

Document Number



DERBI Cover Page



244915

DN0024593

Dow AgroSciences Confidential

Document Type	Scientific Area	Document Status	PTR Number	Pages	Date Report Issued
RES	IIA DISCOVERY RESEARCH	NEW	NONE	14	16-MAY-2007

Report Title

AAD12 amino-acid homology search for similarity to toxins

Author(s) and ID's

R. A. HERMAN (U098402)
I. M. LARRINUA (U679885)

Department

Report / File Number(s)

Materials

X08; ; ; TRANSGENIC (PLANTS/ORGANISMS) + UNSPECIFIED; + ; ; VAL

Keywords

ARYLOXYALKANOATE DIOXYGENASE 12; PROTEIN; GEN-BANK; EXPRESSION IN PLANTS

Compound Number(s)	Batch/Lot Number(s)	Bayer Codes
AAD12		

Lab Notebooks

Language(s)

Performing Laboratory Name

ENGLISH

DOW AGROSCIENCES LLC,
INDIANAPOLIS, IN, USA

Data Requirement(s)

--

Geo Area and

Country (s) of Study		Protocol Number(s)	Study Number(s)	Published?	Date Study Completed
NA	US		071035	NO	16-MAY-2007
				Vertebrate?	GLP?
				N/A	NO

Reviewer(s)

--

SUMMARY

(In accordance with 40 CFR part 152, this summary is available
for public release after registration)

STUDY TITLE

AAD12 Amino-Acid Homology Search for Similarity to Toxins

DATA REQUIREMENTS

None

AUTHOR(S)

I. M. Larrinua, R. A. Herman

STUDY COMPLETED ON

16-May-2007

PERFORMING LABORATORY

Regulatory Laboratories—Indianapolis Lab
Dow AgroSciences LLC
9330 Zionsville Road
Indianapolis, Indiana 46268-1054

LABORATORY STUDY ID

071035

AAD12 Amino-Acid Homology Search for Similarity to Toxins

SUMMARY

The AAD12 (aryloxyalkanoate dioxygenase 12) protein was evaluated for amino-acid sequence similarity to known toxins. A global sequence similarity search of this protein sequence against the GenBank non-redundant protein dataset was conducted in an effort to identify potential safety concerns by examining the characteristics of related proteins. The AAD12 similarity search identified 618 proteins. None of the similar proteins returned by the search identified safety concerns that might arise from the expression of AAD12 in plants.

STUDY TITLE

AAD12 Amino-Acid Homology Search for Similarity to Toxins

DATA REQUIREMENTS

None

AUTHOR(S)

I. M. Larrinua 317 337-3123
[ILarrinua@dow.com]
R. A. Herman

STUDY COMPLETED ON

16-May-2007

PERFORMING LABORATORY

Regulatory Laboratories—Indianapolis Lab
Dow AgroSciences LLC
9330 Zionsville Road
Indianapolis, Indiana 46268-1054

LABORATORY STUDY ID

071035

STATEMENT OF NO DATA CONFIDENTIALITY CLAIMS

Compound: AAD12

Title: AAD12 Amino-Acid Homology Search for Similarity to Toxins

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

Company: Dow AgroSciences LLC

Company Agent: P. L. Hunst

Title: Global Regulatory Trait Leader

Signature: *Penny L. Hunst*

Date: 4/27/2007

*In the United States, the above statement supersedes all other statements of confidentiality that may occur elsewhere in this report.

THIS DATA MAY BE CONSIDERED CONFIDENTIAL IN COUNTRIES OUTSIDE THE UNITED STATES.

STATEMENT OF COMPLIANCE WITH GOOD LABORATORY PRACTICE STANDARDS

Title: AAD12 Amino-Acid Homology Search for Similarity to Toxins

Study Initiation Date: 02/26/2007

Experimental Start Date: 02/26/2007

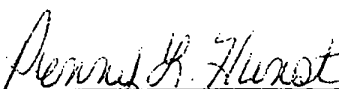
Experiment Termination Date: 03/30/2007

This report represents data generated after the effective date of the EPA FIFRA Good Laboratory Practice Standards.

United States Environmental Protection Agency
Title 40 Code of Federal Regulations Part 160
FEDERAL REGISTER, August 17, 1989

Organisation for Economic Co-Operation and Development
ENV/MC/CHEM(98)17, Paris January 26, 1998

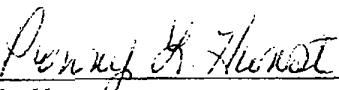
This study does not meet requirements of 40 CFR Part 160.



P. L. Hunst
Sponsor
Dow AgroSciences LLC

4/27/2007

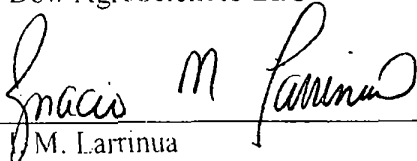
Date



P. L. Hunst
Submitter
Dow AgroSciences LLC

4/27/2007

Date



M. Larrinua
Study Director/Author
Dow AgroSciences LLC

5/16/2007

Study Completion Date

QUALITY ASSURANCE STATEMENT

Compound: AAD12

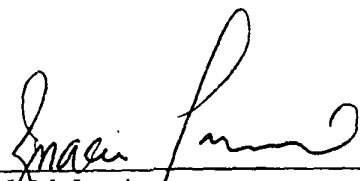
Title: AAD12 Amino-Acid Homology Search for Similarity to Toxins

Study Initiation Date: 02/26/2007

Study Completion Date: 16 May 2007

NON-GLP STUDY

SIGNATURE PAGE



I. M. Larrinua

Author

Dow AgroSciences LLC

14-May-2007

Date



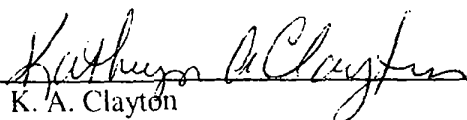
R. A. Herman

Co-author

Dow AgroSciences LLC

16-May-2007

Date



K. A. Clayton

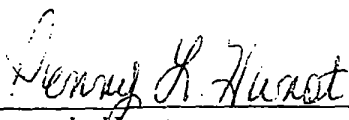
Global Biotechnology Regulatory Sciences

Group Leader

Dow AgroSciences LLC

09-May-2007

Date



Penny L. Munst

Global Regulatory Trait Leader

Dow AgroSciences LLC

5/16/2007

Date

TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT	7
INTRODUCTION.....	8
METHODS AND MATERIALS	8
RESULTS AND DISCUSSION	9
CONCLUSIONS	9
REFERENCES.....	10
FIGURE 1. AMINO ACID SEQUENCE OF AAD1.....	11
FIGURE 2. PERL SCRIPT	11

AAD12 Amino-Acid Homology Search for Similarity to Toxins

ABSTRACT

The AAD12 (aryloxyalkanoate dioxygenase 12) protein was evaluated for amino-acid sequence similarity to known toxins. A global sequence similarity search of this protein sequence against the GenBank non-redundant protein dataset was conducted in an effort to identify potential safety concerns by examining the characteristics of related proteins. The AAD12 similarity search identified 618 proteins. None of the similar proteins returned by the search identified safety concerns that might arise from the expression of AAD12 in plants.

INTRODUCTION

The AAD12 (aryloxyalkanoate dioxygenase 12) protein was evaluated for amino-acid sequence similarity to known toxins. A global sequence similarity search of this protein sequence (Figure 1) against the GenBank non redundant protein dataset was conducted in an effort to identify potential safety concerns by examining the characteristics of related proteins. This data base was posted on February 17th, 2007 and had 4,626,804 sequences with 1,596,079,149 amino acids.

METHODS AND MATERIALS

The similarity search was conducted using the BLASTP 2.2.10 algorithm (1) with a cutoff expectation (E) value of 1.0. AAD12 is an enzyme and preliminary blast results indicated that there would be more than 500 significant hits which is the limit of the web based blast interface. Therefore blastp was run on the command line using the following set of parameters –
`/usr/local/blast-2.2.10/bin/blastall -p blastp -d /usr/local/blast-2.2.10/db/blastlibs/nr -i /genomics/AAD12.txt -o /genomics/AAD12results02222007.txt -e 1 -v 5000 -b 5000 -F ""`.
Accession numbers were extracted from the description line of the BLAST result and used to download the GenPept file for each accession using batch Entrez (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=protein&cmd=search&term=>). Only one kind of error was reported by Batch Entrez. There were three records from the Structure database rather than the Protein database. These could be accessed from the Structure database and the proper GenPept file downloaded.

A perl script RegParse2.pl (Figure 2) was written which parsed the Genpept file and extracted, as tab delimited records, the locus the protein product and the region_name(s). There is only one possible protein product but many possible region_names. If there was no protein product, the script inserted a tab. The script does not handle multiline names, so these were completed manually. The tab delimited file could be imported into Excel and manipulated to sort proteins with similar protein products or region_names together.

RESULTS AND DISCUSSION

The AAD12 similarity search identified 618 proteins. These proteins were annotated with an enzyme function, because the GenPept page for this protein contained a well characterized enzyme activity in its Features annotation under product or region_name. These proteins can be broken down into a few major subclasses. The largest subclass, containing 474 proteins, was identified as tauD or taurine dioxygenases. These are proteins involved in the degradation of taurine. (2). The next largest class, with 138 members, was clavaminic acid synthetases or "CAS-like"(3). There were 2 tolC proteins (XX) which are known efflux pumps (4). The last four proteins were: 1) a (S)-2-(2,4-dichlorophenoxy)propionate, 2-oxoglutarate dioxygenase (5); 2) a pvcB protein which is a known CAS like protein (see accession page of NP_968348); 3) a inosine-uridine preferring nucleoside hydrolase (6); and 4) a hypothetical protein with no functional annotation. Inspection of the BLink results accessible from the GenPept page of this protein reveals significant homology (9e-150) to both CAS and tauD. None of these protein classes are known toxins.

CONCLUSIONS

None of the similar proteins returned by the search identify safety concerns that might arise from the expression of AAD12 in plants.

REFERENCES

1. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215:403-10.
2. Eichhorn E., Van Der Ploeg J. R. , Kertesz M. A. , Leisinger T. (1997)
Characterization of alpha-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. *J. Biol Chem* 272: 23031-23036.
3. Zhang Z., Ren J.-S., Stammers D. K., Baldwin J. E., Harlos K, Schofield C.J. (2000)
Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase.; *Nat Struct Biol* 7:127-133.
4. Koronakis V, Sharff A, Koronakis E, Luisi B, Hughes C. (2000) Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature*. Jun 22;405(6789):914-9.
5. Schleinitz, K.M., Kleinsteuber, S., Vallacys, T. Babel, W. (2004) Localization and Characterization of Two Novel Genes Encoding Stereospecific Dioxygenases Catalyzing 2(2,4-Dichlorophenoxy)propionate Cleavage in *Delftia acidovorans* MC1. *Appl Environ Microbiol* Sep;70(9):5357-65.
6. Gopaul D.N. , Meyer S.L. , Degano M. , Sacchettini J.C. , Schramm V.L. (1996) Inosine-uridine nucleoside hydrolase from *Crithidia fasciculata*. Genetic characterization, crystallization, and identification of histidine 241 as a catalytic site residue. *Biochem* 35 5963-5970.

FIGURE 1. AMINO ACID SEQUENCE OF AAD12

MAQTTLQITPTGATLGATVTGVHLATLDDAGFAALHAAWLQHALLIFPGQHLSNDQQITFAKRF
GAIERIGGGDIVAISNVKADGTVRQHS PAEWDDMMKVIVGNMAWHADSTYMPVMAQGAVFSAEV
VPAVGGRTCFADMRAAYDALDEATRALVHQRSARHSLVYSQSKLGHVQQAGSAYIGYGMTTAT
PLRPLVKVHPETGRPSLLIGRHAHAIPGMDAAESERFLEGLVDWACQAPRVHAHQWAAGDVVVW
DNRCLLHRAEPWDFKLPRVMWHSRLAGRPETEGAALV

FIGURE 2. PERL SCRIPT

```
#!/usr/bin/perl -w
#written by Ignacio Larrinua
#Feb 28th, 2007
#extract LOCUS and region_name from GenPept files

$file1 = $ARGV[0];
open (INFILE, "<$file1") or die "Can't open that file";

$file2 = $ARGV[1];
open (OUTFILE, ">$file2") or die "Can't open that file"; #10

while (<INFILE>) {
    $productnumber = 0;
    if (/^LOCUS/) {
        @locus = split (/\\s+/, $_);
        chomp $locus[1];
        $productnumber++;
    }
    elsif (/^\\s+\\/product/) {
        @product = split ("=", $_);
        chop $product[1]; #20
    }
    elsif (/^\\s+\\/region_name/) {
        @name = split ("=", $_); #20
        chomp $name[1];
        push @region, $name[1];
        undef @name;
    }
    elsif ($_ eq "//\\n"){
        print OUTFILE "$locus[1]\\t";
        $productnumber = @product;
        if ($productnumber == 0){
            print OUTFILE "\\t"; #30
        }
        else {print OUTFILE "$product[1]\\t";}
        $regionnumber = @region;
        for ($i = -1; $i <= ($regionnumber-2); $i++){
            print OUTFILE "$region[$i]";
        }
        print OUTFILE "\\n";
        undef @locus;
    }
}
```

```
        undef @product;  
        undef @region;}  
    else {next;}  
}
```

```
close INFILE;  
close OUTFILE;
```