

**ISOLATION OF THE ASTACIN-LIKE
METALLOPROTEASE CODING GENE (ASTL)
AND ASSESSMENT OF ITS INSECTICIDAL
ACTIVITY**

By

MERVAT RAGAB MANSOUR ALI DIAB

B. Sc. Agric. Sci. (Biotechnology), Fac. Agric., Cairo Univ., 2005

M.Sc. Agric. Sci. (Genetics), Fac. Agric., Cairo Univ., 2015

THESIS

**Submitted in Partial Fulfillment of the
Requirements for the Degree of**

DOCTOR OF PHILOSOPHY

In

**Agricultural Sciences
(Genetics)**

**Department of Genetics
Faculty of Agriculture
Cairo University
EGYPT**

2020

CONTENTS

	Page
INTRODUCTION	1
REVIEW OF LITERATURE	4
1. Insect pests	4
2. Insect pest control methods	5
a. Chemical insecticides.....	5
b. Integrated pest management (IPM)	6
c. Bioinsecticides.....	6
3. Spider venoms	9
a. Toxin peptide folding.....	10
b. Insecticidal targets of spider venoms and mechanisms.....	10
c. Isolation of spider toxin peptides and their activity as biopesticides	11
4. The Astacin-like metalloprotease	12
a. The <i>astl</i> gene.....	14
5. Expressing the Spider-Venom Toxins	14
6. Cotton leafworm (<i>Spodoptera littoralis</i>) (Boisduval)	16
a. The control of <i>S. littoralis</i>	16
7. Rice weevil (<i>Sitophilus oryzae</i>)	17
MATERIALS AND METHODS	18
1. Spiders collection and preparation of venom glands	18
2. Total RNA extraction	18
a. Total RNA isolation using the Trizole reagent.....	19
3. First strand cDNA synthesis	20
4. Cloning of the potential genes encoding different spider toxins	20
a. Computer analysis of the toxin genes from different spider species	20
b. Determination of the conserved regions using Clustal W2 software	22
c. The properties of the designed primers	22
d. Primers used for the amplification of the conserved regions of the genes encoding spider toxins.....	22

CONTENTS (continued)

e. PCR amplification	23
f. Purification of the PCR products.....	24
g. Cloning of the amplified fragments deduced to encode the potential toxin genes	24
h. Preparation of competent cells of <i>E.coli</i>	26
i. Transformation procedure for <i>E.coli</i>	27
j. Selection of the positive clones.....	27
k. Plasmid isolation	28
5. Sequencing and analysis of the fragments of the genes encoding the spider toxins	30
6. Rapid amplification of cDNA ends (RACE) of the gene encoding the Astacin-like metalloprotease coding gene (<i>astl</i>).....	30
a. Amplification of the 5' end of the <i>astl</i> gene	31
b. Amplification of the 3' end of the <i>astl</i> gene.....	32
c. Cloning and sequencing of the amplified ends	32
7. The assembly of the full length of the Astacin-like metalloprotease coding gene (<i>astl</i>).....	33
8. Sequences and phylogenetic analysis.....	34
9. Identification of the spider species using the DNA barcode technology.....	35
10. GNA snowdrop lectin gene assembly.....	36
11. Cloning of the active domain of the gene coding for the astacin like metalloprotease (<i>Ha-astl</i>) in bacterial and yeast expression vectors.....	38
a. Cloning and expression of the 486 bp containing the active domain of the gene coding for astacin like metalloprotease (<i>astl</i>) into bacterial expression vectors	38
b. Cloning of the active domain of the astacin like metalloprotease (<i>astl</i>) coding gene in yeast expression vectors and transformation in yeast <i>Pichia pastoris</i>	45
12. Cloning of the active domain of GNA lectin and the fused astl/GNA in yeast expression vector.....	50

CONTENTS (continued)

a. Cloning of the GNA lectin gene in yeast expression vector.....	50
b. Cloning of the <i>Ha-astl/GNA</i> fused sequence in yeast expression vector.....	51
13. Transformation and screening of pPICZαAGNA and pPICZαAHaastl/GNA plasmid into the yeast "<i>Pichia pastoris</i>" cells.....	54
14. Confirmation of the cloned genes by sequence analysis.....	54
15. Expression of the Ha-astl, GNA and Ha-astl/GNA proteins.....	55
16. Detection of gene expression	56
a. SDS-PAGE analysis.....	56
b. Western blotting analysis.....	57
c. Sandwich ELISA test.....	59
17. Assessment of the toxicity of the Ha-astl and Ha-astl/GNA proteins.....	60
a. Assessment of the toxicity on <i>Spodoptera littoralis</i>	60
b. Assessment of the toxicity on <i>Sitophilus oryzae</i>	62
18. Statistical analysis.....	62
RESULTS.....	64
1. Investigation of the genes sequence encoding the toxic genes using the bioinformatics approach.....	64
a. Bioinformatics analysis to search the sequences for the toxic genes in other spiders	64
b. Determination of the conserved regions.....	64
c. Designing the degenerate primers	64
2. Synthesis of the first strand cDNA	65
3. Amplification, cloning and sequencing of the toxin peptide genes isolated from the spider species.....	65
a. Amplification of fragments of the toxin peptide genes using the degenerate primer sets	65

CONTENTS (continued)

b. Cloning of the selected fragments in pGEM-T-easy vector.....	66
c. Sequence analysis of the selected fragments potentially encoding the toxin peptides	71
4. Identification of the spider species of Sal using the DNA barcode technology	73
5. Isolation and sequencing of the 5' and 3' ends of the Ha-astacin-like metalloprotease coding gene	75
a. Amplification and sequencing of the 5' end of the Ha-astl gene.....	76
b. Amplification and sequencing of the 3' end of the Ha-astl gene.....	76
c. Amplification of the full length of the Ha-astl gene.....	79
6. The sequence analysis of the Ha-astl gene.....	80
7. The alignment of the gene encoding the astacin like metalloprotease.....	81
8. Snow drop Lectin (GNA) assembly.....	81
9. Cloning and expression of the Ha-Astacin-like metalloprotease gene in bacterial expression vectors (pGEX-4T-3 and pQE30).....	86
a. Cloning of the active domain of the astl gene in bacterial expression vectors, pGEX-4T-3 and pQE30.....	86
b. Expression of the Ha-astl gene harbored by the bacterial expression vectors (pGEX-4T-3 and pQE30).....	87
10. Cloning, transformation and expression of the astl gene in the yeast expression vectors pPICZαA and pPICZ A.....	93
a. Cloning of the Ha-astl gene in the yeast expression vectors pPICZ α A and pPICZ A.....	93
b. Transformation of the recombinant vectors (pPICZ α A-astl) and (pPICZ A-astl) in the yeast " <i>Pichia pastoris</i> " KM71H strain.....	95
c. Expression of the astl clond in pPICZ α A and pPICZ A in the yeast " <i>Pichia pastoris</i> " "KM71H strain".....	96
11. Detection of the expressed astl protein by Sandwich ELISA.....	98

CONTENTS (continued)

12. Cloning, transformation and expression of the GNA lectin gene in the pPICZαA	99
13. Fusion of the astI and the GNA lectin active domains.....	101
a. .Cloning of the active domain of astI in pSK ⁺ vector	102
b. Cloning of the GNA fragment in the pSK ⁺ -astI vector	102
14. Cloning of the fused <i>astI</i>/<i>GNA lectin</i> genes in the yeast expression vector pPICZαA and its expression in yeast cells.....	104
a. Cloning of the fusion astI/GNA in the yeast expression vector pPICZ α A.....	104
b. Transformation of the fused fragment "astI/GNA" in the yeast " <i>Pichia pastoris</i> " "KM71H strain.....	105
c. Expression of the fused fragment astI/GNA in the yeast " <i>Pichia pastoris</i> " "KM71H strain"	108
15. Detection of the expression of the astI, GNA and astI/GNA in the yeast "<i>Pichia pastoris</i>" "KM71H strain" using the Western blot analysis.....	108
16. The insecticidal activity of Ha-astI and Ha-astI/GNA proteins on <i>Spodoptera littoralis</i> and <i>Sitophilus oryzae</i>.....	110
a. The insecticidal activity of the Ha-astI and the Ha-astI/GNA proteins on <i>Spodoptera littoralis</i>	110
b. The insecticidal activity of Ha-astI and the fused protein Ha-astI/GNA on <i>S.oryzae</i>	115
DISCUSSION.....	120
SUMMARY.....	127
REFERENCES.....	140
ARABIC SUMMARY.....	
ARABIC ABSTRACT.....	

LIST OF TABLES

No.	Title	Page
1	The names, nucleotide sequences and the annealing temperature of the thirty five primers used for amplification of the different genes coding for spider toxins and the expected size of the amplified fragments.....	25
2	Name and nucleotide sequences of the RACE primers for the astacin-like metalloprotease coding gene.....	33
3	The names and sequences of the primers used to identify the spider species, the expected size of the amplified fragment in (bp) and the names of the target gene	36
4	Names and sequences of the 10 lectin oligonucleotides	39
5	List of primers used for the cloning of astI and the fusion of astI and GNA in pSK+ vector and in bacterial and yeast expression vectors.....	40
6	List of the universal primers	42
7	The amplified PCR fragments from the 18 sets of primers demonstrating the clone number, the spider species, the name of primer combination and the expected sizes	68

LIST OF TABLES (continued)

8	The effect of the injection of 100ng and 300ng of astl on the fourth instar larvae after 120 h	110
9	The effect of different concentrations of the GNA, the Ha-astl and the Ha-astl/GNA proteins taken <i>per os</i> on the first instar larvae of <i>Spodoptera littoralis</i> for four days.....	112
10	The effect of different concentrations of the GNA, the astl and the astl/GNA proteins taken <i>per os</i> on the second instar larvae of <i>Spodoptera littoralis</i> for four days.....	114
11	The effect of different concentrations of the GNA, the astl and the astl/GNA proteins taken <i>per os</i> on the adults of <i>Sitophilus oryzae</i>	118

LIST OF FIGURES

No.	Title	Page
1	pGEM-T Easy Vector map.....	26
2	Strategy for synthesizing the assembled lectin gene fragments	37
3	Map of the bacterial expression vector pGEX-4T-3.....	40
4	Map of the bacterial expression vector pQE30.....	41
5	Map of the yeast expression vector pPICZ α A	46
6	Map of the yeast expression vector pPICZA.....	47
7	Transformed yeast colonies grown on YPD plates for 2 days at 28-30°C	49
8	Map of the pBluescriptSK(+)-vector (pSK+).....	52
9	The steps of immersing the castor leaves in the astI and fused proteins.	61
10	The steps of wheat seeds treatment with the desired protein	63
11	PCR analysis of the cDNA isolated from the five spider species using the 18 primer sets	67
12	PCR Screening (A) for Scy fragment (B) for Spa fragment (C) for Spa fragment	67

LIST OF FIGURES (continued)

13	Screening PCR for the selected Gna fragments using specific and universal primers	70
14	PCR Screening for the Sal fragments using the specific primers..	70
15	PCR Screening for the eight Pho fragments	72
16	The nucleotide sequence of the Sal fragment no. 5.....	72
17	alignment sheet on NCBI-BLAST website showing the similarity of the 119 bp of astl with other published sequences.....	73
18	PCR analysis to identify the spider species using two different barcodes.....	74
19	The partial sequence of the 28S ribosomal RNA gene of the spider <i>Hasarius adansoni</i>	75
20	The partial sequence of the mitochondrial cytochrome c oxidase subunit 1 (<i>COI</i>) gene of the spider <i>Hasarius adansoni</i>	75
21	(A) The electrophoretic separation of the 5' RACE of Ha-astacin like metalloprotease gene reactions.....	77
22	The nucleotide sequence of the 5' end of the Ha- <i>astl</i> gene.	77
23	The electrophoretic separation of the 3' RACE of Ha-astacin like metalloprotease coding gene reactions.	78

LIST OF FIGURES (continued)

24	The nucleotide sequence of the 3' end of the Ha- <i>astl</i> gene.....	79
25	The electrophoretic separation of the full length of Ha-astacin like metalloprotease coding gene reactions	80
26	Nucleotides and deduced amino acids sequence of <i>Hasarius adansoni</i> Astacin-like metalloprotease	82
27	Amino acids alignment of the Ha-astacin like metalloprotease toxin with other spiders	83
28	Phylogenetic tree of the amino acids of the <i>Hasarius adansoni</i> Astacin-like metalloprotease (Ha-astl) toxin with its corresponding in other spiders	84
29	Electrophoretic separation of the ssDNA and dsDNA for GNA fragments.	85
30	PCR amplification of the <i>GNA lectin</i> gene. (B) Screening for some transformed clones for lectin by universal primers (Sp6,T7 promoter) then specific primers (F1, R10)	85
31	Confirmation PCR for the positive clones of GNA for three positive clones using specific and universal primers (Sp6,T7 promoter)	85
32	The nucleotide sequence of the synthesized snowdrop <i>GNA lectin</i> gene	86
33	PCR amplification of the active domain of the astl gene to be cloned in the pGEX-4T-3 and the pQE30...	88

LIST OF FIGURES (continued)

34	(A). Restriction digestion of the active domain of astl (B). Digestion of the bacterial expression vectors pGEX-4T-3.....	88
35	PCR screening for the positive clones of astl active domain in pGEX-4T-3 (A) and pQE30 (B)	89
36	PCR screening for the positive clones of astl active domain in the pGEX-4T-3 and pQE30	90
37	The expression of the GST (C) and GST-astl harbored by the intact vector, pGEX-4T-3 (1-5) through 4 days on 12% SDS-PAGE gel.....	90
38	Western blot analysis for the induced empty vector (pGEX-4T-3) and three recombinant vector (pGEX-4T-3 –astl). (A) Western blot analysis for the Purified astl from the pGEX-4T-3-astl recombinant vector (B).....	91
39	The expression of the intact vector, pQE30 and the four recombinant vectors (pQE30–astl) through 4 days on SDS-PAGE gel.....	92
40	(A) Western blot analysis for the induced intact vector, pQE30 (C) and four of recombinant vectors, pQE30–astl (B) Western blot analysis for the purified astl from the pQE30-astl recombinant vector.	93
41	PCR amplification of the active domain of astl fragment to be cloned in pPICZ α A and in pPICZA.....	94

LIST OF FIGURES (continued)

42	PCR screening for the the active domain of astI fragment cloned in pPICZA α (A) and in pPICZA (B)	94
43	Restriction digestion analysis using the <i>EcoRI</i> restriction enzyme for native vector pPICZ α A and the recombinant vectors, pPICZA α -astI	94
44	Restriction digestion analysis using the <i>SacI</i> restriction enzyme for the recombinant plasmids pPICZA α -astI...	95
45	PCR for the potential transformed yeast colonies.....	97
46	(A) PCR analysis to detect the transformed yeast colonies using the universal primers (5'AOX1, 3'AOX1). (B) Nested PCR using the specific primer astIF1 and the universal primer 3'AOX1.....	97
47	The expression of recombinant and non recombinant yeast colonies on SDS-PAGE under different experiments	98
48	Sandwich ELISA showing the average values of negative control and His proteins, positive control, clone 9, clone17, clone 18 and clone 19 at OD 450.....	99
49	(A) PCR screening for 12 yeast colonies that potentially carried the positive clones for pPICZ α A-active domain of GNA (B) Nested PCR for the colonies	100
50	Sandwich ELISA showing the average values of the negative control and the recombinant plasmid pPICZ α A-GNA at OD 450.....	101

LIST OF FIGURES (continued)

51	Restriction digestion analysis for the pSK ⁺ vector and active domain of astl by <i>EcoRI</i> and <i>PstI</i>	102
52	PCR screening for the active domain of astl in pSK ⁺ vector using universal and specific primers.....	103
53	Confirmation PCR for three positive clones by different combinations of primers.....	103
54	Restriction digestion analysis for the pSK ⁺ vector , pSK ⁺ vector –astl vector and GNA fragments by <i>PstI</i> and <i>XbaI</i> restriction enzymes.	103
55	PCR screening for the fusion of the astl/GNA in the pSK ⁺ vector	104
56	Confirmation PCR for two positive clones of the fusion in pSK ⁺ using different combination of primers.....	104
57	PCR analysis for the fused fragment "Ha-astl/GNA"...	105
58	PCR screening for the fused fragment "Ha-astl/GNA" in the pPICZαA vector.....	105
59	Confirmation PCR for three positive clones of the fused fragment "Ha-astl/GNA" in pPICZαA using different combinations of primers	106
60	Restriction digestion analysis of the pPICZαA-astl/GNA using the enzyme <i>BstxI</i>	106
61	PCR screening for the pPICZαA-Ha-astl/GNA in yeast colonies	107

LIST OF FIGURES (continued)

62	Nested PCR for the pPICZ α A-Ha-astI/GNA in yeast colonies	107
63	Confirmation PCR for four positive clones for pPICZ α A-astI/GNA in yeast colonies	108
64	Sandwich ELISA showing the average values of negative control and the pPICZ α A-astI/GNA at OD 450.....	109
65	Western blot analysis for the induced yeast colonies containing the recombinant vectors pPICZ α A-astI, pPICZ α A-GNA and pPICZ α A-astI/GNA.....	109
66	Histogram revealing the mortality ratio of the first instar of <i>Spodoptera littoralis</i> after 96 hours.....	113
67	Histogram revealing the mortality ratio of the second instar of <i>Spodoptera littoralis</i> after 96 hours.....	115
68	Retardation in growth for the treated larvae fed on Ha-astI and fused protein astI/GNA.....	116
69	The development of treated larvae (T) compared to the control (C)	117
70	The fed castor leaves by the control (C) and the treated larvae (T)	117
71	Histogram revealing the mortality ratio of the adults of <i>Sitophilus oryzae</i> after fourteen days	119

Name of Candidate: Mervat Ragab Mansour Ali Diab **Degree:** Ph.D.

Title of Thesis: Isolation of the Astacin-like metalloprotease coding gene (ASTL) and assessment of its insecticidal activity.

Supervisors: Dr. Ebtissam H.A. Hussein

Dr. Ahmed M. A. Mohammed

Dr. Mahmoud M. Ahmed

Department: Genetics

Approval: 27 /7/ 2020

ABSTRACT

Biopesticides are considered ecofriendly alternatives to the chemical pesticides for controlling agricultural pests. Spider venoms are a cocktail of thousand peptide toxins, and could have insecticidal activity against a wide range of insect orders. The present investigation was conducted with the main goal of assessing the efficacy of spider-based toxin peptides for pest control and to determine the probability of using these toxins to produce a new biopesticide that is friendly to the environment. Twenty five fragments from genes coding for putative toxins were amplified from five spider species. Sequence analysis revealed that out of the twenty five sequences, only one fragment showed high similarity with the astacin-like metalloprotease toxin gene. Using the barcoding technology, the spider species of this fragment was identified as *Hasarius adansoni*. The full length of *astl* cDNA was cloned from this spider species using the RACE technology. Sequencing of the cloned *Ha-astl* cDNA proved that its full length includes 802 bp with 714bp open reading frame encoding for 238 amino acids. The predicted molecular weight of the encoded protein is 27.33 kDa with two disulfide bridges between cysteine residues at positions 87-238 and 108-128 and three zinc binding sites at histidine residues at positions 136,140 and 146. A 486bp of the catalytic domain was cloned and expressed by the yeast expression system *Pichia pastoris*. In an attempt to enhance the insecticidal activity of the toxic protein, the astacin like metalloprotease toxin was fused to the GNA snowdrop lectin in the same frame and expressed in *Pichia pastoris*. The insecticidal activity of the *Ha-astl* and the *Ha-astl*/GNA proteins was determined towards two species of agricultural insects from two different orders, cotton leaf worm, *Spodoptera littoralis* (Lepidoptera: Noctuidae) and rice weevil, *Sitophilus oryzae* (Coleoptera: Curculionidae). The bioassay was performed using three concentrations (100,500 and 1000 µg/ml) for four days for *S. littoralis* and 14 days for *S. oryzae*. In *S. littoralis*, the highest insecticidal effect was recorded for the larvae fed for 4 days on the fused protein at the highest concentration (1000 µg/ml). The mortality

percentages were $78.6\% \pm 4.16$ and $71.66\% \pm 3.51$ for first and second larval instars, respectively. This was followed by the larvae fed on astl protein after four days as $69.3\% \pm 2.51$ and $65\% \pm 2.64$ mortality for the first and second larval instars, respectively. In addition, growth retardation and different types of larvae abnormalities were noticed. While in *S. oryzae* the mortality ratio induced by the Ha-astl protein was $46.6\% \pm 0.5$, $48\% \pm 1.7$ and $64\% \pm 3.0$ for the three concentrations 100, 500, 1000 $\mu\text{g/ml}$, respectively. In addition, lower mortality ratios were observed for the adults treated with the three fused protein (Ha-astl/GNA) concentrations, i.e., $41.3\% \pm 2.0$, $42.6\% \pm 2.8$ and $49.3\% \pm 2.0$, respectively.

Key words: Zinc metalloproteases, astl, GNA, Fusion technology. Insect control, *Spodoptera littoralis*, *Sitophilus oryzae*.