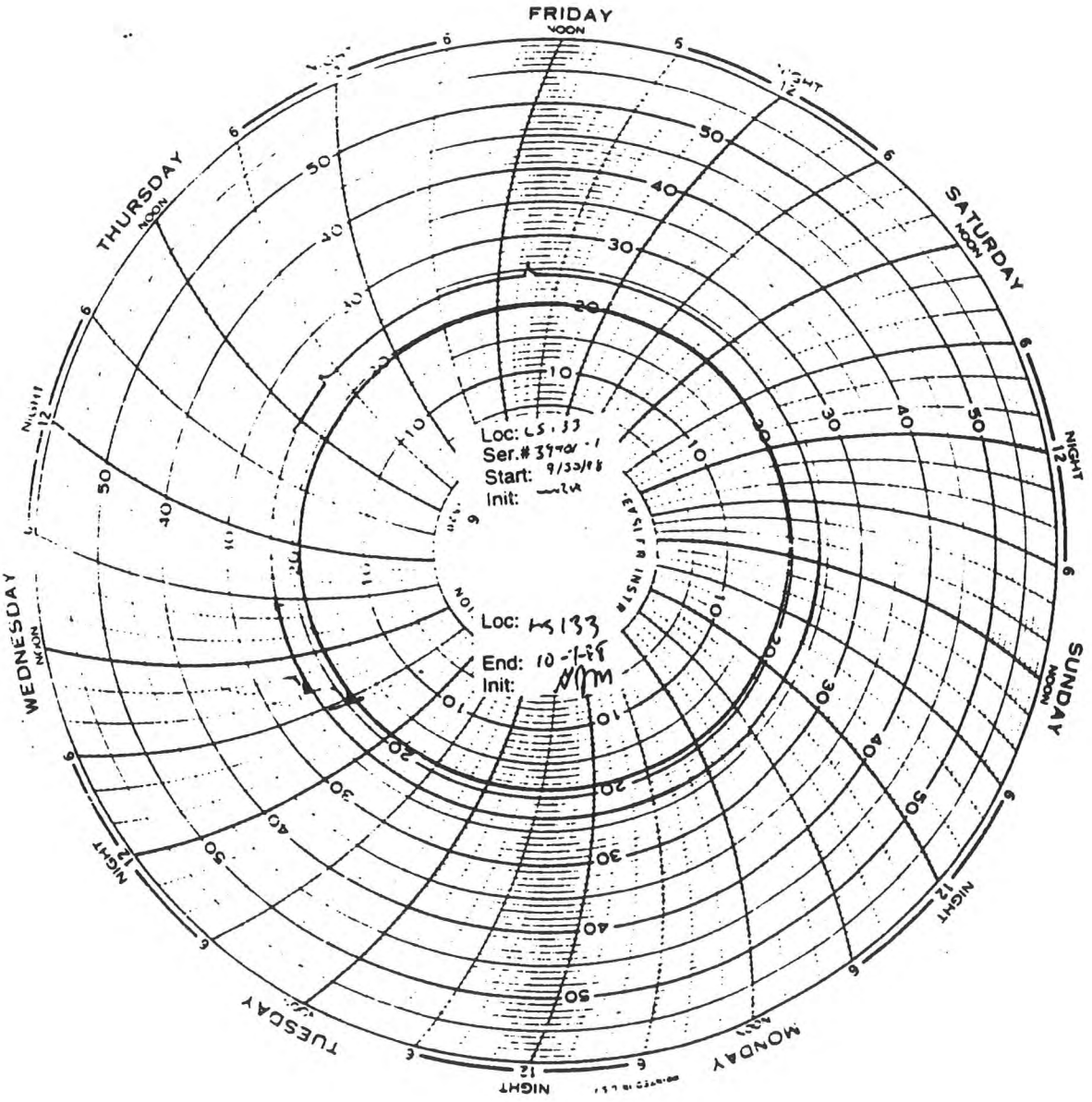
Temperature/Relative Humidity Chart Recording for Forma Incubator 9/23/98 to 9/30/98

Note: Outside tracing is dry bulb temperature. Inner tracing is wet bulb temperature. Actual relative humidity can be calculated by referring to a standard wet bulb/dry bulb temperature/relative humidity conversion chart.



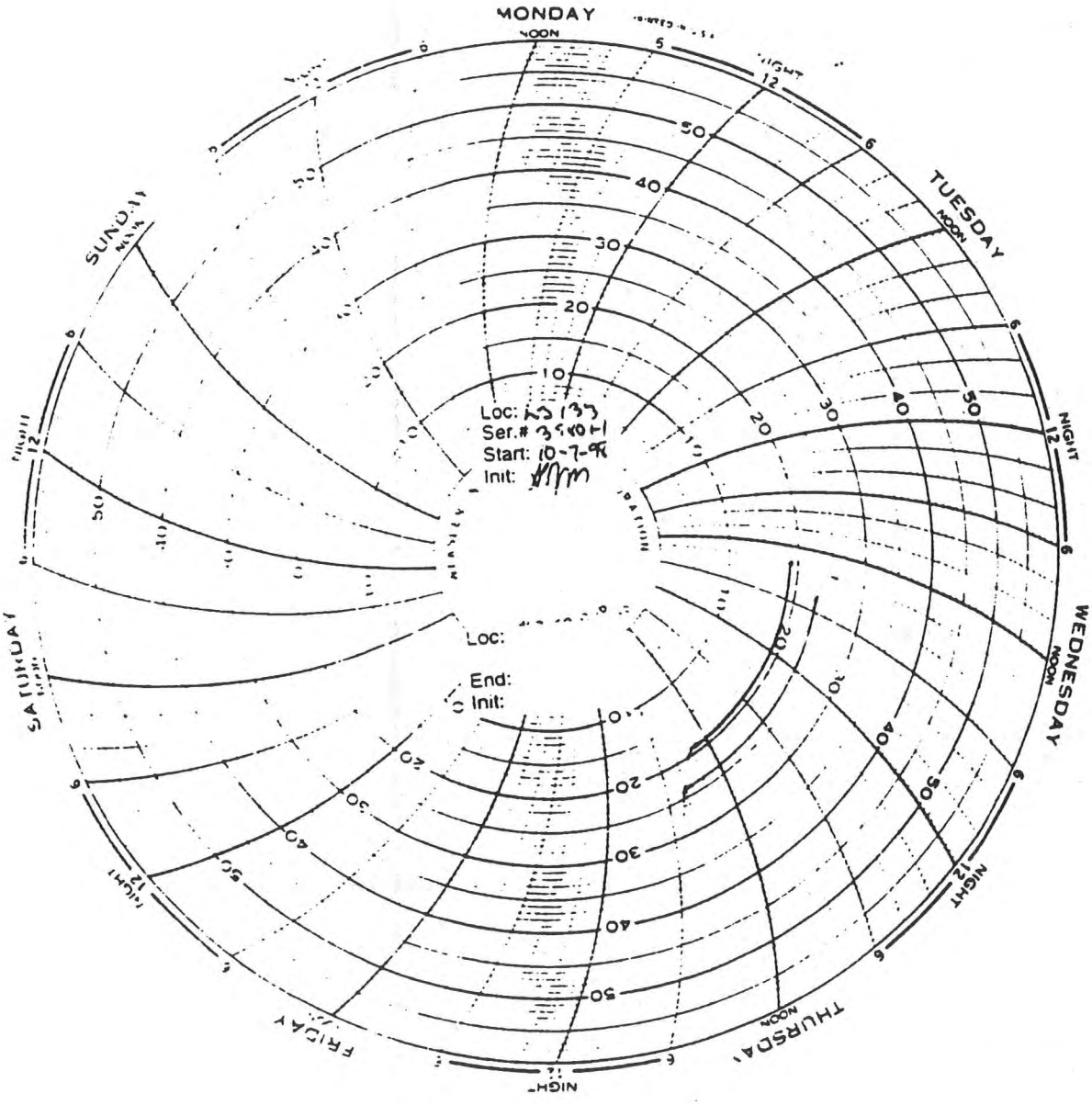
Temperature/Relative Humidity Chart Recording for Forma Incubator 9/30/98 to 10/7/98

Note: Outside tracing is dry bulb temperature. Inner tracing is wet bulb temperature. Actual relative humidity can be calculated by referring to a standard wet bulb/dry bulb temperature/relative humidity conversion chart.



Temperature/Relative Humidity Chart Recording for Forma Incubator 10/7/98 to 10/8/98

Note: Outside tracing is dry bulb temperature. Inner tracing is wet bulb temperature. Actual relative humidity can be calculated by referring to a standard wet bulb/dry bulb temperature/relative humidity conversion chart.



APPENDIX B

Original Data

Tobacco budworm Dose-Mortality Data

Dipel

Rate	# Dead (and Stunted) / total number					
	Bioassay Initiation Date					
ng Powder/cm2	10-Sep	25-Sep	1-Oct	2-Oct	Total	Dead
0	21/47, 27/47	0/48, 4/47	5/47, 3/48	0/47, 2/48	285	14
0.1	23/45					
0.3	31/46	16/46	17/46	11/48	140	44
1	38/43	37/48	19/44	28/48	140	84
3	44/46	39/46	39/47	41/45	138	119
10	45/47	47/48	43/44	47/47	139	137
30	45/45	48/48	44/45	47/48	141	139
100		48/48	44/44	47/47	139	139

Notes: Data not used - HCM

Bacterial Cry1F

Rate						
	Bioassay Initiation Date					
ng Powder/cm2	10-Sep	25-Sep	1-Oct	2-Oct	Total	Dead
0	21/47, 27/47	0/48, 4/47	5/47, 3/48	0/47, 2/48	285	14
0.1	19/48					
0.3	20/48	1/47	8/47	1/47	141	10
1	17/47	3/45	8/47	3/48	140	14
3	19/48	7/48	10/46	4/48	142	21
10	27/48	38/46	15/46	15/47	139	68
30	42/48	44/47	32/46	21/47	140	97
100	47/48	46/47	38/45	28/47	139	112
300		43/44	43/45	43/48	137	129
500	46/46	48/48	47/47	45/47	142	140

Notes: Data not used - HCM

Plant Cry1F

Rate						
	Bioassay Initiation Date					
ng Powder/cm2	10-Sep	25-Sep	1-Oct	2-Oct	Total	Dead
0	21/47, 27/47	0/48, 4/47	5/47, 3/48	0/47, 2/48	285	14
3	20/44	0/48	3/45	0/47	140	3
10	22/47	1/48	2/47	3/48	143	6
30	25/48	3/48	3/47	1/47	142	7
100	28/45	3/48	7/47	4/48	143	14
300	37/47	14/48	13/48	7/48	144	34
1000	40/45	33/48	22/46	17/47	141	72
3000	46/47	44/48	31/46	25/44	138	100
5000	41/41	45/48	35/46	36/42	136	116

Notes: Data not used - HCM

Plant Control	10-Sep	17-Sep	25-Sep	1-Oct	2-Oct
0		5/48			
(See notes)	16/46	8/48	3/48	1/46	1/47
Rate	5000	5000	5000	5000	5000

Note Explanations HCM - High control mortality

Fall Armyworm Dose-Mortality Data

Dipel

Dead (and Stunted) / total number

Rate ng Powder/cm2	Bioassay Initiation Date 11-Sep	17-Sep	24-Sep	Total	Dead
0	5/46	5/48, 3/48	1/48, 3/48	238	17
0.1	8/45	5/48		93	13
0.3	2/48	5/47	8/48	143	15
1	11/48	7/48	11/48	144	29
3	11/48	16/48	20/48	144	47
10	29/48	25/48	24/47	143	78
30	29/48	33/48	39/48	144	101
100			43/48	48	43

Notes:

Bacterial Cry1F

Rate ng Powder/cm2	Bioassay Initiation Date 11-Sep	17-Sep	24-Sep	1-Oct	Total	Dead
0	5/46	5/48, 3/48	1/48, 3/48	4/48, 5/48	334	26
0.1		4/48	3/48	8/45	141	15
0.3	15/48	4/48	8/47	4/48	191	31
1	17/48	6/48	2/48	5/45	189	30
3	22/48	16/48	11/48	12/44	188	61
10	15/48	10/47	18/48	21/46	189	64
30	25/48	29/48	22/48	33/47	191	109
100	31/48	42/48	42/46	40/48	190	155
500	40/48	47/48	47/48	43/44	188	177

Notes:

Plant Cry1F

Rate ng Powder/cm2	Bioassay Initiation Date 11-Sep	17-Sep	24-Sep	1-Oct	Total	Dead
0	5/46	5/48, 3/48	1/48, 3/48	4/48, 5/48	288	21
3		1/47	3/48	4/48	143	8
10		6/48	5/48	8/47	143	19
30	27/48	8/48	3/48	9/46	142	20
100	21/47	7/48	3/48	14/48	144	24
300	28/48	11/48	12/46	12/48	142	35
500	39/48					
1000	42/46	27/48	15/47	17/46	141	59
3000	46/48	35/42	38/47	26/48	137	99
5000		45/48	40/47	32/48	143	117

Notes: Data not used - NLM

Plant Control	11-Sep	17-Sep	24-Sep	1-Oct
	18/48	1/48	2/48	2/43
Rate	5000	5000	5000	5000

Note Explanations NLM - Non linear mortality

European Corn Borer Dose-Mortality Data

Dipel

Dead (and Stunted) / total number

Rate	Bioassay Initiation Date			Total	Dead
ng Powder/cm2	4-Sep	11-Sep	25-Sep		
0	3/45	3/47	3/46, 1/45	183	10
0.1	5/48	6/44	6/47	139	17
0.3	9/46	12/39	15/47	132	36
1	18/48	17/39	26/47	134	61
3	28/44	34/41	41/44	129	103
10	44/46	36/39	48/48	133	128
30	43/46	28/28	47/47	121	118

Notes:

Bacterial Cry1F

Rate	Bioassay Initiation Date			Total	Dead
ng Powder/cm2	4-Sep	11-Sep	25-Sep		
0	3/45	3/47	3/46, 1/45	183	10
0.1	7/46	2/21	1/45	112	10
0.3	4/47		6/46	93	10
1	7/47	19/47	11/46	140	37
3	9/48	18/38	27/47	133	54
10	22/47	24/30	41/45	122	87
30	22/48	39/40	47/48	136	108
100	39/48	27/27	48/48	123	114
500	43/46	29/29	48/48	123	120

Notes:

Plant Cry1F

Rate	Bioassay Initiation Date				Total	Dead
ng Powder/cm2	4-Sep	11-Sep	25-Sep	2-Oct		
0	1/44	3/47	3/46, 1/45	3/45, 4/46	226	12
0.1	1/46				46	1
0.3	3/47				47	3
1	3/48				48	3
3	6/45	0/9	1/48	7/45	138	14
10	2/48	1/11	3/47	5/47	142	10
30	5/45	2/10	8/46	4/46	137	17
100	6/47	2/10	10/46	13/46	139	29
300		2/9	29/43	16/47	90	45
500	8/48				48	8
1000		8/10	41/44	30/47	91	71
3000		10/10	41/45	42/43	88	83
5000		10/10	46/48	48/48	96	94
10000						

Notes: Data not used, NEI

Plant Control	4-Sep	11-Sep	25-Sep	2-Oct
	3/48	3/41	1/46	6/47
Rate:	500	5000	5000	5000

Note Explanations:

NEI - Not enough insects

Corn Earworm Dose-Mortality Data

Dipel

Rate	# Dead (and Stunted) / total number					
	Bioassay Initiation Date					
ng Powder/cm2	10-Sep	17-Sep	25-Sep		Total	Dead
0	0/47, 1/48	4/47, 3/46	0/48, 0/48		284	8
0.1	7/46	16/47	13/48		141	36
0.3	7/47	19/46	13/48		141	39
1	23/46	27/46	28/48		141	78
3	35/48	36/48	39/48		144	110
10	38/47	45/48	45/47		142	128
30	44/48	46/46	47/48		142	137

Notes:

Bacterial Cry1F

Rate						
	Bioassay Initiation Date					
ng Powder/cm2	10-Sep	17-Sep	25-Sep	2-Oct	Total	Dead
0	0/47, 1/48	4/47, 3/46	0/48, 0/48	2/48, 5/48	285	14
0.1	1/48					
0.3	26/47					
1	1/48	5/47	1/48	1/47	142	7
3	2/48	5/48	2/48	2/47	143	9
10	3/47	12/46	1/48	5/48	142	18
30	1/48	24/48	1/48	6/48	144	31
100	3/45	16/47	9/48	14/47	142	39
300		22/47	13/47	18/48	142	53
500	34/47	31/45	20/48	25/48	141	76
1000		33/47	34/48	35/48	143	102
	Data not used.					
Notes	NLM					

Plant Cry1F

Rate						
	Bioassay Initiation Date					
ng Powder/cm2	10-Sep	17-Sep	25-Sep	2-Oct	Total	Dead
0	0/47, 1/48	4/47, 3/46	0/48, 0/48	2/48, 5/48	287	8
3	7/48					
10	5/47	42/47	3/48	3/48	143	11
30	9/47	14/47	2/48	1/48	143	12
100	5/48	15/48	3/48	7/48	144	15
300	10/46	15/48	0/48	8/48	142	18
1000	11/47	24/48	5/48	8/47	142	24
3000	22/48	41/48	2/47	13/48	143	40
5000	20/46	16/43	5/48	14/48	142	39
10000		33/47	11/48	24/48	96	35
		Data not used.				
Notes:		NLM				

Plant Control	10-Sep	17-Sep	25-Sep	2-Oct
	10/45	19/48	2/46	4/48
Rate	5000	10000	10000	10000

Note Explanations:

NLM- non-linear mortality observed

Black cutworm Dose-Mortality Data

Dipel

Dead (and Stunted) / total number

Rate	Bioassay Initiation Date						
ng Powder/cm2	11-Sep	18-Sep	25-Sep	1-Oct	2-Oct	Total	Dead
0	7/46	4/46, 2/48	3/47, 4/47	3/48, 1/48	1/47, 1/45	376	19
0.1	17/46						
0.3		9/46	3/48	0/47		141	9
1	22/46	8/48	2/46	2/47	3/45	186	15
3		11/47	3/46	3/47	2/45	185	19
10	13/41	13/48	7/48	7/48	3/47	191	30
30		10/46	13/47	14/47	8/44	184	45
100		24/48	17/48	20/47	17/48	191	78
300					33/48	48	33

Notes: Data not used - Preliminary

Bacterial Cry1F

Rate	Bioassay Initiation Date					
ng Powder/cm2	11-Sep	18-Sep	25-Sep	1-Oct	Total	Dead
0	7/46	4/46, 2/48	3/47, 4/47	3/48, 1/48	284	17
0.1	11/47					
1	12/47	4/47	1/47	5/48	142	10
3		12/46	0/47	0/48	141	12
10	9/48	10/47	2/48	5/48	143	17
30	19/44	20/48	4/45	4/47	140	28
100	16/46	23/48	2/47	0/48	143	25
300		32/44	6/46	12/46	136	50
500	21/44	37/46	13/46	22/45	137	72
1000		42/48	22/47	29/47	139	83
5000		48/48			48	48

Notes: Data not used - Preliminary

Plant Cry1F

Rate	Bioassay Initiation Date					
ng Powder/cm2	11-Sep	18-Sep	25-Sep	1-Oct	Total	Dead
0	7/46	4/46, 2/48	3/47, 4/47	3/48, 1/48	284	17
10	13/43		3/48	1/47	95	4
30		3/48	5/47	2/48	143	10
100	13/47	7/48	2/47	1/48	143	10
300		2/45	3/48	3/48	141	8
1000	8/45	7/45	4/48	4/48	141	15
3000		11/46	3/47	0/43	136	14
5000	12/43	15/47	9/48	6/47	142	30
10000		32/48	10/48	3/47	143	45
20000		35/45			45	35

Notes: Data not used - Preliminary

Plant Control	11-Sep	18-Sep	25-Sep	1-Oct
	14/48	21/47	12/48	2/46
Rate	5000	20000	10000	10000

Note Explanations:

Appendix N: Report of [redacted], Pioneer, Biological Testing

The following materials were received, bacterial cell protein 8/25/98 from Steve Evans, Bot. #1 plant material, negative control and Bot. #2 Cry1F plant material were both received 9/3/98 from Jacque Rivas. The plant material was stored at room temperature while the bacterial cell prep was stored at 4deg. C. Their were no unexpected surprises with any of the materials. The experiment was started on 9/14/98 and completed on 9/21/98. LC50's were obtained by linear regression using the JMP program and shown in the attached pages. The LC50's were 0.14ppm for the bacterial cell protein and 0.11ppm for the Cry1F plant tissue. The negative control was not active and had no LC50.

[redacted] 10/13/98

Date 9/14/98

G/P Study ECB Cry1F dose response LC_{50} on plant tissue

Study # 980918.ECB

Introduction: Wanted to confirm using Bioassay the level of Cry1F in plant tissue. We used an LC_{50} to compare to known levels.

Materials and methods: Diet incorporation: one rep. microtray 1
200 ul Southland diet/well
19 wells/treatment
1 Larvae/well
7 days at 26°C

Treatments:

980914.ECB

Plant tissue

		Actual wt.
Bot.#1	Neg. control	32.61mg
Bot.#2	Pos. Cry1F	32.55mg

1.54ug/mg=1.54mg/g

1/1.45mg/g=.06494
649.4mg powder=1mgCry1F
32.47mg powder=0.05mgCry1F
0.05mg Cry1F/10ml PBS=0.005ug/ulCry1F

Bacterial cell protien Cry1F(tr)

11.4%Cry1F 2.32mg
11.4%=114mg/g

1/114mg/g=0.0088
8.8mg powder=1mgCry1F
0.5mgCry1F/10mlPBS=0.05ug/ul
1ml 0.05ug/ulCry1F+9ml PBS=10ml 0.005ug/ulCry1F
All three preps plant tissue Bot. #1&2 and Bacterial cell protien Start conc. 0.005ug/ul
Bot#1was calculated and measured based on Bot.#2 even though no Cry1F was present.

Cry1F

Rate ppm	Stock ul	PBS ul	Diet ul
0	0	300	3700
0.03	24	276	3700
0.05	40	260	3700
0.07	56	244	3700
0.1	80	220	3700
0.3	240	60	3700
0.5	400	0	3600

[redacted] 9/18/98 Witnessed _____
Date Date

Non-Cry1F plant tissue ammount of leaf tissue used same as for Cry1F plant tissue

Rate ppm	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Mean	Std.Dev.	Std.Err.	%mortality
0	8.4	3.6	7.8	0.7	0.3	D	9.5	7.9	9.9	D	9.7	5.8	8.4	4.5	6.375	3.39066	0.9062	14%
0.03	9.5	5.7	4.3	7.7	9.2	4.5	0.3	8.3	0.1	9.5	9.7	8.6	D	D	6.45	3.48334	0.931	14%
0.05	11.5	10.1	8.1	10.4	8.3	8.5	11.4	11.7	1.8	9.9	3.9	1.1	D	D	8.0583	3.74711	1.0015	14%
0.07	3.3	8.3	8.1	9.2	8.4	2.9	10.6	7.6	9.6	8	0.6	1.8	0	D	6.0308	3.72008	0.9942	7%
0.1	9.2	9.7	14.5	7.4	4.5	0.8	14	7.7	9.3	2.2	8.7	11	9.8	8.3	8.3643	3.85718	1.0309	0%
0.3	4.2	9	9.8	3	14.4	13.9	8.6	12.1	0.9	9.3	7.2	8.7	D	D	8.425	4.13062	1.104	14%
0.5	7.3	8.8	4.3	8.8	8.1	7.8	9	9.4	9.7	D	D	D	D	D	7.9111	1.76383	0.4714	36%

Cry1F plant tissue LC50=0.11ppm

Rate ppm	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Mean	Std.Dev.	Std.Err.	%mortality
0	0.5	3.8	10.1	5.7	13.1	2.8	4.8	10.2	10.6	0.3	19	11.4	12.1	D	8.0154	5.64419	1.4817	7%
0.03	0	0	0.2	0.6	0	0	0	0	0	0	0	D	D	D	0.0727	0.16488	0.0494	21%
0.05	D	D	D	D	D	D	0	0	0	0	0	0	0	0	0	0	0	43%
0.07	D	D	D	D	D	D	D	D	D	D	0	0	0	0	0	0	0	71%
0.1	D	D	D	D	D	D	D	D	D	0	0	0	0	0	0	0	0	57%
0.3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	#DIV/0!	#DIV/0!	#DIV/0!	100%
0.5	D	D	D	D	D	D	D	D	D	D	D	D	D	D	#DIV/0!	#DIV/0!	#DIV/0!	100%

Bacterial cell protien LC50=0.14ppm

Rate ppm	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Mean	Std.Dev.	Std.Err.	%mortality
0	7.9	7.4	6.2	5.3	6.3	7.7	0.6	10	1.9	3.7	D	D	D	D	5.7	2.89521	0.7738	29%
0.03	0.3	0.4	1.2	D	D	D	0	0	0	0	0	0	0	0	0.1727	0.36903	0.0986	21%
0.05	0.5	D	D	D	0	0	0	0	0	0	0	0	0	0	0.0466	0.15076	0.0403	21%
0.07	D	D	D	D	D	D	0	0	0	0	0	0	0	0	0	0	0	43%
0.1	D	D	D	D	D	D	D	0	0	0	0	0	0	0	0	0	0	50%
0.3	D	D	D	D	D	D	D	D	D	D	D	D	D	D	#DIV/0!	#DIV/0!	#DIV/0!	100%
0.5	D	D	D	D	D	D	D	D	D	D	D	D	D	D	#DIV/0!	#DIV/0!	#DIV/0!	100%

[redacted]

Witnessed

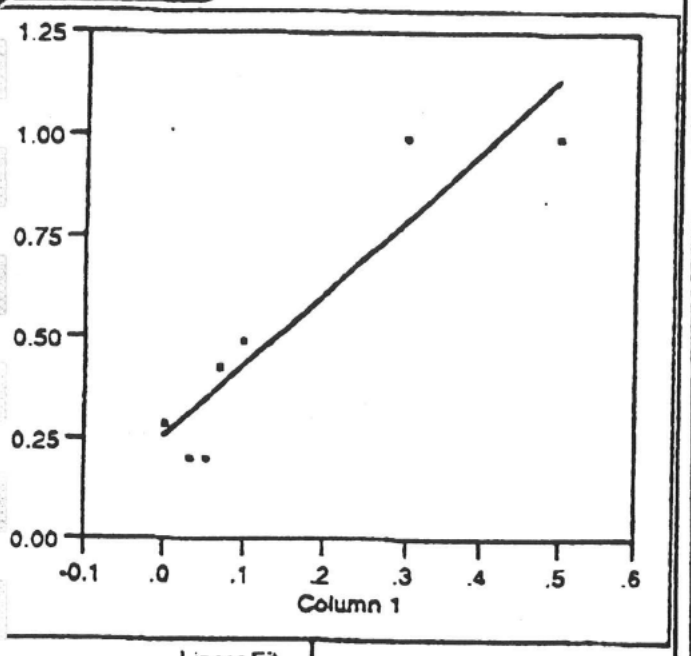
Date

Page 297 of 298

Date 9/19/98

[redacted]

[redacted]



Fit

ummary of Fit

uare	0.858344
uare Adj	0.830013
ot Mean Square Error	0.142266
of Response	0.52
rvations (or Sum Wgts)	7

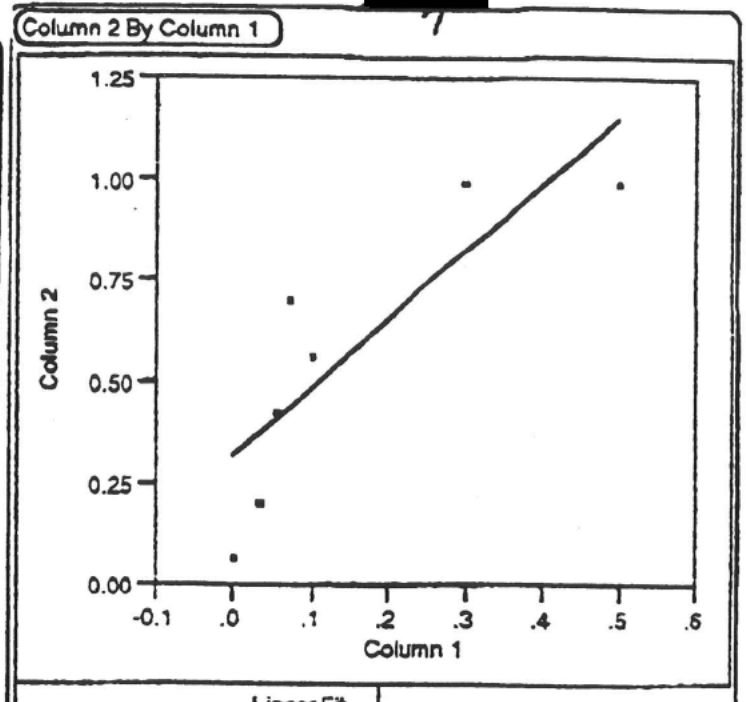
CrVIF Back, Pinto2

sis of Variance

ce	DF	Sum of Squares	Mean Square	F Ratio
del	1	0.61320125	0.613201	30.2969
	5	0.10119875	0.020240	Prob>F
al	6	0.71440000		0.0027

rameter Estimates

	Estimate	Std Error	t Ratio	Prob>|t|
ercept	0.2578735	0.071828	3.59	0.0157
nn 1	1.74751	0.317483	5.50	0.0027



Linear Fit

Summary of Fit

RSquare	0.712948
RSquare Adj	0.655538
Root Mean Square Error	0.212776
Mean of Response	0.57
Observations (or Sum Wgts)	7

CrVIF Plant Tissue

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	0.56223108	0.562231	12.4185
Error	5	0.22636892	0.045274	Prob>F
C Total	6	0.78860000		0.0168

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	0.319004	0.107428	2.97	0.0312
Column 1	1.6733068	0.474834	3.52	0.0168

$50 = 1.750(X) + 0.258$

$0.19\ ppm = LC50$

$0.50 = 1.67(X) + 0.32$

$0.11\ ppm = LC50$

[redacted]

Date

Witnessed _____

Date