Molecular mechanism of action of *Lysinibacillus sphaericus* Binary toxin protein components

By

MAHIMA SHARMA

Enrollment No. - LIFE01201204008

Bhabha Atomic Research Centre, Mumbai

A thesis submitted to the

Board of Studies in Life Sciences

In partial fulfillment of requirements

for the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE

December, 2017

Homi Bhabha National Institute

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Mahima Sharma entitled "Molecular mechanism of action of *Lysinibacillus sphaericus* binary toxin protein components" and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Chairman – Prof. Basuthkar J. Rao	**Date:**
Guide / Convener – Prof. Vinay Kumar	**Date:**
Examiner – Prof. Punit Kaur	**Date:**
Member 1- Dr. Anand Ballal	**Date:**
Member 2- Dr. Sahayog Jamdar	**Date:**

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.

I/We hereby certify that I/we have read this thesis prepared under my/our direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date:

Place:

Guide

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Mahima Sharma

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

Mahima Sharma

List of Publications arising from the thesis

Journal

1. "An oligomeric complex of BinA/BinB is not formed in-situ in mosquito-larvicidal *Lysinibacillus sphaericus* ISPC-8", Hire R., Sharma M., Hadapad A.B., Kumar V., *J. Invertebr. Pathol*., **2014**, *122*, 44–47.
2. "PEGylation enhances mosquito-larvicidal activity of *Lysinibacillus sphaericus* Binary toxin", Sharma M., Hire R., Hadapad A.B., Gupta G.D., Kumar V., *Bioconjugate Chemistry,* **2017**, *28*, 410–418.
3. "Receptor protein of *Lysinibacillus sphaericus* mosquito-larvicidal toxin displays amylomaltase activity", Sharma M., Gupta G.D., Kumar V., *Insect Biochem. Mol. Biol.,* **2017**, *93*, 37-46.
4. "Mosquito-larvicidal BinA toxin displays affinity for glycoconjugates: basis of BinA mediated cytotoxicity", Sharma M., Gupta G.D., Kumar V. *(Under review)*
5. "Crystallization and preliminary structural analysis of Cqm1, a BinAB receptor protein", Sharma M., Lakshmi A., Gupta G.D., Kumar V. (*Manuscript under preparation*).

Abstracts

1. "Polyethylene glycol conjugation enhances mosquito-larvicidal activity of *Lysinibacillus sphaericus* BinA protein", Sharma M., Hire R., Hadapad A.B., Gupta G.D., Kumar V., *Biophysical Journal,* **2017**, *112*, 49a.

Conferences

1. "Characterization of carbohydrate binding domain of BinA protein from *Lysinibacillus sphaericus*", Sharma M., Gupta G.D., Kumar V., **Indo-US conference: Advances in Enzymology: Implications in health, physics and therapeutics**. 17–19 Jan. 2017, ACTREC, Mumbai, India.
2. "Polyethylene glycol conjugation enhances mosquito-larvicidal activity of *Lysinibacillus sphaericus* BinA protein", Sharma M., Hire R., Hadapad A., Gupta G.D., Kumar V., **61stAnnual meeting of Biophysical Society**. 11–15 Feb. 2017, New Orleans, Louisiana, USA.
3. "*Lysinibacillus sphaericus* BinA protein: insight into its mode of action", Sharma M., Gupta G.D., Kumar V., **Annual Symposium of the Indian Biophysical Society**. 23-25 March 2017, IISER Mohali, Punjab, India.

GenBank Submissions

- **KY929304.** *Culex quinquefasciatus* isolate Trombay binary toxin receptor protein (Cqm1) gene, complete cds.
- **KY929305.** *Culex quinquefasciatus* isolate Mankhurd binary toxin receptor protein (Cqm1) gene, complete cds.
- **MG211157.** Cqm1_Synthetic, gene sequence optimized for expression in *E. coli* of *Culex quinquefasciatus* Cqm1 protein.

Mahima Sharma

"If you think you are too small to make a difference, try sleeping with a mosquito"

…. Dalai Lama

ACKNOWLEDGEMENTS

The last five years have been a period of intense learning for me, not only in the scientific arena, but also on a personal level. With this note of thanks, I would like to extend my sincere gratitude to all those who have supported and helped me throughout this period.

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Vinay Kumar, Head of Division, Radiation biology & health sciences division, BARC, for his continuous support, patience, motivation, and immense knowledge. His guidance helped me in all the times of research and writing of this thesis. I am extremely indebted to him for his invaluable guidance in orienting my scientific thoughts, writing skills and shaping my research career. I could not have imagined having a better advisor and mentor for my Ph.D. study.

Besides my supervisor, I would like to sincerely thank my doctoral committee: Chairperson, Prof. Basuthkar J. Rao, Department of Biological Sciences, TIFR, and members, Dr. Anand Ballal, Molecular biology division, BARC and Dr. Sahayog Jamdar, Food technology division, BARC, who have generously given their time and their insightful comments and encouragement, and for the hard questions which incented me to widen my research from various perspectives.

I am immensely grateful to Dr. Gagan Deep Gupta for generously sharing his meticulous research and insights and for the stimulating discussions that supported and expanded my own work experience.

I express my sincere thanks to Dr. Ramesh Hire and Dr. Ashok Hadapad, who have been active collaborators in the research work. I heartily appreciate their precious time and support.

I thank all my lab members, Ms. Shweta, Dr. Subhash Bihani, Dr. Lata Panicker, Dr. Mukesh Kumar, Dr. Vishal Prashar and Dr. Amit Das for their support and well wishes during this period. Working together has been a memorable experience. I also thank Mr. S. Jadhav for his prompt help with the store supplies and other related technical support.

I would also like to thank Dr. H. D. Sarma (BARC), Dr. Ajay Saini (BARC), Dr. AVSS Narayan Rao (BARC), Dr. Ashok Verma (ACTREC), Dr. V. Prasanna (ACTREC) and Dr. Sanjeeva (IIT Bombay) for their help and giving access to their research facilities.

I am extremely grateful to Anuradha and all my friends from 56th Batch of Bioscience training school, BARC, whose friendship, hospitality and knowledge have supported and entertained me over the years of our friendship.

I am very thankful to Dr. Hema Rajaram, Dean, Life Sciences, Homi Bhabha National Institute (HBNI) for all the help and support in managing and completing the Ph.D.-related official paperwork as well as for academic advice. I am also grateful to HBNI and Department of Atomic energy (DAE), India, for providing the fellowship to carry out my research work. I also thank HBNI and Department of Biotechnology (DBT), India, for giving me the international travel grant to attend and present my research work at the international conference.

I would take this moment to express my thankfulness once again to Prof. Vinay Kumar and Dr. Rupinder Bhatia for their immense care, hospitality and delectable dinners.

Last but not the least; I would like to thank my family, my wonderful parents and my wonderful beloved sister, for their unconditional love and immeasurable support in all my life endeavors. Thanks for always being there for me.

Thank you very much!

Mahima Sharma

CONTENTS

Homi Bhabha National Institute

SYNOPSIS OF Ph. D. THESIS

1. **Name of the Student:** Mahima Sharma

2. **Name of the Constituent Institution:** Bhabha Atomic Research Centre

3. **Enrollment No.:** LIFE01201204008

4. **Title of the Thesis:** Molecular mechanism of action of mosquito-larvicidal *Lysinibacillus sphaericus* binary toxin protein components

5. **Board of Studies:** Life Sciences

SYNOPSIS

(Limited to 10 pages in double spacing)

Mosquito borne diseases cause millions of deaths every year. *Aedes*, *Anopheles* and *Culex* are the most common species transmitting serious human diseases like Zika, Chikungunya, Dengue, Yellow fever, Malaria and West Nile fever etc. Although many physical and chemical methods for mosquito control are available, developing resistance among mosquitoes demands development of novel and alternative approaches. Also, chemical insecticides are harmful to the ecosystem. Several strains of *Lysinibacillus sphaericus* (formerly known as *Bacillus sphaericus*), a gram positive, aerobic, spore producing bacterium found throughout the world in soil and aquatic environments, are found to be toxic to *Culex* and *Anopheles* mosquito larvae and are poorly or not toxic to larvae of *Aedes* species. The bacterium has been evaluated by World Health Organization for its high efficiency for controlling mosquito-borne diseases and for its safety against non-target organisms (**1**).

The high larvicidal activity of *L. sphaericus* is due to the presence of binary (BinAB) toxin which is produced as parasporal crystalline inclusions during the early stages of sporulation. Binary toxin is composed of two highly conserved pro-BinA (41.9 kDa) and pro-BinB (51.4 kDa) proteins. Once ingested by susceptible larvae, the crystalline inclusions are solubilized in the alkaline pH of larval midgut, followed by activation of both pro-BinA and pro-BinB by the proteases. Active proteins (BinA, 39 kDa; BinB, 42 kDa) bind specifically to brush border membrane fractions at midgut epithelium of mosquito larvae. The receptor for active BinAB has been identified as a 66 kDa α-glucosidase in *Culex* (Cqm1/Cpm1; 580 AA) (**2**). BinB is a receptor binding subunit and BinA confers toxicity. It is accepted that a heterotetrameric complex, BinA2.BinB2, assembles on the receptor and triggers a set of cytopathological events on internalization, like formation of pores in the membrane, cytoplasmic vacuolation, and autophagy (**3**). The actual mechanism of larval death, however, remains unclear. These proteins display maximum toxicity at equimolar concentration. Also, a preformed covalent complex of BinA and BinB, synthesized by glutaraldehyde crosslinking, and high concentration of BinA alone displayed toxicity against Culex (**4,5**). The resistance that develops over frequent use of *L. sphaericus* is raising serious question for the long-term use of *L. sphaericus* as larvicide. Mutations in the genes encoding the receptor appear to be primary mechanism of resistance. The recombinant *L. sphaericus* expressing *B. thuringiensis* larvicidal proteins seems to reduce the potential for resistance (**6**).

The crystal structure of BinAB reveals presence of a globular N-terminal β-trefoil domain and β-rich C-terminal domain in BinA and BinB proteins (**7**). The β-trefoil scaffold is highly conserved architecture of some sugar binding proteins (lectins) and the β-rich C-terminal domain shares similarity with aerolysin type β-pore forming toxins. The knowledge of structure-activity relationship is expected to provide platform for engineering these proteins with higher toxicity and with broader target specificity. The primary objectives of

the present thesis are structural and biochemical analyses of the functional domains of binary toxin component proteins, understanding the molecular mechanism of toxicity, understanding the interaction of BinAB proteins with the receptor protein, and to "rationally" improve the efficacy of BinAB toxin. A highly toxic local isolate of *L. sphaericus* (ISPC-8) was used in the present studies. The **Chapter 1** of the thesis provides introduction and summary of the literature. And **Chapter 2** of the thesis discusses Methods used in the present study.

The *L. sphaericus* spore-crystals contain high molecular weight proteins (~125 kDa, 110 kDa), in addition to the binary toxin components. The identity of these high molecular weight complexes has been intensely debated. It has been argued that the high molecular weight complexes are preformed stable oligomers of BinA and BinB (**8**). **Chapter 3** of the thesis discusses the characterization of these high molecular weight complexes purified from the local isolate ISPC-8 (**9**). The spore–crystal pellet of *L. sphaericus* ISPC-8 (serotype 5a5b and phage type 3) was isolated from the culture grown in the standard nutrient broth (NB) medium supplemented with 0.3% sugar cane molasses. Equal quantities of BinAB and high molecular weight protein/complex were observed in the spore crystals and these proteins co-purified on anion-exchange chromatography. The high molecular weight complex was purified to homogeneity by three stage column chromatography methods (anion exchange, and gel filtration using Superdex 200 and Superdex 75 columns) and appeared as two protein bands (125 kDa and 110 kDa) on 5% SDS-PAGE gel. These were probed with proteomics and biophysics tools. Contrary to earlier expectations, peptide mass fingerprinting and MS/MS analyses revealed with high confidence (MASCOT score 158, Expect score 3.9E–09) that the two SDS-PAGE bands of high molecular weight proteins correspond to surface layer protein (SlpC). The bioinformatics analysis revealed that SlpC (molecular weight of 125.2 kDa estimated from protein sequence) is highly conserved in different strains of *L. sphaericus*. The SlpC protein of *L. sphaericus* is predicted to be secreted with a leader peptide that is

removed between Ala-30 and Ala-31 of the primary translation product and is a precursor of the 110 kDa peptide (**10**). The purified SlpC protein from the local isolate was found to exist as a dimer. Its molecular mass was estimated to be ~282 kDa by size exclusion chromatography and 243.8± 59.8 kDa by dynamic light scattering experiments. The purified SlpC was not toxic to *Culex* larvae up to 100 µg/ml protein concentration, but showed poor toxicity (20 % mortality) at 150 µg/ml, as compared to control experiments with BinA/BinB mixture from spore crystal (LC_{50} ~ 6 ng/ml).

The active BinAB toxin from *L. sphaericus* binds to surface receptor protein (Cqm1), which is bound to the midgut epithelial cells via a glycosyl-phosphatidylinositol (GPI) anchor. The Cqm1 protein has been classified as a member of glycoside hydrolase family 13 (GH 13) of the CAZy database. It shares 74 % identity with Aam1 receptor protein from *Aedes* species. However, Aam1, but not Cqm1, has been observed to be glycosylated *in vivo* (**11**). **Chapter 4** discusses the characterization of the ordered domain (residues 23-560) of Cqm1. BinAB susceptible *Culex quinquefasciatus* colony was established from local *Culex* strains found in Trombay and Mankhurd regions, Mumbai, India, and were reared without exposure to *L. sphaericus*. The *cqm*1 gene from BinAB susceptible *Culex quinquefasciatus* mosquito colony was PCR amplified and sequenced (GenBank accession number: KY929304). The local strain showed polymorphism at six nucleotide sites (GenBank accession number: KY929305). The *cqm*1 gene (1840 bp) contains a 1740 bp coding DNA sequence (CDS), which encodes for 580 amino acid long Cqm1 protein. Due to difficulties encountered in cDNA preparation, *cqm*1 gene with CDS of 1740 bp was synthesized chemically with codons optimized for expression in *E. coli* host. The *cqm*1 gene was cloned into different expression vectors using restriction-digestion based cloning and ligation independent cloning (LIC) approaches. Fusion systems including glutathione S-transferase and trigger factor as solubilizing tags were also tried. However, none of the attempts could successfully yield fully folded soluble

protein. For expressing soluble Cqm1 protein, truncated *cqm*1 gene (1614 bp) lacking 66 bases from the 5'-end (coding for N terminal signal peptide) and 60 bases from the 3'-end (coding for C terminal GPI anchor) was PCR amplified and cloned into pNIC28-Bsa4 vector by LIC method, and the encoded protein was purified using *E. coli* expression system using three-stage column chromatography methods (immobilized metal ion affinity, anion exchange, and gel filtration using Superdex 200 column). The protein displayed hydrolytic activity against 4-nitrophenyl α-D-glucopyranoside (α-pNPG) and not against 4-nitrophenyl β-D-glucopyranoside. Values for the Michaelis-Menten kinetics parameters towards α-pNPG substrate were estimated to be 0.44 $\mu mol.min^{-1}.mg^{-1}$ (*K*m) and 1.9 s^{-1} (*k*cat). The specific activity was estimated to be 3.2 Units /mg. The optimum temperature (probed in the range 25-60 ^oC) and pH (probed for pH range 4.6-11.0) for hydrolytic activity against α-pNPG were observed to be about 37 ^oC and pH 7.2, respectively. The observed pH optimum relates well to the physiological conditions found in the posterior half of larval midgut with near neutral pH. Fluorescently labeled BinAB were earlier observed to co-localize in the posterior midgut region of *Culex* and *Aedes* larvae (**12**). The melting temperature *T*m for Cqm1, as observed by thermofluor-shift assay, was 51.5 ^oC and Ca^{2+} was observed to provide structural stability to the protein. The size exclusion chromatography, dynamic light scattering and Native-PAGE revealed the dimeric status of Cqm1 that is consistent with the apical localization of GPI-anchored Cqm1 in lipid rafts. Presence of aerolysin domain in BinA and BinB proteins and dimeric status of the recombinant Cqm1 protein also corroborates the suggestion that oligomer of Cpm1 (Cqm1 ortholog from *Culex pipiens*) promotes the opening of ionic pores by interacting with BinAB. Further, toxic BinA component did not inhibit α-glucosidase activity of Cqm1, while BinB reduced the activity by nearly 50%. Since BinB alone does not exert toxic effect, it rules out the possibility discussed earlier (**2**) that larvicidal activity may be due to inhibition of receptor biochemical activity.

Notably, thin layer chromatography experiments established Cqm1 as α-glucosidase competent to cleave α-1,4-glycosidic bonds of maltose and maltotriose with high glycosyltransferase activity. The glycosyltransferase activity results in formation of maltose-oligomers from shorter maltodextrins by new α-1,4-linkages. The Cqm1 protein also hydrolyses glycogen and sucrose, but not isomaltose and trehalose. This is a novel and new finding. The observed hydrolysis and synthesis of maltose-oligomers is consistent with open and accessible active-site consisting of a catalytic triad (Asp-224, Glu-290 and Asp-358). The bioinformatics structural model was constructed using MODELLER and SWISS PDB tools, as attempts to crystallize Cqm1 were not successful. These observed amylomaltase and sucrase activities suggest that Cqm1 may be involved in carbohydrate metabolism in mosquitoes.

Carbohydrates, as free oligosaccharides or as glycoconjugates, interact with carbohydrate-binding proteins like lectins and play important roles in many biological events. Earlier competitive-assay studies had suggested that sugars influence the toxicity of BinA (**13**). Also, the presence of three carbohydrate binding modules in BinA structure has been reported recently **(7)**. **Chapter 5** of the thesis discusses interaction of BinA with glycans and with the receptor protein, Cqm1. BinA did not display hemagglutination and hemolytic activities, hence novel high-throughput approaches based on fluorescence using acrylodan as an extrinsic fluorophore and thermofluor shift assay using SYPRO-orange dye were used to characterize carbohydrate specificity of BinA. A total of 34 saccharides (monosaccharides, disaccharides, polysaccharides, and glycoproteins) were used for the initial high-throughput screening. The chosen ligand set included glycans known to bind most of the lectins. The promising glycans were identified based on significant change (>10%) in the fluorescence intensity. No change in thermofluor shift assay, however, was observed for any ligand, may be owing to the nature of BinA/ligand interactions. Binding of simple sugars could not be

confirmed in Isothermal Titration Calorimetry experiments. However, changes in the water absorption spectral region of FTIR (near 3000 cm^{-1}) revealed that L-fucose and N-acetyl-D-lactosamine bind to BinA. Further, glycoproteins (fetuin, asialofetuin and thyroglobulin) were observed by Surface Plasmon Resonance to show strong binding towards BinA with K_D values less than a micromolar. This study thus demonstrates that BinA is competent to bind diverse and structurally different glycosylated proteins. BinA, however, did not show direct interaction with Cqm1 (purified without GPI anchor) as was evident from size exclusion chromatography and chemical (glutaraldehyde) crosslinking experiments. These results are suggestive of the mechanism for cytotoxic effects of BinA. It may bind diverse glycosylated proteins inside the cell through carbohydrate binding domain. The interaction of aerolysin domain of BinA with N-glycan core of GPI anchor for receptor recognition has been reported for several bacterial toxins (1**4**). The insect proteome analysis reveals that glycosylated proteins are abundant in insects and are essential for larval survival (**15**). Our hypothesis also rationalizes non-toxic effect of BinA on *Aedes aegypti*.

Instability, short half-lives and rapid proteolytic digestion can limit use of BinA/BinB as an effective insecticide in real-life application. PEGylation of therapeutic proteins and peptides, by covalent attachment of polyethylene glycol chains to the target molecules, is proving to be highly potent tool to prolong bio-availability, increase stability, and reduce immunogenicity. **Chapter 6** demonstrates the beneficial effect of PEGylation on mosquito-larvicidal activity of BinA (**16**). Polymer conjugation was achieved using methoxy polyethylene glycol isocyanate (mPEG-ISC-750) at two different pH values (pH 7.2 and 8.5). At pH 7.2, the N-terminus amine group can be expected to react with isocyanate due to its low pKa (7.7 $\pm$ 0.5) generally found in proteins. At pH 8.5, however, amine group of accessible lysine residues (pKa, 10.4 $\pm$ 1.1) can also form covalent linkage. Two isoforms of the biopolymers were purified to homogeneity by three-stage column chromatography

methods. These were highly water soluble and resistant to trypsin and proteinase K proteases. In comparison, we found BinB, isolated from *L. sphaericus* spore crystals, to be susceptible to degradation by these proteases. PEGylation in both the isoforms was confirmed by immunostaining using anti-PEG and anti-PEG (methoxy group) antibodies. Further MALDI TOF analysis confirmed presence of single PEG moiety in both the isoforms. The mono-PEGylated BinA isoforms displayed preservation of the toxin structure and improved thermal stability as revealed by the circular dichroism analysis. The thermal stability of tertiary structure improved by about 3–5 ^{o}C, as was evident from thermal denaturation studies by thermofluor shift assay. The observed stability in the tertiary structure was found to be independent of the site of PEG conjugation. Notably, PEGylation enhanced BinA toxicity by nearly 6-fold. The PEGylated BinA isoforms alone, without BinB, displayed high larvicidal activity (LC_{50} value of ~3.4 ng/mL) against 3rd instar *Culex* larvae, which compares favorably against the LC_{50} value of ~5 ng/ml observed for an equimolar mixture of activated BinA and BinB proteins. This is the first study probing the effect of PEGylation on the activity of a safe and world-wide used *L. sphaericus* binary toxin. BinA can be synthesized easily through recombinant technology and easily PEGylated. Taken together, these studies suggest that the use of PEGylated BinA alone, without BinB, may prove to be a more effective biological control agent. This is in-line with WHO recommendations on improvement and efficient implementation of mosquito control interventions to control the serious health risk posed on the society by the outbreak of mosquito-borne diseases.

The main conclusions from the present study are discussed in the **Chapter 7**, which also summarizes some key aspects in understanding the molecular mechanism of action of *L. sphaericus* binary toxin. Since *L. sphaericus* spores are used widely for mosquito control and are not toxic to non-target organisms, ensuring that pure BinAB is used in mosquito control programs is essential. The present work clearly establishes the identity of the high molecular

weight complex from the spore crystals, which co-purifies with binary toxin, as the surface layer protein (SlpC) of *L. sphaericus*. As no homologous *slp*C-like gene is detected in *E. coli* genome, *L. sphaericus* BinA protein produced in large and pure form using *E. coli* expression system can be employed for mosquito-control programs. The present thesis establishes for the first time that Cqm1 is an amylomaltase with high glycosyltransferase activity. This property and the observation that Cqm1 can hydrolyze glycogen and sucrose is suggestive of its role in carbohydrate metabolism in mosquitoes. Further, it has been demonstrated here that BinA binds the diverse glycosylated proteins with high affinity. The cytotoxicity of BinA may be due to its binding to partially folded or functionally critical glycosylated proteins. N-glycosylation has been noted to be an essential post-translational modification for larval metamorphosis and development. Disruption of these processes through RNAi or chemical inhibitors has been found to be lethal to larvae (**15**). Further, BinA may also be expected to bind strongly and non-specifically to the glycosylated receptor protein (Aam1) of *Aedes*. The non-specific high-affinity binding to Aam1 may inhibit internalization of BinA rendering *Aedes* larvae refractory to binary toxin. Introducing mutations in the carbohydrate binding domain of BinA, to improve larvicidal activity and broaden specificity of the toxin, might compromise its activity against already known targets. In such a scenario, chemical modification of BinA toxin and/or construction of Ls and Bti chimeric strains (**17**) appears to be more favorable approach for toxin improvisation. PEGylation has proven to be a potent tool for chemical modifications of drugs in real-life applications. FDA approval of the use of PEG in drug development has made it a popular method of choice. The present work reports first attempt to improve the toxicity of BinA protein through PEGylation with a rewarding 6-fold increase in its larvicidal activity (**16**). More of such innovative approaches are needed to improve the target specificity of such efficient insecticides to combat mosquito borne

diseases. The present thesis lays down stepping stones in elucidating the mechanism of action of binary toxin and for toxin modification to achieve higher toxicity.

References:

1. Berry C. *J. Invertebr. Pathol.,* **2012**, 109:1-10.
2. Silva-Filha MH, Nielsen-LeRoux C, Charles JF. *Insect Biochem. Mol. Biol*., **1999**, 29:711-721.
3. Opota O, Gauthier NC, Doye A, Berry C, Gounon P, Lemichez E, Pauron D. *PLoS One*, **2011**, 6: e14682.
4. Hire RS, Hadapad AB, Dongre TK, Kumar V. *J. Invertebr. Pathol*., **2009**, 101:106-111.
5. Kale A, Hire RS, Hadapad AB, D'Souza SF, Kumar V. *Insect Biochem. Mol. Biol*., **2013**, 43:1045-1054.
6. Park HW, Bideshi DK, Federici BA. *Appl. Environ. Microbiol.*, **2003**, 69:1331-1334.
7. Colletier JP, Sawaya MR, Gingery M, et al. *Nature*, **2016**, 539:43-47.
8. Smith AW, Cámara-Artigas A, Brune DC, Allen JP. *J. Invertebr. Pathol*., **2005**, 88:27-33.
9. Hire RS, Sharma M, Hadapad AB, Kumar V. *J. Invertebr. Pathol.*, **2014**, 122:44-47.
10. Bowditch RD, Baumann P, Yousten AA. *J. Bacteriol*. **1989**, 171:4178-4188.
11. Ferreira LM, Romão TP, de-Melo-Neto OP, Silva-Filha MH. *Insect Biochem. Mol. Biol*., **2010**, 40:604-610.
12. Lekakarn H, Promdonkoy B, Boonserm P. *J. Invertebr. Pathol*., **2015**, 132:125-131.
13. Broadwell AH, Baumann P. *Appl Environ Microbiol.*, **1987**, 53:1333-1337.
14. MacKenzie CR, Hirama T, Buckley JT. *J. Biol. Chem*., **1999**, 274:22604-22609.
15. Walski T, De Schutter K, Van Damme EJ, Smagghe G. *Insect Biochem. Mol. Biol*., **2017**, 83:21-34.
16. Sharma M, Hire R, Hadapad AB, Gupta GD, Kumar V. *Bioconjugate Chemistry,* **2017**, 28:410–418.
17. Bideshi DK, Park HW, Hice RH, Wirth MC, Federici BA. Sci Rep. **2017,** 7:11282.

Publications in Refereed Journal:

a. Published

- **Sharma M**, Hire R, Hadapad AB, Gupta GD, Kumar V. PEGylation enhances mosquito-larvicidal activity of *Lysinibacillus sphaericus* Binary toxin. *Bioconjugate Chemistry,* **2017,** 28:410–418.
- **Sharma M**, Hire R, Hadapad AB, Gupta GD, Kumar V. Polyethylene glycol conjugation enhances mosquito-larvicidal activity of *Lysinibacillus sphaericus* BinA protein. *Biophysical Journal,* **2017,** 112:49a.
- Hire R, **Sharma M**, Hadapad AB, Kumar V. An oligomeric complex of BinA/BinB is not formed in-situ in mosquito-larvicidal Lysinibacillus sphaericus ISPC-8. *J. Invertebr. Pathol*., **2014**, 122:44–47.

b. Accepted:

c. Communicated:

- Sharma M, Gupta GD, Kumar V. Receptor protein of *Lysinibacillus sphaericus* mosquito-larvicidal toxin displays amylomaltase activity. *Insect Biochem. Mol. Biol.,* **2017** (*Under revision*).
- Sharma M, Gupta GD, Kumar V. Mosquito-larvicidal BinA toxin binds glycosylated proteins with high affinity. *(Under preparation)*

Other Publications:

a. Book/Book Chapter

b. Conference/Symposium

- Sharma M. et al, Characterization of carbohydrate binding domain of BinA protein from *Lysinibacillus sphaericus*. **Indo-US conference: Advances in Enzymology: Implications in health, physics and therapeutics**. 17–19 Jan. 2017, ACTREC, Mumbai, India.
- Sharma M. et al, Polyethylene glycol conjugation enhances mosquito-larvicidal activity of *Lysinibacillus sphaericus* BinA protein. **61st Annual meeting of Biophysical Society**. 11–15 Feb. 2017, New Orleans, Louisiana, USA.
- Sharma M. et al, *Lysinibacillus sphaericus* BinA protein: insight into its mode of action. **Annual Symposium of the Indian Biophysical Society**. 23-25 March 2017, IISER Mohali, Punjab, India.

ABBREVIATIONS

aa	:	Amino acid residues
AEC	:	Anion exchange chromatography
Amp	:	Ampicillin
APS	:	Ammonium per sulphate
BCIP	:	5-bromo-4-chloro-3-indolyl phosphate
bp	:	Base pair
BSA	:	Bovine serum albumin
Cam	:	Chloramphenicol
CD	:	Circular dichroism
CDS	:	Coding DNA sequence
CV	:	Column volume
Cys	:	Cysteine
D/W	:	Distilled Water
deg	:	Degrees
DLS	:	Dynamic light scattering
dmol	:	Decimole
dNTP	:	Deoxyribonucleoside triphosphate
DSF	:	Differential scanning fluorimetry
DTNB	:	5,5'-dithio-bis-(2-nitrobenzoic acid)
DTT	:	Dithiothreitol
EDTA	:	Ethylenediaminetetraacetic acid
EGTA	:	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid
FTIR	:	Fourier-transform infrared spectroscopy
g	:	Gram
GPI	:	Glycosylphosphotidylinositol
hr	:	Hour
IMAC	:	Immobilized metal ion affinity chromatography
IPTG	:	Isopropyl β-thiogalactopyranoside
ITC	:	Isothermal titration calorimetry
Kan	:	Kanamycin
kb	:	Kilo base pairs
*K*D	:	Equilibrium dissociation constant
kDa	:	Kilo-Dalton
L	:	Litre
LB	:	Luria Bertani broth
LIC	:	Ligation independent cloning

M	:	Molar
MALDI-MS	:	Matrix-assisted laser desorption/ionization-Mass spectrometry
mg	:	Milligrams
min	:	Minutes
mL	:	Milliliter
mM	:	Millimolar
Mr	:	Molecular weight
MWM	:	Molecular weight marker
NBT	:	Nitro-blue tetrazolium
nM	:	Nanomolar
O/N	:	Over night
OD	:	Optical Density
ORF	:	Open reading frame
PAGE	:	Polyacrylamide gel electrophoresis
PCR	:	Polymerase Chain Reactions
PEG	:	Polyethylene glycol
Pi	:	Inorganic phosphate
pI	:	Iso-electric point
pmol	:	Picomole
PMSF	:	Phenylmethylsulfonyl fluoride
*p*NPP	:	*p*-Nitro phenyl phosphate
RCF	:	Relative centrifugal force
RMSD	:	Root mean square deviation
rpm	:	Rotation per minutes
SDS	:	Sodium dodecyl sulfate
SEC	:	Size exclusion chromatography
sec	:	Second
SPR	:	Surface plasmon resonance
TAE	:	Tris-Acetate-EDTA
TE	:	Tris-EDTA
Tet	:	Tetracycline
Trp	:	Tryptophan
α-pNPG	:	4-nitrophenyl α-D-glucopyranoside
β-pNPG	:	4-nitrophenyl β-D-glucopyranoside
µg	:	Microgram
µL	:	Microlitre
µM	:	Micromolar

List of Figures

List of Tables

Chapter 1

Introduction

Mosquito-borne diseases

Vectors are living organisms that transmit infectious disease-causing pathogens. Vector-borne diseases account for more than 17% of all infectious diseases (http://www.who.int). Of all disease-transmitting vectors, mosquitoes cause the greatest menace. Their ability to carry and spread diseases to humans causes millions of deaths every year. *Aedes*, *Anopheles* and *Culex* are the most common examples of mosquitoes found worldwide. They transmit pathogens ranging from parasites to viruses and are responsible for spreading many debilitating human diseases. For instance, *Aedes aegypti* transmits Dengue, Zika, Chikungunya; *Anopheles* transmits malaria; while *Culex* transmits West Nile fever, encephalitis and filariasis. Dengue is the most common mosquito-borne viral disease, with 50– 100 million cases worldwide each year in more than 100 countries, with 3 billion people living in dengue endemic countries [1]. Zika is an emerging threat and was declared a global public health emergency by World Health Organization. Malaria is endemic in 91 countries, with about 40% of the world population at risk (http://www.who.int). It causes more than 400,000 deaths every year globally. In 2016, there were an estimated 216 million cases of malaria, an increase of about 5 million cases over 2015 [2]. In India, mosquito-borne diseases have been on the rise across the country, with chikungunya recording 300% increase in cases between 2012– 2015. Meanwhile, dengue cases increased 157%, followed by Japanese encephalitis cases with 124% rise, according to **National Health Profile 2017** report. Steep rise in the incidence of these mosquito-borne diseases also results in a huge socio-economic burden on the annual income of the country in terms of loss of production and medical costs.

Various steps are being taken to fight mosquito-borne diseases. However, the impact of such steps is slow. In the absence of any effective vaccines or drugs against these debilitating diseases, their prevention and control depends exclusively on effective vector

control to minimize the risk of transmission that can be achieved by controlling the mosquito vector population or interruption of human-vector contact. Vector control approaches are generally divided into three categories: environmental, chemical and biological management. Environmental management seeks to transform the surroundings to prevent or minimize vector propagation and human contact with the vector-pathogen by destroying egg/ larval/ pupal habitats, use of mosquito nets, etc. Use of chemical larvicide, fogging etc. comes under chemical management. They proved satisfactory initially, however, the persistence of residual chemical insecticides pollutes the environment and causes development of resistance in mosquitoes, hence challenging their efficiency in the long run. This demands development of novel and alternative approaches which are safe for the environment and sustainable.

Biological approach for vector management

Biological control strategies are environmental friendly and use predatory or parasitic organisms which target different stages of the mosquito life-cycle and reduce the populations by killing the mosquitoes, exploiting mosquito behaviour to improve mosquito mortality, and releasing mosquitoes that are either sterile or unable to transmit disease without adversely affecting the ecosystem.

The history of using microorganisms for biological control dates to the 1960s, when entomopathogenic fungus was first reported to be larvicidal for malaria vector *Anopheles gambiae* [3]. Of the various microorganisms known as potential agents for vector control, bacteria continue to be the most potent for further development as larvicide [4]. Two important bacteria, *Bacillus thuringiensis* subsp. *israelensis* (Bti) and *Lysinibacillus sphaericus* (*Ls*), which produce insecticidal toxins, have proven very successful in mosquito control since their discovery and have been extensively commercialized. They have proven as a successful alternate to chemical insecticides.

***Bacillus thuringiensis* subsp. *israelensis* (Bti)**

Bti is a gram positive, aerobic, spore forming bacterium with larvicidal activity against Dipterans, primarily mosquitoes and blackflies. Bti has been the most successful larvicide for mosquito control and has been in use for more than 30 years. Its larvicidal activity is attributed to δ-endotoxins produced as spherical parasporal inclusion bodies during sporulation (Fig. 1.1) [5]. The parasporal bodies are rich in four larvicidal proteins of primary importance, three crystalline proteins namely Cry11A, Cry4A, Cry4B and a cytosolic protein Cyt1A. Of these, the cytolytic toxin, Cyt1A (27 kDa), seems to be the key factor, as absence of Cyt1A in the toxin complex enables resistance to evolve [6]. Cyt1A adopts a typical cytolysin fold containing a β-sheet held by two surrounding α-helical layers (Fig. 1.2) and its toxicity effect is explained based on the pore-forming model [7].

Both the Cry and Cyt proteins are produced as protoxins which are solubilized inside the larval midgut and activated by midgut proteases [8]. Active Cry toxins recognize specific GPI anchored receptors in the midgut microvilli [8] while, the active Cyt toxin interacts directly with the membrane lipids and resulting in either pore formation [9] or membrane disruption through detergent like interaction [10]. Cry and Cyt toxins act synergistically and show maximum toxicity in combination. However, Bti has been found to be relatively ineffective in polluted water and its residual activity in most habitats is limited to a few days after treatment.

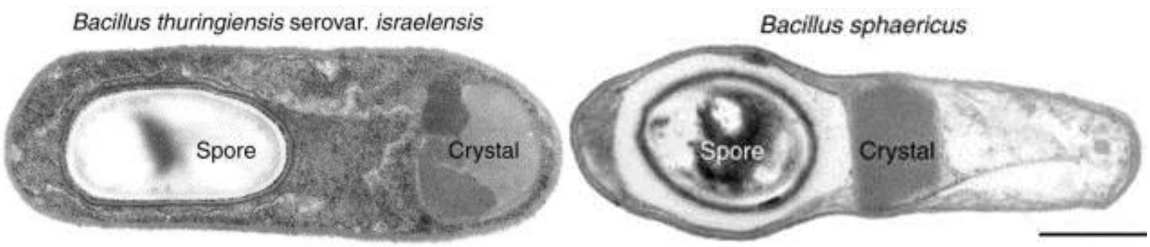


Figure 1.1. Parasporal crystalline bodies of Bti and Ls. [11]

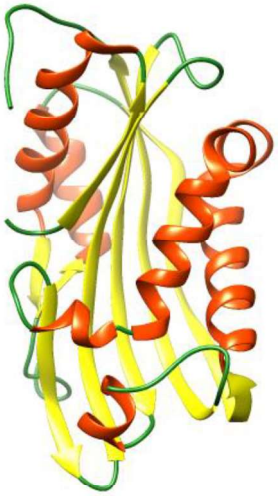


Figure 1.2. Ribbon model of Cyt1A protein, PDB ID: 3RON [7], drawn using Chimera suite [12].

***Lysinibacillus sphaericus* (*Ls*)**

Lysinibacillus sphaericus is a gram positive, aerobic, spore forming bacterium found in a variety of soil and aquatic environments [13]. It was isolated for the first time from fourth-instar larvae of *Culiseta incidens* near Fresno, California [14]. It is characterized by having a spherical terminal spore and by its inability to utilize carbohydrates, except N-acetylglucosamine [15].

Formerly known as *Bacillus sphaericus*, this species was reclassified to a new genus *Lysinibacillus* based on phenotypic traits, mainly on the differences in peptidoglycan composition consisting of lysine and aspartic acid instead of *meso*-diaminopimelic acid, the major component of *Bacillus* cell wall [16]. Ls is known for its high larvicidal activity against *Culex* and *Anopheles* mosquito larvae, but, poor or no toxicity against larvae of *Aedes aegypti.* It was not until the 1970s that the first strains with potential use as mosquito-control agents were discovered [17]. Among the various strains of Ls only few are toxic against mosquitoes. Some of the well-known larvicidal strains examined so far include,

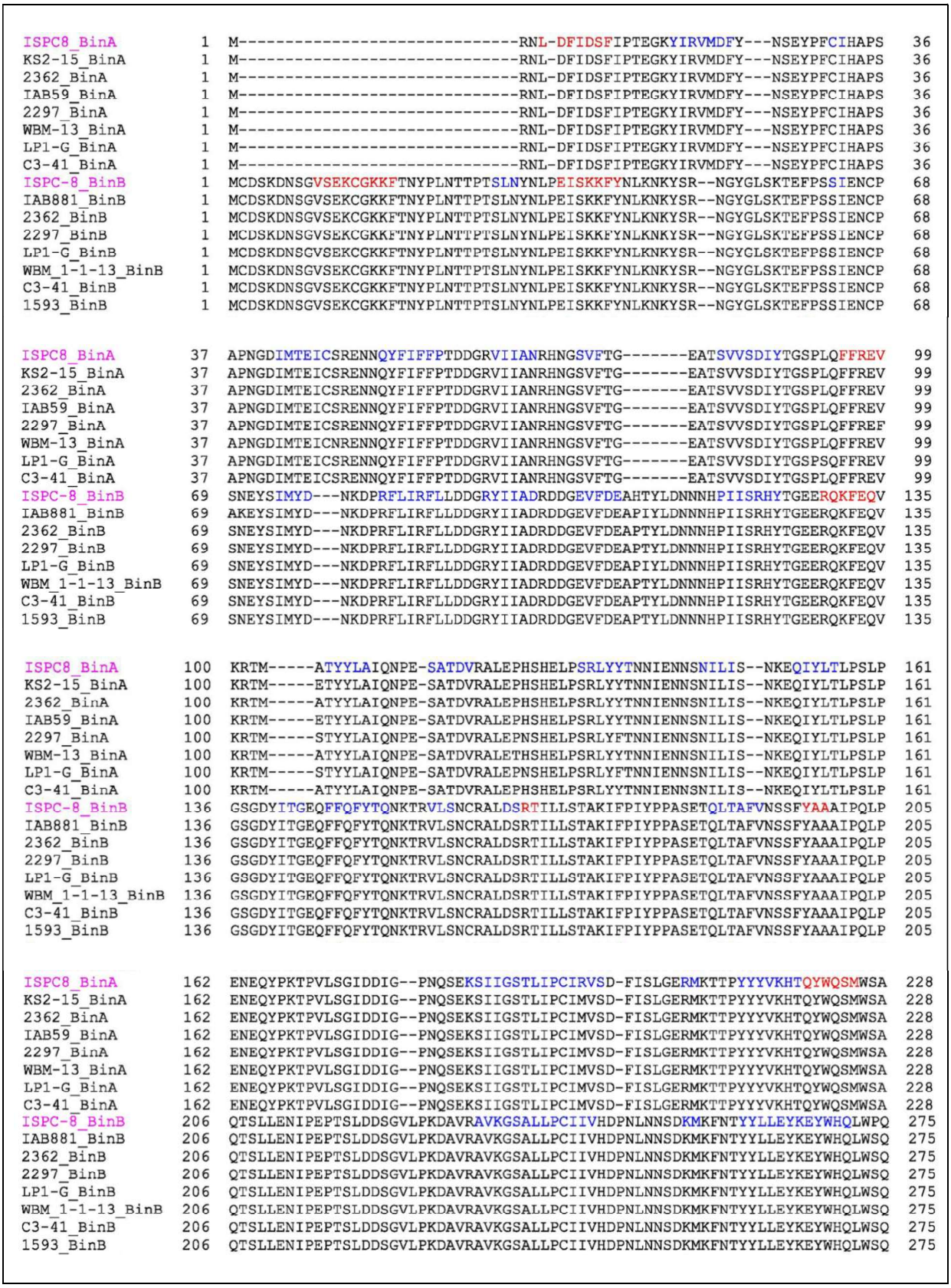
1593, 2297 and 2362 isolated from Indonesia, Sri Lanka and Nigeria, respectively. The bacterium has been evaluated by World Health Organization for its high efficacy in mosquito control [13, 18]. It has been found to be highly specific with no adverse effects on non-target organisms, like honeybees, other mosquito predators, chironomids and other eukaryotic organism, and is considered as environmental-friendly [19−21]. Ls strains have been found to be relatively more effective in polluted water unlike Bti, especially against *Culex* and *Anopheles*, and are found to possess longer residual activity due to longer persistence or recycling [18].

The pathogenic strains of Ls produce many insecticidal protein toxins, like sphaericolysin [22], Mtx toxins (Mtx1 and Mtx2) and others. However, the primary agent responsible for its high larvicidal activity is the binary toxin produced as parasporal crystalline inclusions (Fig. 1.1) during the early stages of sporulation. Another two component toxin is found to exist in pathogenic strains of Ls. One of the components is the Cry48Aa1 protein, which is closely related to the mosquitocidal Cry toxins of Bti, and belongs to the 3-domain Cry toxin family. The second is the Cry49Aa1 protein, which is related to the binary toxin of Ls and to Cry35 and Cry36 of Bti [23]. The two components show high level toxicity only in combination. However, their target range is very narrowly limited to the genus *Culex* [24].

Binary toxin

Binary (BinAB) toxin, produced by toxic Ls strains, is composed of two highly conserved protoxins, pro-BinA (41.9 kDa) and pro-BinB (51.4 kDa) proteins, encoded by two highly conserved genes, *bin*A and *bin*B [25, 26] (Fig. 1.3). Once ingested by susceptible larvae, the crystalline inclusions are solubilised in the alkaline pH of larval midgut and the protoxins are activated by midgut proteases [27]. Active proteins (BinA, 39 kDa; BinB, 42 kDa) bind to their specific receptors in the midgut microvilli of mosquito larvae [28]. BinB is the receptor binding component while BinA confers toxicity. Their specific receptor has been identified as a 66 kDa glycosylphosphatidylinositol (GPI) anchored α-glucosidase, named Cpm1 and Cqm1 from *Culex pipiens* and *Culex quinquefasciatus*, respectively, located on the apical membrane of the midgut epithelium cells [28, 29]. Presentation of the receptor is essential for toxin activity. The two protein components of binary toxin act synergistically and display maximum larvicidal activity at equimolar concentrations [30, 31]. However, BinA and PEGylated BinA, alone without BinB, also show high larvicidal activity, although at relatively high doses [32−34].

It is generally believed that a heterotetramer, BinA2.BinB2, is formed at the receptor that translocates inside the cell [35, 36, 37]. Existence of a preformed stable oligomer of binary toxin inside the spores of *L. sphaericus* [34] has been debated by the scientific community. Such a pre-formed oligomer was not reported for the recombinant proteins when co-expressed in *E. coli*, or in a stoichiometric mixture of recombinant proteins *in vitro* [32]. But a covalent complex of BinA and BinB, synthesized by glutaraldehyde crosslinking, showed maximal activity reported for any stoichiometric ratio of the two component proteins [32].


ISPC8_BinA 1 M----------RNL-DFIDSFIPTEGKYIRVMDFY---NSEYPFCIHAPS 36
KS2-15_BinA 1 M----------RNL-DFIDSFIPTEGKYIRVMDFY---NSEYPFCIHAPS 36
2362_BinA 1 M----------RNL-DFIDSFIPTEGKYIRVMDFY---NSEYPFCIHAPS 36
IAB59_BinA 1 M----------RNL-DFIDSFIPTEGKYIRVMDFY---NSEYPFCIHAPS 36
2297_BinA 1 M----------RNL-DFIDSFIPTEGKYIRVMDFY---NSEYPFCIHAPS 36
WBM-13_BinA 1 M----------RNL-DFIDSFIPTEGKYIRVMDFY---NSEYPFCIHAPS 36
LP1-G_BinA 1 M----------RNL-DFIDSFIPTEGKYIRVMDFY---NSEYPFCIHAPS 36
C3-41_BinA 1 M----------RNL-DFIDSFIPTEGKYIRVMDFY---NSEYPFCIHAPS 36
ISPC-8_BinB 1 MCDSKDNSGVSEKCGKKFTNYPLNTTPTSLNYNLPEISKKFYNLKNKYSR--NGYGLSKTEFPSSIENCP 68
IAB881_BinB 1 MCDSKDNSGVSEKCGKKFTNYPLNTTPTSLNYNLPEISKKFYNLKNKYSR--NGYGLSKTEFPSSIENCP 68
2362_BinB 1 MCDSKDNSGVSEKCGKKFTNYPLNTTPTSLNYNLPEISKKFYNLKNKYSR--NGYGLSKTEFPSSIENCP 68
2297_BinB 1 MCDSKDNSGVSEKCGKKFTNYPLNTTPTSLNYNLPEISKKFYNLKNKYSR--NGYGLSKTEFPSSIENCP 68
LP1-G_BinB 1 MCDSKDNSGVSEKCGKKFTNYPLNTTPTSLNYNLPEISKKFYNLKNKYSR--NGYGLSKTEFPSSIENCP 68
WBM_1-1-13_BinB 1 MCDSKDNSGVSEKCGKKFTNYPLNTTPTSLNYNLPEISKKFYNLKNKYSR--NGYGLSKTEFPSSIENCP 68
C3-41_BinB 1 MCDSKDNSGVSEKCGKKFTNYPLNTTPTSLNYNLPEISKKFYNLKNKYSR--NGYGLSKTEFPSSIENCP 68
1593_BinB 1 MCDSKDNSGVSEKCGKKFTNYPLNTTPTSLNYNLPEISKKFYNLKNKYSR--NGYGLSKTEFPSSIENCP 68
ISPC8_BinA 37 APNGDIMTEICSRENNQYFIFFPTDDGRVIIANRHNGSVFTG-------EATSVVSDIYTGSPLQFFREV 99
KS2-15_BinA 37 APNGDIMTEICSRENNQYFIFFPTDDGRVIIANRHNGSVFTG-------EATSVVSDIYTGSPLQFFREV 99
2362_BinA 37 APNGDIMTEICSRENNQYFIFFPTDDGRVIIANRHNGSVFTG-------EATSVVSDIYTGSPLQFFREV 99
IAB59_BinA 37 APNGDIMTEICSRENNQYFIFFPTDDGRVIIANRHNGSVFTG-------EATSVVSDIYTGSPLQFFREV 99
2297_BinA 37 APNGDIMTEICSRENNQYFIFFPTDDGRVIIANRHNGSVFTG-------EATSVVSDIYTGSPLQFFREF 99
WBM-13_BinA 37 APNGDIMTEICNRENNQYFIFFPTDDGRVIIANRHNGSVFTG-------EATSVVSDIYTGSPLQFFREV 99
LP1-G_BinA 37 APNGDIMTEICSRENNQYFIFFPTDDGRVIIANRHNGSVFTG-------EATSVVSDIYTGSPSQFFREV 99
C3-41_BinA 37 APNGDIMTEICSRENNQYFIFFPTDDGRVIIANRHNGSVFTG-------EATSVVSDIYTGSPLQFFREV 99
ISPC-8_BinB 69 SNEYSIMYD---NKDPRFLIRFLLDDGRYIIADRDDGEVFDEAHTYLDNNNHPIISRHYTGEERQKFEQV 135
IAB881_BinB 69 AKEYSIMYD---NKDPRFLIRFLLDDGRYIIADRDDGEVFDEAPIYLDNNNHPIISRHYTGEERQKFEQV 135
2362_BinB 69 SNEYSIMYD---NKDPRFLIRFLLDDGRYIIADRDDGEVFDEAPTYLDNNNHPIISRHYTGEERQKFEQV 135
2297_BinB 69 SNEYSIMYD---NKDPRFLIRFLLDDGRYIIADRDDGEVFDEAPTYLDNNNHPIISRHYTGEERQKFEQV 135
LP1-G_BinB 69 SNEYSIMYD---NKDPRFLIRFLLDDGRYIIADRDDGEVFDEAPIYLDNNNHPIISRHYTGEERQKFEQV 135
WBM_1-1-13_BinB 69 SNEYSIMYD---NKDPRFLIRFLLDDGRYIIADRDDGEVFDEAPTYLDNNNHPIISRHYTGEERQKFEQV 135
C3-41_BinB 69 SNEYSIMYD---NKDPRFLIRFLLDDGRYIIADRDDGEVFDEAPTYLDNNNHPIISRHYTGEERQKFEQV 135
1593_BinB 69 SNEYSIMYD---NKDPRFLIRFLLDDGRYIIADRDDGEVFDEAPTYLDNNNHPIISRHYTGEERQKFEQV 135
ISPC8_BinA 100 KRTM-----ATYYLAIQNPE-SATDVRALEPHSHELPSRLYYTNNIENNSNILIS--NKEQIYLTLPSLP 161
KS2-15_BinA 100 KRTM-----ETYYLAIQNPE-SATDVRALEPHSHELPSRLYYTNNIENNSNILIS--NKEQIYLTLPSLP 161
2362_BinA 100 KRTM-----ATYYLAIQNPE-SATDVRALEPHSHELPSRLYYTNNIENNSNILIS--NKEQIYLTLPSLP 161
IAB59_BinA 100 KRTM-----ETYYLAIQNPE-SATDVRALEPHSHELPSRLYYTNNIENNSNILIS--NKEQIYLTLPSLP 161
2297_BinA 100 KRTM-----STYYLAIQNPE-SATDVRALEPNSHELPSRLYFTNNIENNSNILIS--NKEQIYLTLPSLP 161
WBM-13_BinA 100 KRTM-----ATYYLAIQNPE-SATDVRALETHSHELPSRLYYTNNIENNSNILIS--NKEQIYLTLPSLP 161
LP1-G_BinA 100 KRTM-----STYYLAIQNPE-SATDVRALEPNSHELPSRLYFTNNIENNSNILIS--NKEQIYLTLPSLP 161
C3-41_BinA 100 KRTM-----ATYYLAIQNPE-SATDVRALEPHSHELPSRLYYTNNIENNSNILIS--NKEQIYLTLPSLP 161
ISPC-8_BinB 136 GSGDYITGEQFFQFYTQNKTRVLSNCRALDSRTILLSTAKIFPIYPPASETQLTAFVNSSFYAAAIPQLP 205
IAB881_BinB 136 GSGDYITGEQFFQFYTQNKTRVLSNCRALDSRTILLSTAKIFPIYPPASETQLTAFVNSSFYAAAIPQLP 205
2362_BinB 136 GSGDYITGEQFFQFYTQNKTRVLSNCRALDSRTILLSTAKIFPIYPPASETQLTAFVNSSFYAAAIPQLP 205
2297_BinB 136 GSGDYITGEQFFQFYTQNKTRVLSNCRALDSRTILLSTAKIFPIYPPASETQLTAFVNSSFYAAAIPQLP 205
LP1-G_BinB 136 GSGDYITGEQFFQFYTQNKTRVLSNCRALDSRTILLSTAKIFPIYPPASETQLTAFVNSSFYAAAIPQLP 205
WBM_1-1-13_BinB 136 GSGDYITGEQFFQFYTQNKTRVLSNCRALDSRTILLSTAKIFPIYPPASETQLTAFVNSSFYAAAIPQLP 205
C3-41_BinB 136 GSGDYITGEQFFQFYTQNKTRVLSNCRALDSRTILLSTAKIFPIYPPASETQLTAFVNSSFYAAAIPQLP 205
1593_BinB 136 GSGDYITGEQFFQFYTQNKTRVLSNCRALDSRTILLSTAKIFPIYPPASETQLTAFVNSSFYAAAIPQLP 205
ISPC8_BinA 162 ENEQYPKTPVLSGIDDIG--PNQSEKSIIGSTLIPCIRVSD-FISLGERMKTTPYYYVKHTQYWQSMWSA 228
KS2-15_BinA 162 ENEQYPKTPVLSGIDDIG--PNQSEKSIIGSTLIPCIMVSD-FISLGERMKTTPYYYVKHTQYWQSMWSA 228
2362_BinA 162 ENEQYPKTPVLSGIDDIG--PNQSEKSIIGSTLIPCIMVSD-FISLGERMKTTPYYYVKHTQYWQSMWSA 228
IAB59_BinA 162 ENEQYPKTPVLSGIDDIG--PNQSEKSIIGSTLIPCIMVSD-FISLGERMKTTPYYYVKHTQYWQSMWSA 228
2297_BinA 162 ENEQYPKTPVLSGIDDIG--PNQSEKSIIGSTLIPCIMVSD-FISLGERMKTTPYYYVKHTQYWQSMWSA 228
WBM-13_BinA 162 ENEQYPKTPVLSGIDDIG--PNQSEKSIIGSTLIPCIMVSD-FISLGERMKTTPYYYVKHTQYWQSMWSA 228
LP1-G_BinA 162 ENEQYPKTPVLSGIDDIG--PNQSEKSIIGSTLIPCIMVSD-FISLGERMKTTPYYYVKHTQYWQSMWSA 228
C3-41_BinA 162 ENEQYPKTPVLSGIDDIG--PNQSEKSIIGSTLIPCIMVSD-FISLGERMKTTPYYYVKHTQYWQSMWSA 228
ISPC-8_BinB 206 QTSLLENIPEPTSLDDSGVLPKDAVRAVKGSALLPCIIVHDPNLNNSDKMKFNTYYLLEYKEYWHQLWPQ 275
IAB881_BinB 206 QTSLLENIPEPTSLDDSGVLPKDAVRAVKGSALLPCIIVHDPNLNNSDKMKFNTYYLLEYKEYWHQLWSQ 275
2362_BinB 206 QTSLLENIPEPTSLDDSGVLPKDAVRAVKGSALLPCIIVHDPNLNNSDKMKFNTYYLLEYKEYWHQLWSQ 275
2297_BinB 206 QTSLLENIPEPTSLDDSGVLPKDAVRAVKGSALLPCIIVHDPNLNNSDKMKFNTYYLLEYKEYWHQLWSQ 275
LP1-G_BinB 206 QTSLLENIPEPTSLDDSGVLPKDAVRAVKGSALLPCIIVHDPNLNNSDKMKFNTYYLLEYKEYWHQLWSQ 275
WBM_1-1-13_BinB 206 QTSLLENIPEPTSLDDSGVLPKDAVRAVKGSALLPCIIVHDPNLNNSDKMKFNTYYLLEYKEYWHQLWSQ 275
C3-41_BinB 206 QTSLLENIPEPTSLDDSGVLPKDAVRAVKGSALLPCIIVHDPNLNNSDKMKFNTYYLLEYKEYWHQLWSQ 275
1593_BinB 206 QTSLLENIPEPTSLDDSGVLPKDAVRAVKGSALLPCIIVHDPNLNNSDKMKFNTYYLLEYKEYWHQLWSQ 275

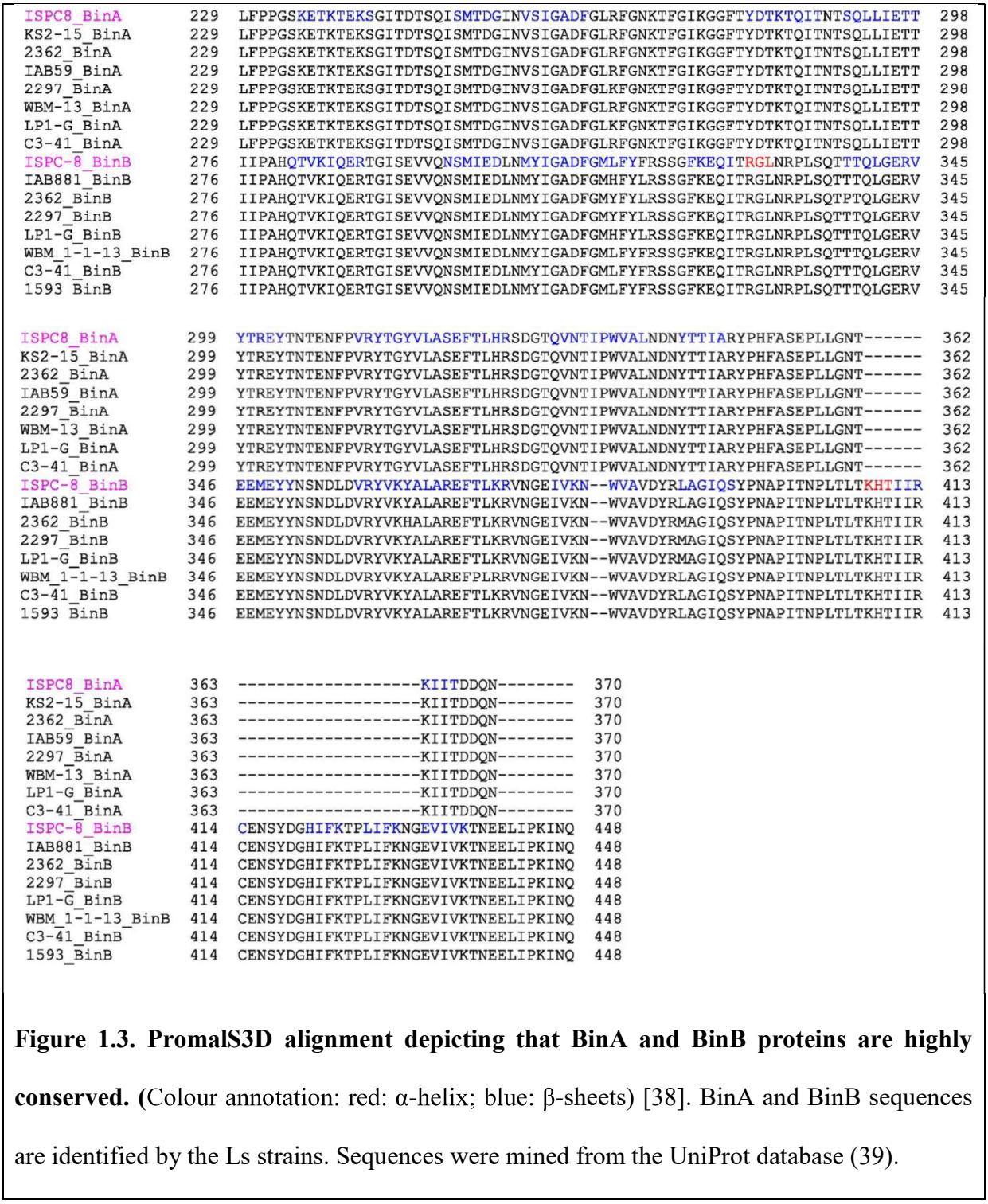


Figure 1.3. PromalS3D alignment depicting that BinA and BinB proteins are highly conserved. (Colour annotation: red: α-helix; blue: β-sheets) [38]. BinA and BinB sequences are identified by the Ls strains. Sequences were mined from the UniProt database (39).

BinAB crystal structure

The absence of structural information about BinAB components limited our understanding of the structure-activity relationship which is also essential for rationally improvising the toxin. The recently determined structure of BinAB of *L. sphaericus* revealed that both BinA and BinB subunits are comprised of an N-terminal β-trefoil domain (BinA residues 1–155, BinB residues 1–198) and a C-terminal pore-forming domain (BinA residues 156–370, BinB residues 199–448) (Fig. 1.4) [40].

The β-trefoil scaffold is highly conserved architecture of some sugar binding proteins (lectins) and the β-rich C-terminal domain shares similarity with aerolysin type β-pore forming toxins. Similar information was revealed from the crystal structure of recombinantly produced BinB [41]. The most notable structural differences between BinA and BinB are in the carbohydrate-binding modules, where all three modules (α, β and γ) of BinA appear structurally capable of binding carbohydrate; unlike BinB in which pseudo-three-fold symmetry of trefoil is distorted due to loop insertion [40]. These differences may contribute to the distinct roles of BinA and BinB in toxicity and suggests that BinB has a lesser role in carbohydrate binding. Earlier, several carbohydrates were observed to reduce BinA toxicity against mosquito cells [27]. The structure information also revealed the presence of four pH switches that facilitate crystal dissolution and release of soluble protoxins in the larval midgut [40]. A large heterodimeric interface is evident from the crystal structure that rearranges following proteolytic activation and helps in anchoring BinA to the receptor bound BinB [40].

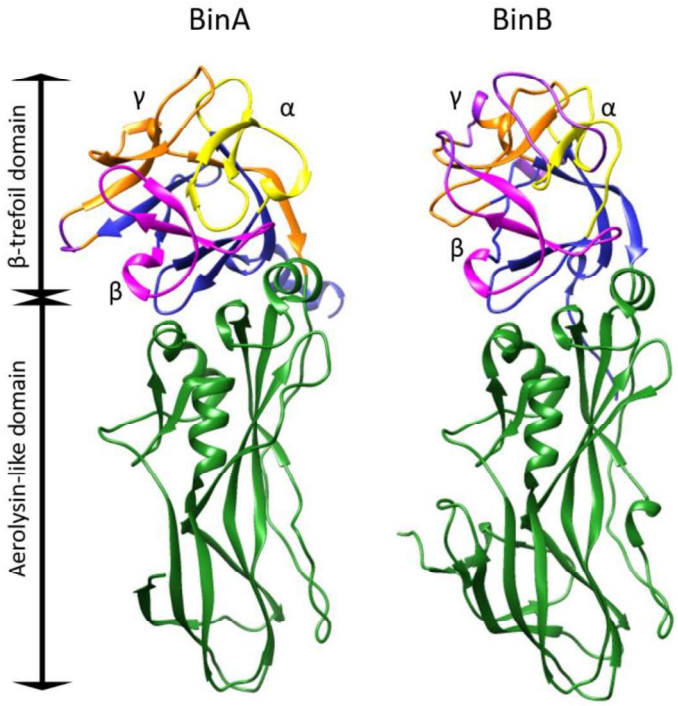


Figure 1.4. Ribbon model of BinA & BinB proteins, PDB ID: 5FOY, drawn using Chimera suite [12]. (yellow, magenta and orange : α, β and γ trefoil modules, respectively; purple : loop insertions ; blue : trefoil barrel) [40].

Carbohydrate binding module

Lectins are carbohydrate binding proteins that recognize carbohydrates, as free oligosaccharides or as glycoconjugates. These interactions are expected to play important role in biological events, like receptor recognition or adhesion processes involving lectins. Insight into the activities of lectins at the molecular level requires detailed understanding of the carbohydrate-binding module (CBM) present in these proteins. CBM is defined as a contiguous, independently folding sequence of amino acids which is involved primarily in either targeting a catalytic fragment to specific carbohydrate or concentrating catalytic fragments in proximity of a substrate [42]. CBM are recognized by their characteristic conserved protein fold such as β-sandwich, β-trefoil, Cysteine knot, Cellulose binding and OB (oligonucleotide/oligosaccharide binding) folds etc. However, presence of such protein

folds is not predictive of the function or ligand specificity of the corresponding lectin [42]. β-sandwich and β-trefoil are the most frequent folds found in lectins. β-trefoil fold contains 12 strands of β-sheet, forming six hairpin turns which arrange to form a *β*-barrel with a triangular cap on one end of the β-barrel called the 'hairpin triplet'. Each subunit of this fold (α, β and γ) is a contiguous amino acid sequence with a four β-strand, two-hairpin structure having a trefoil shape (Fig. 1.5). The β-trefoil fold has a pseudo-3-fold symmetry [43] and provides functional carbohydrate-binding sites in each of the three trefoil subdomains. The multivalency offers significantly enhanced affinities [44].

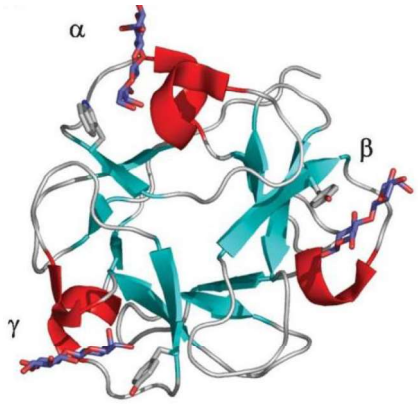


Figure 1.5. Typical β-trefoil domain (α, β and γ modules) with bound ligands. PDB ID: 1KNM. Adapted from Boraston et al., Biochem. J., 2004, 382, 769–781

Aerolysin-like pore-forming Domain

Pore forming toxins (PFTs) are one of the largest and widely distributed families among the bacterial toxins. PFTs can exist in both water-soluble and membrane-inserted forms. They are broadly classified into two groups, α-PFTs and β-PFTs, depending upon whether membrane spanning is achieved using α-helices or β-strands, respectively [45].

Aerolysin, produced by the *Aeromonas hydrophila*, is the founding member of β-pore-forming toxins (β-PFTs). It shares high sequence identity with α-toxin from *Clostridium*. Aerolysin is produced as an inactive precursor known as proaerolysin, which

is activated by proteolytic processing. The toxin then binds to its specific GPI anchored receptors. Once bound to the receptor and activated, the toxin undergoes polymerization forming ring like structures that get inserted into the membrane and form pores. An "aerolysin-like domain" or "aerolysin fold" has been defined which is highly conserved and can be found throughout all the kingdoms. In bacteria, aerolysin folds are present either as a single lobe or as a pore forming domain fused to an N-terminal lectin domain, as found in aerolysin (Fig. 1.6). Aerolysin binds to its receptor through interaction with the glycan core of the GPI anchor of its target receptor [46].

BinA, BinB and Cry49 from Ls and Cry35 and Cry36 from Bti belong to the Toxin_10 family of proteins (Pfam accession number: PF05431). The recently published crystal structures of Cry35Ab1 [47] and BinAB [40] have shown presence of aerolysin-like fold in these proteins. This shows that the Toxin_10 family of proteins may also represent β pore-forming toxins.

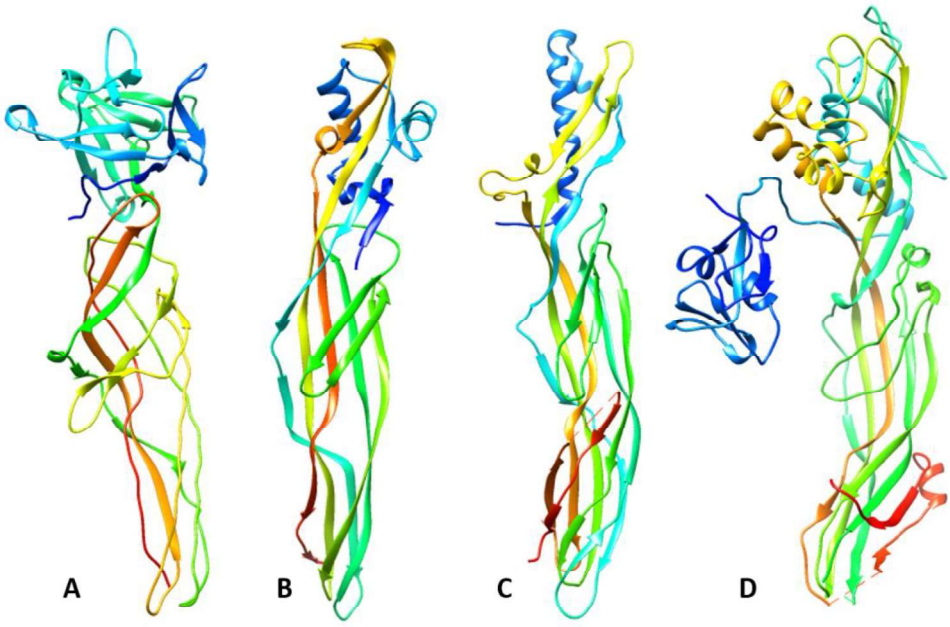


Figure 1.6. Conserved common core for aerolysin-like β-PFTs. A) Hemolytic lectin, PDB ID: 1W3F; **B)** Parasporin, PDB ID: 2ZTB; **C)** Epsilon-toxin, PDB ID: 1UYJ; **D)** Proaerolysin, PDB ID: 1PRE [48]. Ribbon models drawn using Chimera suite [12].

Glycosylphosphatidylinositol (GPI) anchor

Glycosylphosphatidylinositol (GPI) anchor is a glycolipid structure and represents a C-terminal post-translational modification to attach the modified protein on the cell membrane. This post-translational modification is found ubiquitously in eukaryotes and in some archaea but not in eubacteria. The core of GPI anchor consists of phosphatidylinositol, glycans comprising one glucosamine and at least three mannose residues, and a terminal phosphoethanolamine, which is amide-bonded to the newly formed carboxyl terminus of the protein during the process of GPI attachment (Fig. 1.7). The glycan side-branches in the GPI backbone vary depending on the organism, cell type, and protein. GPI-anchored proteins are typically associated with membrane microdomains (lipid rafts). They often exist on the cell surface as transient homodimers and may be involved in signal transduction. For each of these properties, the lipid and the glycan moieties of the GPI anchor are crucial [49, 50].

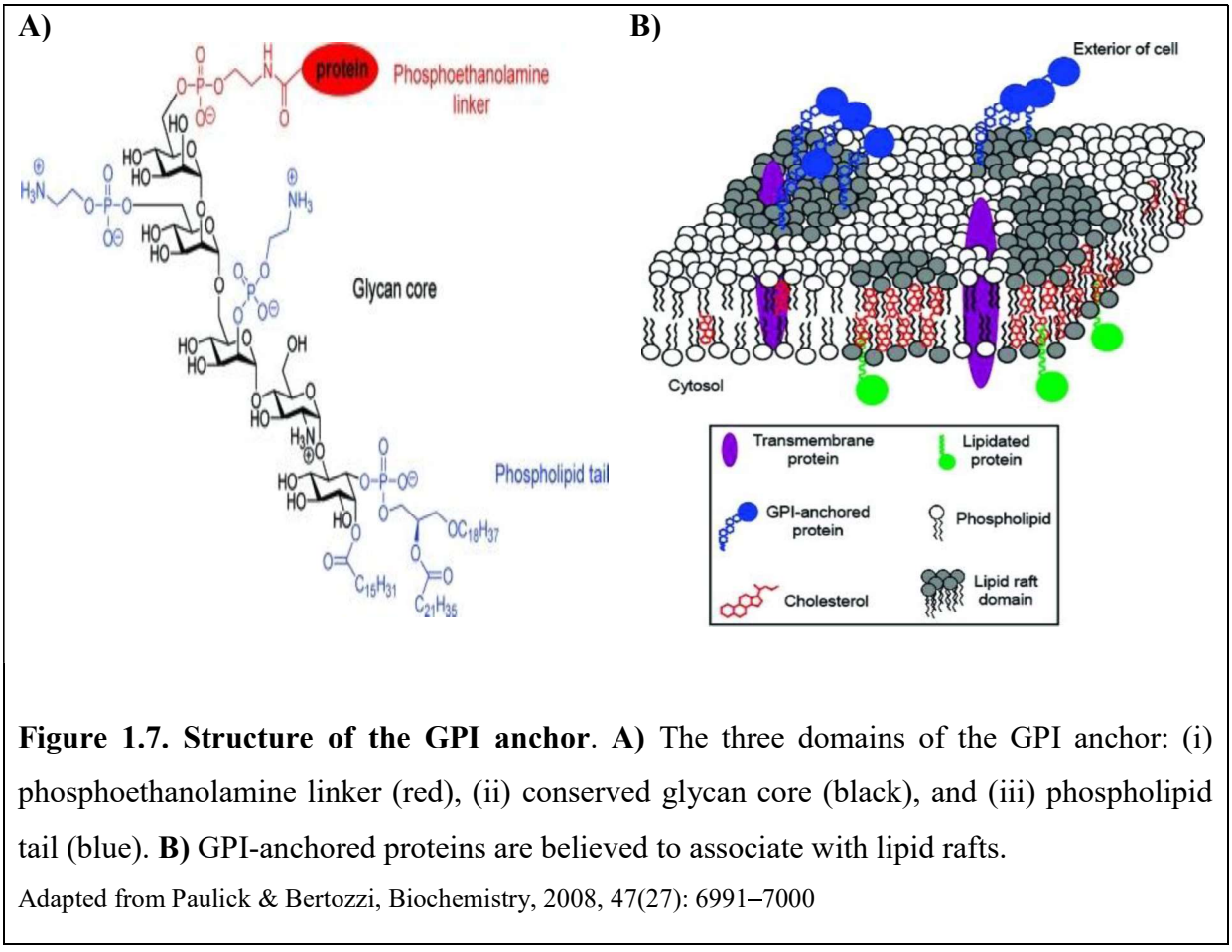


Figure 1.7. Structure of the GPI anchor. **A)** The three domains of the GPI anchor: (i) phosphoethanolamine linker (red), (ii) conserved glycan core (black), and (iii) phospholipid tail (blue). **B)** GPI-anchored proteins are believed to associate with lipid rafts.

Adapted from Paulick & Bertozzi, Biochemistry, 2008, 47(27): 6991–7000

Mode of action of BinAB toxin

Constant efforts are being made world-wide to elucidate the mode of action of binary toxin. It is accepted that a heterotetrameric complex, BinA2.BinB2, assembles on the receptor and triggers a set of cytopathological events following internalization like formation of pores in the membrane [35, 36, 37], and cytoplasmic vacuolation, and autophagy [51], resulting in larval death. A recent study has also suggested that binary toxin might trigger apoptosis via intrinsic or mitochondrial pathway, *in vivo*, contributing to larval death [52]. The study reported enhanced activity of caspase-9 and caspase-3 apoptotic proteins and linked toxicity with mitochondrial response. However, the specific mechanism of larval death remains unclear.

Refractoriness of *Aedes aegypti* to BinAB toxin

Homologous Aam1 receptor protein has been characterized from *Aedes aegypti*, which is refractory to BinAB toxin, and displays 74% identity to Cqm1 (Fig. 1.8) [53]. Similarly, Agm3 has been characterized in *Anopheles gambiae,* which is moderately susceptible to BinAB toxin, and displays 67% identity to Cqm1 (Fig. 1.8) [54, 55].

Many attempts have been made to understand the refractoriness of *A. aegypti* to binary toxin. Initial reports by Nielsen-Leroux and Charles [56] ruled out the possibility of crystal processing as the reason for the resistance. Later it was found that Aam1 is also expressed as GPI anchored protein in the larval stage, but unlike Cqm1, it was glycosylated [53]. However, carbohydrate removal from Aam1 did not affect its binding properties suggesting that differential glycosylation of this protein does not interfere with potential toxin binding site [53]. Further, through extensive site directed mutagenesis it was found that $_{159}$GG$_{160}$ doublet in Cqm1, not conserved in Aam1, is likely associated with the differential binding capacities of Cqm1 and Aam1 with binary toxin [57]. A comparative study of binary toxin internalization in *Culex* and *Aedes* larvae using the fluorescently labelled BinAB

components revealed that the two proteins co-localize on the cell membrane and in cell cytoplasm in *Culex*. However, no internalization was observed when fed to *Aedes* larvae and both the component proteins were localized only on cell membrane [58].

Figure 1.8. PromalS3D alignment for Cqm1, Cpm1, Aam1 and Agm3 highlighting that these proteins are highly conserved. (Colour annotation: red: α-helix; blue: β-sheets) [38].

Developing resistance among mosquitoes against bacterial toxins: A serious concern

Bacterial toxins have proven to be efficient and promising larvicides, however, developing resistance among mosquitoes is posing a serious concern. Binary toxin is a single site directed larvicide as it recognizes a single receptor. This predisposes it to a higher risk of resistance development. The primary cause appears to be the lack of the receptor presentation in the form bound to the midgut membrane. This is the consequence of mutations in the receptor sequence leading to pre-mature translation termination resulting in the synthesis of a truncated protein that lacks the GPI anchor [59−62].

On the contrary, complex mixture of 4 diverse toxins found in Bti and their ability to recognize different receptors reduces the risk of resistance development against Bti [63]; most importantly the presence of Cyt toxin suppresses resistance against Cry proteins in mosquitoes [64].

Improving bacterial toxins to overcome developing resistance

The current focus of the scientific community is to develop more effective bacterial strains with enhanced larvicidal activity and broader target specificity as binary toxin is currently ineffective against *Aedes aegypti* which is one of the major concerns for the human health world-wide. Moreover, evolution of resistant mosquito populations questions the efficacy of existing mosquito control interventions. Different strategies are being attempted to address these issues.

- Developing recombinant bacterial strains

Attempts have been made in the past for developing recombinant bacterial strains with enhanced toxicity by fusing Cry, Cyt from Bti with binary toxins from Ls [65, 66]. In continuation with the earlier work recently a chimeric protein, Cyt1Aa-BinA (69.6 kDa) has been created by fusing Cyt1Aa protoxin from Bti to the BinA protoxin from Ls. Cyt1Aa has

high affinity for mosquito microvilli lipids. The chimeric protein forms a stable parasporal inclusion when expressed in an acrystalliferous strain of Bti and is highly toxic to *Aedes aegypti* apart from *Anopheles gambiae, An. stephensi*, and both BinAB-sensitive and BinAB-resistant strains of *Culex quinquefasciatus*. It has proven as a successful attempt in expanding the target spectrum of BinA [34].

- Bioconjugation strategies

Bioconjugation is the chemical crosslinking of the biomolecule with a synthetic or semi-synthetic moiety. Bioconjugation offers therapeutic opportunities and is gaining attention in the field of biopharmaceuticals. One such example is PEGylation. PEGylation is the covalent attachment of poly (ethylene glycol) to proteins of diagnostic and therapeutic importance to improve their pharmacokinetic properties. It was introduced in 1977 by Abuchowski and co-workers when they observed superior immunogenic properties of bovine serum albumin (BSA)-PEG conjugates compared to the unmodified protein [67]. Later it was demonstrated that PEGylation results in enhanced circulation lifetimes and reduces immunogenicity [68]. Since then PEGylation has been demonstrated to improve pharmacological properties of proteins. PEG has been classified as “Generally Regarded As Safe” (GRAS) by FDA. PEGylation stabilizes the protein by enhancing its thermal, chemical and enzymatic stability [69] and increases the size [70], thereby reducing clearance from the bloodstream [71]. Increased half-life of the protein drug reduces its dosage frequency [72]. Several FDA-approved PEG conjugated proteins have been released as drugs [73]. Notably, the number of PEG moieties and their site of conjugation are critical to the efficacy and activity of the protein. Considerable efforts are made to achieve site specific conjugation.

Thesis objectives

The primary objectives of the present thesis are structural and biochemical analyses of the functional domains of binary toxin component proteins, understanding the molecular basis of toxicity, understanding the interaction of BinAB proteins with the receptor protein, and to "rationally" improve the efficacy of BinAB toxin. A highly toxic local isolate of *L. sphaericus* (ISPC-8) was used in the present studies.

Chapter 2

Materials & Methods

The present chapter summarizes the general protocols used during experimentation.

2.1. MATERIALS

All the chemicals used were of analytical grade or molecular biology grade. Molecular biology grade chemicals, inorganic salts, organic chemicals and solvents of highest purity grade were from Sisco Research Laboratories (SRL), India or Sigma, USA or Roche, Switzerland. The plastic-wares were from Tarson. All the glass-wares were from Borosil. Luria-Bertani broth and Luria-Bertani agar were procured from Himedia, India. The enzymes, DNA molecular weight standards and Protein molecular weight standards were from New England Biolabs, USA. Plastic disposable filter assemblies (0.22 μm) were from Sartorius, Germany. Dialysis tubing and tubing clamps were from Spectrum Scientific, USA. Immobilon^{TM}-P Polyvinylidene Difluoride transfer membranes (for immunoblotting of proteins) and ultrafiltration units (Centricon and Centriplus) were obtained from Millipore Corporation, USA. Other important chemicals and materials have been given in the individual chapters.

2.2. REAGENT AND BUFFERS

All the reagents and buffers were prepared in double distilled water. The buffers and reagents were prepared as described in the Laboratory manual- Molecular biology cloning [74]. The reagents were autoclaved or prepared in autoclaved water and the buffers were filter sterilized using 0.22 μm syringe filters and stored after autoclaving.

2.3. BACTERIAL STARINS, PLASMIDS & DNA PRIMERS

The list of the standard bacterial strains, plasmids and DNA primers used in the present study are given in **Table 2.1, 2.2** and **2.3** respectively.

Table 2.1: List of bacterial strains used during the present study

Strain name	Genotype	Remarks
Lysinibacillus sphaericus ISPC 8	serotype 5a5b and phage type 3	Isolated from dead larvae of *Culex* species collected from field [75]
DH5α	F^- *rec*A1 *end*A1 *gyr*A96 *rel*A1 *thi*-1 *sup*E44 *hsd*R17 *Δlac*U169(Φ80 *lac*ZΔM15)	Lab collection; Cloning host
XL10-Gold	Tet^rΔ*(mcrA)183* Δ*(mcrCB-hsdSMR-mrr)173* *end*A1 *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac* Hte [F′ *proAB* $lacI^q$ZΔM15Tn*10* (Tet^r) Amy Cam^r]	Stratagene; Cloning host
BL21(DE3)	F^- *dcm ompT hsdS* (r_B-m_B-) *gal* λ (DE3)	Novagen; expression host
BL21Star (DE3)	F^-*ompT hsdS*B ($r_B^-m_B^-$) *gal dcm rne131*(DE3)	Novagen; expression host
BL21(DE3) pLysS	F^-*ompT hsdS* (r_B^- m_B^-) *gal dcm* λ(DE3) pLysS (Cam^r) (λ(DE3): *lac*I, *lac*UV5-T7 gene 1, *ind*1, *sam*7, *nin*5)	Novagen; expression host

Table 2.2: List of plasmids used/derived in the present study

Name of plasmids	Size (bp)	marker	Remarks
pET-28a(+)	5369	Kan	Novagen
pCold-TF	5769	Amp	TaKaRa
pGEX-4T-3	4968	Amp	GE Healthcare
pNIC28-Bsa4	7284	Kan	Addgene plasmid # 26103
pET28a-*bin*A	6425	Kan	present study, truncated active *bin*A gene cloned into *Nde*I and *Bam*HI site of pET28a(+)
pET28a-*bin*B	6716	Kan	present study, full length *bin*B *gene* cloned into *Nde*I and *Bam*HI site of pET28a(+)
pNIC28-Bsa4-*cqm*1	6967	Kan	present study, truncated *cqm*1 CDS cloned into pNIC28-Bsa4 vector using ligation independent cloning method.

Table 2.3: List of the DNA primers* used for the present study.
(*Restriction enzyme site engineered are underlined).

Primer	Sequence (5' to 3')	Remarks
tBinA F	ATATACATATGAAGTACATTCGC GTTATGG	Forward and reverse primers for cloning truncated *bin*A gene into *Nde*I and *Bam*H1 sites of pET28a(+)
tBinA R	TATAGGATCCTTATGTAATAATCT TTG	
BinB F	GATGAAGAACATATGTGCGATTC AAAAGAC	Forward and reverse primers for cloning *bin*B gene into *Nde*I and *Bam*H1 sites of pET28a(+)
BinB R	AGTTGGATCCTTACTGGTTAATTT TAGGTATTAA	
Cqm F	TATTACATATGCGACCGCTGGGA GCTTTG	Forward and reverse primers for amplifying full length *cqm*1 gene from *Culex* larvae
Cqm R	ATCGCAAGCTTACACGAAAATAT ACCTGGCC	
Cqm1 F	ATTATACATATGCGCCCGCTGGGT GCTC	Forward and reverse primers for cloning chemically synthesized full length *cqm*1 CDS into *Nde*I and *Hind*III sites of pET28a(+) and pCold-TF vectors
Cqm1 R	AATTAAGCTTTTAGACGAAGATG TAGCGTGCC	
CQM-d22-F	TAC TTC CAA TCC ATG CGT GAA CCG GAC TCA AAA G	Forward and reverse primers for cloning truncated-*cqm*1 CDS into pNIC28-Bsa4 vector using LIC.
CQM-560-R	TAT CCA CCT TTA CTG CGC GGA ACT CAC GGT CGC	
rdm3	CCGGCGTAGAGGATCGAGAT	pET plasmid primer binds upstream of RBS
pcg1	TAGTTATTGCTCAGCGGTGG	pET plasmid primer binds further downstream of transcription terminator

2.4. METHODS

2.4.1. Construction of recombinant plasmids

Molecular cloning involves a set of experimental methods to assemble recombinant DNA molecules and to direct their replication within host organisms. Construction of recombinant plasmid through molecular cloning involves following steps-

Choice of host organism and cloning vector (plasmid)

⇩

Preparation of vector DNA (gene fragment)

⇩

Preparation of DNA to be cloned (PCR, restriction-digestion)

⇩

Creation of recombinant DNA (Ligation)

⇩

Introduction of recombinant DNA into host organism (Transformation)

⇩

Selection of organisms containing recombinant DNA

⇩

Screening for clones with desired DNA inserts and biological properties (Colony PCR).

2.4.1.1. Plasmid purification

In the present study, plasmid purification, from the bacterial host cells (usually DH5α and XL10-gold strains), was done using the plasmid purification kit from Qiagen which is based on the modified alkaline lysis method.

2.4.1.2. Purification and concentration of DNA using phenol:chloroform method

PCR amplified DNA was purified and concentrated using the phenol:chloroform method. Tris-EDTA (TE) buffer was added to DNA solution to be purified to a final volume of 400 μL. Sodium acetate buffer (pH 5.2) was added to a final concentration of 0.3 M and mixed properly. Equal volume of phenol:chloroform:isoamylalcohol mixture (25:24:1) was added and mixed thoroughly by pipetting for 1 min before centrifuging at 12,000 rpm for 5

min. The upper aqueous layer was transferred to a fresh tube and the DNA was precipitated with 2 volumes of ethanol followed by incubation for 1 h at -20 ºC and centrifugation at 14000 rpm, for 15 min. The pellet was washed with 70 % ethanol, dried and dissolved in minimum volume of TE buffer. For purification of particular size DNA fragment, the DNA was extracted from gel-excised band and extracted as per the protocol of the QIAEX II kit from Qiagen. The purified DNA was finally eluted in a low salt buffer.

2.4.1.3. Qualitative and quantitative detection of DNA through Agarose gel electrophoresis

Purification, amplification and restriction-digestion of plasmid DNA and genes fragments, during cloning experiments, was monitored by agarose gel electrophoresis (AGE). Typically, 0.8 – 1% agarose gels were prepared in 1X TAE buffer with 0.5 µg/mL concentration of ethidium bromide.

2.4.1.4. Primer designing

The primer sequences were designed from the target sequences and were validated for primer dimer, % GC and melting temperature (Tm) using the Multiple Primer Analyzer program from ThermoFisher Scientific (www.thermofisher.com/in/thermo-scientific/thermo-scientific-web-tools/multiple-primer-analyzer.htmL). Each primer was typically between 18-33 nucleotides in length, having 35 – 60% GC content (Tm between 45 – 70 ºC) and possessed no complimentary sequence/runs of bases towards 3' end. Desired restriction enzyme sites were incorporated towards 5' end of the primers with the sufficient flanking sequence that is required for the efficient cutting by the restriction enzyme. In all cases, care was taken to ensure that the last 4 to 5 bases of the primer, at the 3' end, were highly specific to the fragment to be amplified.

2.4.1.5. Polymerase chain reaction (PCR)

PCR reactions for gene amplification were carried out using high fidelity *Phusion* DNA polymerase enzyme. For selection of positive recombinant through colony PCR, *Taq* DNA polymerase was used instead of *Phusion* polymerase. The PCR reaction mix was prepared keeping the standard concentrations of the components. DNA polymerase enzyme was added at the end. Annealing temperature varied from 54—60 ^{o}C depending upon the primer-template complementarity and GC content. Extension time varied based on the length of amplified product, typically, 30 to 60 sec/kb. For *Taq* polymerase, denaturation and extension temperatures were 94 ^{o}C and 68 ^{o}C, respectively. PCR was carried out for 35 cycles.

2.4.1.6. Restriction digestion

The conditions employed for restriction digestion described in this thesis are essentially those recommended by the manufacturer. Typically, 20—30 units of restriction enzyme were added to the 50 μL reaction mixture (0.5—1 μg DNA) followed by incubation for 4 to 6 hrs. at 37^{o}C. The digested DNA was resolved on 0.8% agarose gel. The DNA marker (100 bp and/or 1kb DNA ladder, NEB) were used to estimate sizes of the products. The DNA was further purified by either gel extraction or phenol:chloroform methods followed by alcohol precipitation.

2.4.1.7. DNA Ligation

In a typical ligation reaction, the restriction digested vector and insert were mixed together in the ratio of 1:3 (molar ratio). The amount of DNA used for ligation was varied between 100 to 500 ng. After addition of T4 DNA ligase the reaction mixture was incubated at 16 ^{o}C overnight, and was used directly for transformation.

2.4.1.8. Ligation independent cloning

Ligation independent cloning (LIC) is an alternative method to the traditional restriction digestion/ ligase based method of cloning. It makes use of the 3'—> 5' activity of T4 DNA polymerase to create very specific 10—15 base single overhangs in the expression vector as well as in the insert DNA. The annealing of the insert and the vector is performed in the absence of ligase by simple mixing of the DNA fragments. The protocol for LIC, followed in the present thesis, was as described by Savitsky et al [76]. The LIC vector used in the present studies is pNIC-Bsa4 (procured from Addgene, plasmid no. 26103; gift from Opher Gileadi). To create an insert with complementary overhangs to the LIC vector, the following primers were used:

Forward primer: TACTTCCAATCCATG-gene of interest

Reverse primer: TATCCACCTTTACTG-gene of interest (reverse complement)

The annealed product was transformed into XL-10 gold competent cells using the protocol discussed in section 2.4.1.9. Positive constructs were identified through colony PCR performed as per the protocol discussed in section 2.4.1.10.

2.4.1.9. Transformation

Prior to transformation, *E. coli* cells were made competent using the calcium chloride method. Transformation was done by heat shock method at 42 ºC with 50–100 ng of DNA and 100 µL of competent cells. 50–100 µL or concentrated mixtures were plated on LB agar plates with appropriate antibiotics and incubated O/N at 37 ºC. The transformed bacterial cell colonies were selected and streaked onto fresh LB agar plates.

2.4.1.10. Selection of positive recombinants by colony PCR

Colony PCR was performed to screen the positively transformed recombinant bacterial colonies. The O/N incubated transformants were transferred to fresh LB agar plates

with appropriate antibiotic. The preparation of PCR reaction mixture (25 μL) and thermal cycling conditions were as described previously except for the addition of template. Using sterile 0.5–10 μL pipette tips, a single colony was touched (~1/10th of colony) and suspended into the reaction mixture. PCR reaction was performed with the plasmid/gene specific primer sets for the amplification of the target ORFs. Colony PCR that yielded a product of appropriate size indicated successful cloning of the respective gene. Plasmid was isolated from positive clones, and the release of appropriately-sized DNA insert was verified by restriction digestion and sequencing.

2.4.1.11. DNA sequencing

Positively recombinant clones were sequenced using an automated DNA sequencer based on Sanger's dideoxy method of DNA sequencing available at a commercial facility. The electrophoretogram was compared with the DNA sequence to confirm the results using BioEdit software.

2.4.1.12. Overlap extension polymerase chain reaction

In the present thesis, overlap extension polymerase chain reaction (OE-PCR), also referred to as *Splicing by overlap extension / Splicing by overhang extension (SOE) PCR,* was performed to splice off the introns from the target gene sequence. The target exons were amplified with special primers having a 5' overhang complementary to the end of the other molecule to be annealed. The PCR was performed as described earlier in section 2.2.1.6. The amplified exons were annealed together through their overlapping ends. A final PCR was carried out with the primers for the 5' and 3' ends of the target gene. The success of the method was determined after analysis of the amplified products on 0.8–1% agarose gel.

2.4.2. Protein purification from high expression systems

2.4.2.1. Overexpression of target gene in pET expression system

The target genes were cloned into pET based vectors under control of strong bacteriophage T7 transcription and (optionally) translation signals. Recombinant constructs were transformed into *E. coli* based expression host cells, like BL21 (DE3), BL21 Star(DE3), pLysS etc., containing a chromosomal copy of the T7 RNA polymerase gene under *lacUV5* control (under IPTG control). Induction of protein expression was achieved by the addition of the IPTG (0.2–1 mM) to the growing culture in the late log phase (OD_{600} ~ 0.7−0.8) under temperature control, which induces T7 RNA polymerase, which in turn transcribes the target DNA in the plasmid. These parameters were optimized for expression of different recombinant proteins. The cells were harvested after 4 h to O/N induction by centrifugation at 6000 rpm for 5 min. The cell pellet was stored at -20 ºC till further use. The expression of the proteins was checked using SDS-PAGE by loading and comparing the uninduced and induced samples.

2.4.2.2. Protein purification

The harvested cells were lysed in lysis buffer using a combination of enzymatic lysis (using lysozyme) and mechanical methods (pulsed sonication). The desired proteins were enriched to apparent homogeneity using various chromatography techniques. Different kinds of column chromatography methods, like immobilized metal affinity column chromatography (IMAC), ion exchange chromatography (IEC) and size exclusion chromatography (SEC), were employed to purify the expressed proteins. Proteins in the present study were mostly purified using a low-pressure liquid chromatography system (Bio-Rad Biologic LP) with manually packed columns at 25 ºC. The elution from the column matrix was monitored online for absorbance at 280 nm wavelength. Eluted protein fractions were adjudged from SDS-PAGE. In the final purification stage using SEC method, Superdex 200 and Superdex

75 columns were used on FPLC (Bio-Rad). The detailed protocols have been mentioned in the relevant chapters.

Immobilized metal ion affinity chromatography (IMAC) method using the Ni-IDA matrix was employed to purify recombinant proteins containing a short affinity tag consisting of poly-histidine residues. The bound protein was eluted by increasing imidazole concentration gradient.

Ion exchange chromatography (AEC) was primarily used to purify and enrich the protein, based on their isoelectric point (pI). The purification was achieved through increasing salt (NaCl) concentration gradient. Q-sepharose (quaternary ammonium group attached to crosslinked agarose beads, a strong anion exchanger was primarily used in the present work.

Size exclusion chromatography (SEC) was employed, to achieve homogeneous purification of the protein, for estimating the molecular weight of the protein in solution and to adjudge the native folding state of the protein in solution. Two different SEC columns were used, namely, Superdex-75^{TM} (GE Healthcare) and Superdex-200^{TM} (GE Healthcare), with the optimum separation ranges from 3 to 70 kDa and 10 to 600 kDa, respectively. The columns were calibrated with gel filtration molecular weight markers (Sigma-Aldrich; Cytochrome C, 12.4 kDa; Carbonic anhydrase, 29 kDa; Ovalbumin, 44.3 kDa; Albumin, 67kDa; Apoferritin, 443 kDa; Thyroglobulin, 669 kDa). The buffers used for SEC were inclusive of certain amount of salt (~100 mM NaCl) to prevent the non-specific ionic interactions with the column matrix or other contaminating proteins.

2.4.2.3. Estimation of protein concentration

UV absorbance was measured at 280 nm using spectrophotometer (JASCO) throughout the work. Protein concentration was roughly calculated using Beer-Lamberts law using molar extinction coefficient calculated from the Swiss-Prot (PROTPARAM) server by submitting the primary protein sequence (https://web.expasy.org/protparam) [77]. The molar extinction

coefficient for the proteins studied in the present thesis, as calculated by PROTPARAM, are 1.24 M^{-1} cm^{-1} for BinA, 1.1 M^{-1} cm^{-1} for BinB and 1.83 M^{-1} cm^{-1} for Cqm1 proteins.

Bradford method [78] was performed as per the instruction of the dye manufacturer. The working range of this method extends from 0.1 –1.4 mg/mL of the protein sample used. A standard curve with known concentrations of Bovine Serum Albumin (BSA) protein was created for the assay. BSA protein standards of appropriate concentrations were prepared, with serial dilution, in the same buffer as that of the test protein. The test protein samples with unknown concentrations were also prepared by dilutions to ensure that the concentrations fall in the range desirable for the assay. Three mL of Bradford reagent (Sigma) were added to 100 µL of protein samples (BSA and test samples) and incubated at RT for 30 min. The absorbance of was measured at 595 nm. The concentrations of unknown samples were calculated from the standard curve prepared using BSA (Fig. 2.1).

Modified Folin-Lowry method [79] works in the linear concentration range of 5 µg/mL to 25 µg/mL. Typically, 200 µL freshly prepared alkaline copper reagent [1% Copper sulfate, 50 µL; sodium-potassium tartrate, 50 µL; and 900 µL D/W] was added to 200 µL of dilute protein samples (5–25 µg protein) and incubated at RT for 10 min. Thereafter, 600 µL of freshly diluted Folins reagent (1:10) was added and the reaction mix was transferred immediately to water bath kept at 50 ºC and incubated for 10 min. The absorbance of the standards (BSA, 5 µg/mL to 25 µg/mL) and unknown samples were measured at 660 nm after the reaction tubes cooled to RT. The concentrations of unknown samples were calculated from the standard curve prepared using BSA (Fig. 2.2).

2.4.2.4. Detection of proteins through Immunoblotting

Immunoblotting was used for detecting the poly-His tag in the purified proteins and for detecting the PEG moiety used for chemical modification of the protein as mentioned in chapter 6. PVDF membrane (Hybond-P, GE Healthcare) used for blotting was pre-

equilibrated as per the manufacturer's instructions. Protein transfer was conducted at constant current (300 mA) for 1-2 hr at 4 ºC in the blotting apparatus (BioRad) containing transfer buffer [25 mM Tri-HCl (pH 8.3), 20 % methanol and 192 mM glycine]. 1% BSA was used as the blocking reagent. The primary antibodies, anti-His antibodies, anti-polyethylene glycol and anti-PEG (methoxy group) monoclonal antibodies produced in Rabbit, were used at 1: 10,000 dilution in blocking reagent. Alkaline phosphatase conjugated anti-rabbit IgG was used as the secondary antibody at 1: 10,000 dilution in 1X Ma/Na buffer (10X stock: 1 M Malic acid, 1.5 M NaCl). NBT/BCIP was used as the chromogenic substrate.

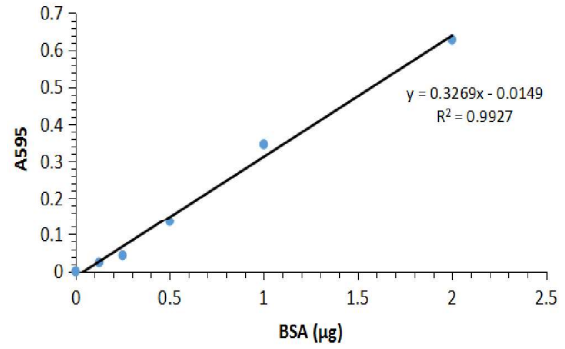


Figure 2.1. Standard curve for protein estimation using Bradford method.

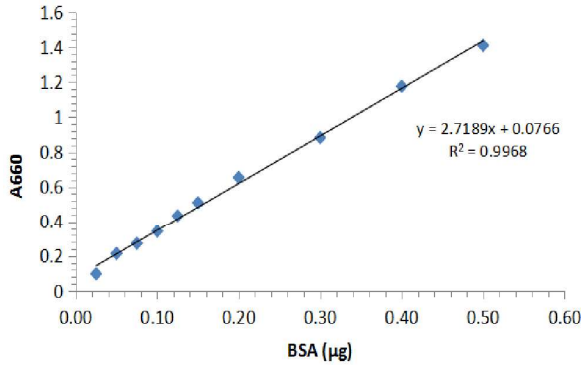


Figure 2.2. Standard curve for protein estimation using modified Folin-Lowry method.

2.4.2.5. Concentrating dilute protein

Ultrafiltration technique was used to concentrate the dilute protein sample using Amicon Ultra (Millipore, USA) centrifugal filter devices of molecular weight cut-off limit of 10 and 30 kDa. The protein was concentrated by centrifugation at 5000 x*g* at 4 ºC. Concentrated proteins were stored at 4 ºC.

2.4.3. Biophysical techniques used for protein characterization

2.4.3.1. Matrix-Assisted Laser Desorption Ionization Time-of-flight (MALDI-TOF) Mass Spectrometry analysis

MALDI-TOF analyses of protein samples for estimation of molecular weights and for protein identification through peptide mass fingerprinting, were mostly carried out at ACTREC, Mumbai and IIT Powai, Mumbai. The protein samples were diluted to 50–100 pmol/µL concentration in buffer [20 mM Tris-HCl pH 8, 25 mM NaCl] and mixed with 2 µL of sinapinic acid matrix. This mixture was spotted on a copper grid and allowed to dry at RT, and subsequently loaded on the spectrometer. External calibration was performed using the Peptide Standard Calibration II. Spectra were processed using the FlexAnalysis software. The peptide mass fingerprinting search was carried using mass peaks observed for Trypsin digested proteins and MASCOT search engine (http://www.matrixscience.com/).

2.4.3.2. Dynamic Light Scattering

In the present study, the dynamic light scattering (DLS) analysis was employed to determine the hydrodynamic radii and % polydispersity of proteins, using the Malvern Zetasizer nano ZS instrument. Typically, 0.5 mL of 0.5−1 mg/mL concentration of the protein was used for the analysis. The buffer and pH conditions are mentioned in the relevant section within the chapters. The sample was spun at a very high speed of 14000 rpm for 15 minutes, before the analysis, to settle down any particulate matter. The analysis was done at

25 ºC. The data analysis was done with the help Malvern Zetasizer software suite (http://www.malvern.com).

2.4.3.3. Circular Dichroism

In the present thesis, all the CD measurements were done on JASCO spectrometer (J-815). Typically, 0.3–0.5 mg/mL concentration of the protein sample in phosphate buffer (pH 7.2) were used for measurements. The variations in the buffer and pH conditions, as per the experiment, are given in the relevant chapters. The CD spectra measurements were done in the far UV-region from 200−260 nm at 20 ºC in 1 mm path length quartz cuvette. The observed ellipticity (in millidegrees) was converted into molar ellipticity ($\Delta\varepsilon$, deg·cm^2/dmol). The secondary structure content (α-helix and β-strand) was estimated using the K2D2 web server (http://www.ogic.ca/projects/k2d2, [80]).

2.4.3.4. Thermofluor shift assay

Differential Scanning Fluorimetry (DSF), often called the thermofluor shift assay, probes the thermal stability of the tertiary structure of a protein as it experiences progressive denaturation. It is used as a high throughput screening method to identify low molecular-weight ligands that bind and stabilize the protein. The temperature at which the protein unfolds is measured by an increase in the fluorescence of a dye like SYPRO orange with affinity for hydrophobic parts of the protein, which are exposed as the protein unfolds. In the present thesis, thermofluor-shift assay was performed as described by Niesen et al [81]. The measurements were performed in a 96 well plate in 25 µL reaction, in triplicates, containing purified protein (2 µM) mixed with freshly diluted SYPRO orange dye (1:1000, v/v). The data was recorded using a real-time PCR (CFX-96, BioRad). The protein samples were heated from 20 to 95 ºC at ramp rate of 1.0 ºC/ min. Thermal unfolding of Hen Egg White (HEW) lysozyme was also probed in the same 96 well plate as an internal control. The

increase in the fluorescence intensity against temperature generated a sigmoidal curve, characteristic for two state transition. The transition temperature or the melting point (T_m) for the protein was determined from the maximum of the first derivative of this sigmoidal curve [81].

2.4.3.5. Surface Plasmon Resonance

Surface plasmon resonance (SPR) has emerged as a potential technique for label free detection of biomolecular interactions in real time. It is being applied for detection of carbohydrate specific interactions of bacterial polysaccharide antigens providing information vital for vaccine development [82]. In a typical SPR experiment, the ligand is immobilized to the sensor surface, the analyte is free in solution and passed over the surface. Changes in the refractive index of the sensor surface owing to binding and dissociation of the analyte is measured as a change in resonance angle and the interaction profile is recorded as a sensogram, response units (RU) plotted against time, over the course of analysis.

In the present thesis, protein-protein and protein-small molecular weight molecules interactions were studied was using SPR. The SPR experiments and analysis were performed using the Biacore T200 SPR (GE Healthcare) and Autolab Esprit SPR systems. Increasing concentration of the analytes were flown over the immobilized ligand. The resulting sensogram was subjected to affinity analysis and kinetic analysis. The equilibrium dissociation constant (*K*D) measured from the affinity analysis indicated the strength of binding at equilibrium. The kinetic analysis estimated the association ('on rate', *k*a) and dissociation rates ('off rate', *k*d), for the binding.

2.4.3.6. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) was carried out to thermodynamically characterize the protein-low molecular weight ligand interactions. A fixed concentration of

the protein was titrated with incremental additions of ligand over time, at constant temperature. The formation of complex, in the form of heat, released or absorbed, needed to maintain isothermal conditions, was measured. The ITC experiments were performed using the MicroCal iTC200 instrument (Malvern, UK). The detailed procedure, with proper experimental conditions, has been described in the relevant chapters.

2.4.3.7. Fluorescence spectroscopy to investigate protein ligand interactions

Fluorescence spectroscopy has found widespread use as a tool to study protein-ligand interactions. It offers the advantage of monitoring the changes in the fluorophore microenvironment as a measure of ligand binding. Tryptophan residues in the proteins are used as intrinsic fluorophores to monitor protein ligand interactions. When the intrinsic tryptophan fluorescence of a protein is not sufficiently sensitive to ligand binding, the protein may be labelled with an extrinsic fluorescence probe. The present study involves the use of intrinsic and extrinsic fluorescence (labelled with an extrinsic fluorophore) of the protein to study the changes in the fluorophore microenvironment in terms of emission intensity or emission wavelength to examine its interaction with different ligands. The fluorescence experiments were performed on a JASCO spectrofluorometer (FP-8500) with 1 cm path-length. INFINITE M200 (Teccan) multimode plate reader was employed for a high throughput screening of ligands for their interaction with protein labelled with an extrinsic fluorophore.

2.4.3.8. Fourier transform infrared (FTIR) spectroscopy

FTIR spectroscopy is a very sensitive and useful tool to monitor the dynamics of protein–ligand interactions. Hydration of proteins is important for their three-dimensional structure and their activity. Ligand binding to biomolecules induces changes in their structure, conformation, and hydration [83]. The present work makes use of the high sensitivity of IR

for water and changes in hydrogen boding owing to ligand binding. In the present study FTIR spectroscopy was employed to validate the interaction of small ligands with the target protein by examining the absorption of water in the 3000−3100 cm^{-1} region. The experiments in the present study were performed on a Vertex 70 FTIR spectrometer (Bruker, Germany) equipped with BioATR system and a HgCdTe (MCT) detector.

2.4.4. Bioassay for protein larvicidal activity analysis

The *Culex* larvae for bioassay performed in chapters 3 and 5 and for amplification of receptor protein Cqm1 gene sequence as discussed in chapter 4 were reared from BinAB susceptible *C. quinquefasciatus* colony, established from local strain found in Trombay and Mankhurd regions, Mumbai, India, without exposure to *L. sphaericus*. The *Culex* nucleus culture for bioassay performed in chapter 6 was obtained from National Institute of Virology, Pune. *Culex* nucleus cultures were maintained at 28 ± 2 °C and 85% relative humidity in a BARC laboratory. Eggs were allowed to hatch in plastic bowls containing 1 L tap water supplemented with 0.13 g sterilized larval food (13:6:1, wheat flour: chickpea flour: yeast extract) [84]. The larvicidal activity of the toxin proteins was assayed against third instar *C. quinquefasciatus* larvae. Ten larvae in 10 mL of autoclaved distilled water were taken in a glass beaker (10 mL capacity). Varying concentrations of the test protein were added to each larval set. Each concentration was tested in duplicates or triplicates. Water and buffer controls were set to account for the effects of environmental conditions and the buffer of the protein. The final mortality observations were taken after 48 hr. Control mortality was corrected. The % mortality was calculated using Abbott formula [85]. The median lethal concentrations (LC_{50}) were calculated at 95% confidence limit by probit analysis [86] using SPSS 12.0 and R program for statistical computing (https://www.r-project.org/; [87]). Further, the analysis of LC_{50} values and confidence limits for different protein samples was performed with Students t-test ($p < 0.05$).

Chapter 3

An oligomeric complex of BinA/BinB is not formed *in-situ* in mosquito-larvicidal *Lysinibacillus sphaericus* ISPC-8

3.1. ABSTRACT

Binary toxin of *L. sphaericus* is composed of two polypeptides; receptor binding BinB and toxic BinA. Both the polypeptides are required for maximal toxicity. It has been suggested that binary toxin exerts toxicity as a heterotetramer constituted by two copies of each of the component polypeptides. It has also been argued that oligomers consisting of two copies of BinA and BinB are pre-formed in *L. sphaericus* spore-crystals. However, recombinant proteins from *E. coli* expression system elute individually as monomers. We purified the likely oligomeric complex from the spore-crystals of highly toxic *L. sphaericus* ISPC-8 strain and probed it with proteomic tools. The analysis showed that the high molecular mass complex in the toxic spore-crystals is composed of only surface layer protein (SlpC). The purified SlpC from the local isolate exists as a dimer and showed poor mosquito-larvicidal activity.

3.2. INTRODUCTION

The Binary toxin of *L. sphaericus* is composed of two polypeptides; receptor binding BinB and toxic BinA. The heterotetramer (BinA2.BinB2) of binary (BinAB) toxin formed on the cell receptor, a GPI-anchored α-glucosidase, inserts into the cell membrane forming membrane pores [35−37, 51] and subsequently causes cell death.

It has been suggested that a stable oligomeric form of BinAB toxin may exist in spore-crystals of *L. sphaericus* 2362 and this oligomer may be playing role in toxicity towards *Culex* larvae [37]. Such a pre-formed oligomer was not reported for the recombinant BinA/BinB proteins co-expressed in *E. coli*, nor was this observed in a stoichiometric mixture of recombinant BinA/BinB proteins *in vitro* [32]. It is thus important that *L. sphaericus* components expressing toxic activity be characterized properly.

The present chapter discusses the characterization of such an oligomer in locally isolated and highly toxic strain of *L. sphaericus* ISPC-8. We did not find any evidence of a pre-formed oligomer of BinA and BinB proteins. Rather, the expected protein complex, purified to homogeneity, was confirmed to be a surface layer protein that displayed poor mosquito-larvicidal activity.

3.3. MATERIALS AND METHODS

3.3.1. Bacterial stain and growth conditions

L. sphaericus ISPC-8 having serotype 5a5b and phage type 3 was isolated from dead larvae of *Culex fatigans*. The bacterial culture was grown till sporulation as described earlier [88]. The pellet containing spore-crystals was washed twice with 1 M NaCl and 10 mM EDTA (pH 7.0) and twice with double-distilled water. The pellet was suspended in sterile distilled water and sonicated for 5 min with 3 sec pulses at 30% amplitude in a Branson digital sonicator (Model 250, Branson Ultrasonics Corporation, CT). The sonicated suspension was centrifuged at 15,000 x*g* for 30 min. The supernatant was discarded and the spore-crystal pellet was used for purification of the target proteins.

3.3.2. Purification of proteins

The purification of proteins was carried out with slight modification of the protocol described by Smith et al [37]. The spore-crystal pellet was suspended in 50 mM NaOH and the suspension was kept shaking on ice for 3 h. The suspension was centrifuged at 15,000 x*g* for 20 min at 4 ºC. The supernatant containing the solubilized binary toxin was dialyzed overnight against buffer A (25 mM Tris-HCl, 10 mM NaCl, 2 mM DTT, pH 9.0). The dialyzed suspension was centrifuged at 15,000 x*g* for 20 min at 4 ºC and the supernatant was loaded onto Q-sepharose column (Bio-Rad laboratories, Hercules, CA). The bound proteins were eluted with a linear gradient of 10–1000 mM NaCl in buffer A over six-column volume. The fractions of high molecular mass proteins were pooled and dialyzed extensively against buffer A. After dialysis, the pooled fractions were concentrated to about 2 mg/mL.

3.3.3. Estimation of molecular size

The purified proteins were loaded on the Superdex^{TM} 200 10/300 GL column (Superdex-200; GE Healthcare) for further purification as well as for molecular weight

determination. Each eluted peak from Superdex-200 molecular sieve column was further loaded onto Superdex-75 column. The Superdex-200 and Superdex-75 columns were calibrated with gel filtration molecular weight markers (Sigma-Aldrich; Cytochrome C, 12.4 kDa; Carbonic anhydrase, 29 kDa; Ovalbumin, 44.3 kDa; Albumin, 67 kDa; Apoferritin, 443 kDa; Thyroglobulin, 669 kDa). The eluted peak of each independent gel filtration experiment was resolved on SDS-PAGE. The purified proteins were concentrated to 5–7 mg/mL and were subjected to dynamic light scattering experiments at 25 ^{o}C with Malvern Zetasizer NanoZS instrument. Three sets of measurements with 18 acquisitions in each set were collected. Data were analyzed with Malvern Zetasizer software suite (http://www.malvern.com).

3.3.4. MALDI-TOF mass spectrometry

The mass spectrometry analysis was performed at the Adelaide Proteomics Centre, University of Adelaide, Australia. Lyophilized protein sample was sonicated in 100 µL of 10% acetonitrile (ACN) / 0.1% trifluoroacetic acid (TFA) to dissolve the protein. About 1 µL of concentrated sample was electrophoresed in 5% acrylamide gel (Bis-Tris buffer). The gel was cut in two halves, one half was stained with silver nitrate and the other stained with Coomassie G-250. The closely spaced high molecular weight (approximately 125 kDa) bands seen in the G-250 stained gel were excised, destained and digested with 100 ng of trypsin per sample according to the "low salt" protocol (Bruker Guide to MALDI Sample Preparation). One microlitre of each sample was applied to a 600 µm Anchor Chip (Bruker Daltonik GmBH, Bremen, Germany) to which 400 ng of α-cyano-4-hydroxycinnamic acid had been applied. MALDI-TOF mass spectra were acquired using a Bruker ultraflex III MALDI-TOF/TOF mass spectrometer (Bruker Daltonik GmBH) operating in reflection mode under the control of the flexControl software (version 3.0, Bruker Daltonik GmBH). Between three and six of the most highly abundant sample ions (i.e. non-trypsin and non-keratin) were

selected as precursors for MS/MS analysis. MALDI-TOF/TOF was performed in the LIFT mode using the same spot on the target.

The peptide mass fingerprinting search was carried using 21 mass peaks observed for trypsin digested proteins and MASCOT search engine (http://www.matrixscience.com/). The specifications were: Taxonomy: eubacteria, databases: NCBInr 20140323 & SwissProt 2014_04, Enzyme: Trypsin, Fixed modifications: Carbamidomethyl (C), Variable modifications: Oxidation (M), peptide mass tolerance and charge state of 1.2 Da and 1+, and missed cleavages: 1. The MOWSE and probability scores calculated by the software were used as the criteria for protein identification.

3.3.5. Toxicity analysis of the high molecular mass protein

The susceptibility of 3rd instar *Culex quinquefasciatus* larvae was checked against *L. sphaericus* spore-crystal mixture (1 x 10^5 spores/mL) in 10 mL of distilled water in triplicate each containing ten larvae. Total mortality was scored after 48 hr. The toxicity of the high molecular mass protein (purified by molecular sieve column chromatography) was tested against the same colony of *C. quinquefasciatus* larvae. Different concentrations of the purified protein were tested in 10 mL of distilled water containing ten third-instar *C. quinquefasciatus* larvae. Two replicates for each concentration (10, 50, 100, 150 μg/mL) were tested. The larval mortality was recorded after 48 hr.

3.4. RESULTS

3.4.1. Purification of proteins

The expected high molecular-weight precursor proteins and BinA/B proteins from spore-crystals of highly toxic local strain of *L. sphaericus* ISPC-8 were purified to homogeneity using ion-exchange and molecular-sieve column chromatography. High molecular weight components (Mr, ~125 kDa), BinB (51.4 kDa) and BinA (41.9 kDa) were found to co-elute on ion-exchange column (Fig. 3.1A, Lane 2). The high molecular weight precursors and BinA/B proteins were clearly resolved by molecular sieve column chromatography using Superdex-200 (Fig. 3.1A & 3.2). The peaks were individually loaded onto Superdex-75 column. Expectedly, the high molecular weight precursor protein eluted in the void volume of Superdex-75 column. Also, BinA and BinB proteins were further resolved on this molecular sieve column.

3.4.2. Molecular size and protein identification

The molecular masses of the proteins were determined based on their elution profile on the molecular sieve Superdex-200 and Superdex-75 columns. The molecular mass of the high molecular weight protein was estimated to be 282 kDa. In contrast, the BinA and BinB proteins elute as monomers from Superdex-200 and Superdex-75 columns with estimated molecular mass of 56 and 40 kDa, respectively (Fig. 3.2). Further, two bands of approximately 125 and 110 kDa were observed for the high molecular-weight proteins when resolved on 5% SDS-PAGE (Fig. 3.1B).

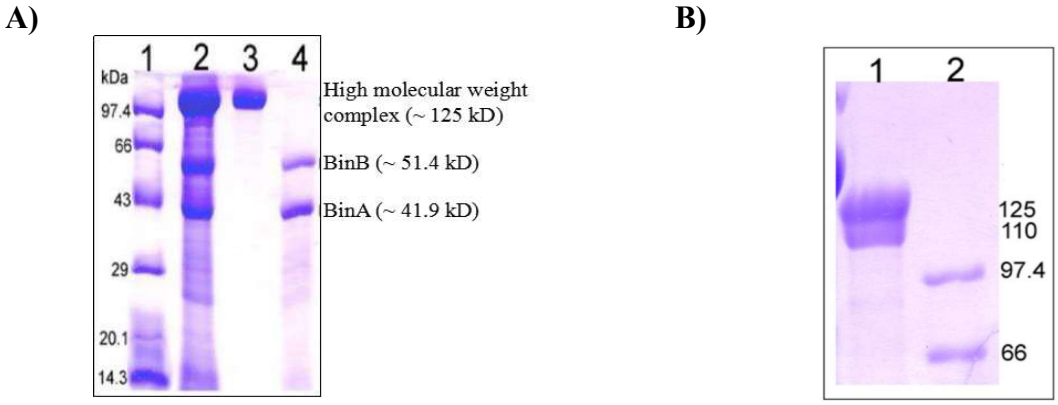


Figure 3.1. Purification of proteins from the spores of *L. sphaericus* ISPC-8. (A) 12% SDS-PAGE analysis of proteins purified from the spore-crystal mixture using ion-exchange and molecular sieve column chromatography. Lane-1, molecular weight markers; Lane-2, spore-crystal proteins purified using anion-exchange chromatography; Lanes-3&4, High molecular weight precursor proteins and BinA/B proteins purified using Superdex-200 column chromatography, respectively. **(B)** 5% SDS-PAGE gel showing two distinct bands at 125 and 110 kDa (Lane-1) of high molecular weight precursors; molecular weight marker (Lane-2).

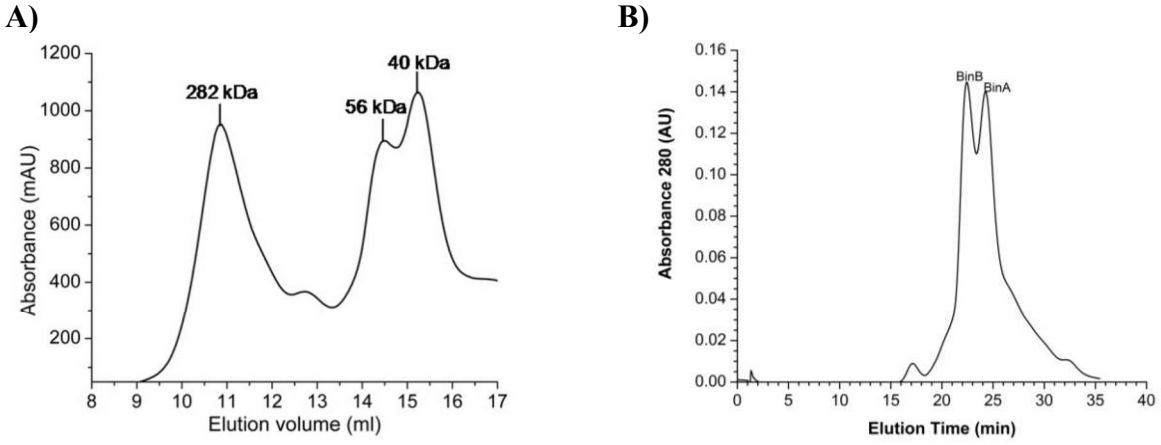


Figure 3.2. Elution profile from gel-filtration chromatography of proteins purified from spore-crystals of *L. sphaericus*. A) The high molecular weight protein elutes from Superdex-200 column corresponding to mass of 282 kDa. **B)** BinA and BinB proteins (corresponding to proteins of masses of 56 and 40 kDa) are well resolved in the elution profiles from Superdex-75 molecular sieve column.

The high molecular weight precursor proteins were further probed with dynamic light scattering experiments and by mass spectrometry. The analysis of the major peak observed in the dynamic light scattering experiments showed that purified high molecular weight proteins were polydispersive (percentage polydispersity ~24.5) and had a molecular weight of about 244 ± 59 kDa corresponding to hydrodynamic radius of about 5.85 nm. The peptide mass fingerprinting and MS/MS analyses revealed with high confidence (MASCOT score, 158; Expect score, 3.9E-09) that the two SDS-PAGE bands of high molecular weight proteins correspond to surface layer protein (SlpC) of *L. sphaericus*. All the 21 peptide masses used in MASCOT search matched with sequence of SlpC with the total protein sequence coverage of 23% (Fig. 3.3).

3.4.3. Toxicity analysis of the high molecular mass protein

The spore-crystal mixture of *L. sphaericus* showed 100% mortality of *Culex* larvae after 48 hr. The toxicity of SlpC (high molecular weight protein) purified from spore-crystal of *L. sphaericus* ISPC-8 was assayed against third instar larvae up to the concentration of 150 µg/mL. The purified SlpC protein was not toxic to *Culex* mosquitoes up to 100 µg/mL protein concentration but showed poor toxicity (20% mortality) at 150 µg/mL, as compared to control experiments without the SlpC protein. In contrast, the BinA/B mixture (purified from spores) and recombinantly expressed active BinA alone have shown toxicity with an LC_{50} value of ~6 ng/mL and ~16 ng/mL, respectively [32, 88].

Figure 3.3. Sequence coverage diagram in the MASCOT analysis of SlpC protein. Blue boxes indicate peptides matched by the peptide mass fingerprinting with the red boxes indicating *b*- and *y*- ions from the MS/MS spectra. The amino acids shown in red constitute the trypsin generated peptide fragments those were detected in MALDI-TOF experiments within ±1.2 Da.

3.5. DISCUSSION

Formation of high-molecular weight precursors (~125 kDa) of BinA/B toxin in *L. sphaericus* 2362 spores and their mosquitocidal activity has been intensely debated. These precursors migrate corresponding to proteins of molecular masses of 110 kDa and 125 kDa in reducing as well as non-reducing PAGE, and hydrodynamic radius of 5.6 ± 1.2 nm has been estimated for these [25, 27, 37]. Attempts have also been made to probe these precursor proteins with biophysical methods and with antisera raised against BinA and BinB proteins. The earlier observed cross-reactivity of precursor proteins with the antiserum of BinA and BinB polypeptides was subsequently rationalized to be due to contamination of precursor peptides in BinA and BinB preparations used for raising antisera [89].

The high molecular weight precursors and binary proteins, purified to homogeneity from spore-crystals of highly toxic local strain of *L. sphaericus* ISPC-8, were clearly resolved by molecular sieve column chromatography using Superdex-200, suggesting that high molecular weight proteins and BinA/B components may not form a stable complex. The molecular mass of the high molecular weight proteins was estimated to be 282 kDa based on their elution profile on the molecular sieve Superdex-200. This matches with dynamic light scattering experiments which showed that purified high molecular weight proteins were polydispersive (percentage polydispersity ~24.5%) and had a molecular weight of about 244 ± 59 kDa. The protein migrates as ~125 kDa protein on 12% SDS-PAGE; however, two bands of approximately 125 and 110 kDa were observed for the high molecular-weight proteins when resolved on 5% SDS-PAGE. The peptide mass fingerprinting and MS/MS analyses revealed with high confidence that the two SDS-PAGE bands of high molecular weight proteins correspond to surface layer protein (GenBank ID: AAA50256.1) of *L. sphaericus*. With the whole genome sequence of at least three *L. sphaericus* species being

available in databases, the MASCOT analysis provides unambiguous identification of high molecular-weight protein of *L. sphaericus* spores as SlpC protein.

The observation of proteins of 110 and 125 kDa on SDS-PAGE in the cell extracts of *Synechococcus*, *B. thuringiensis* and *E. coli* expressing *L. sphaericus bin*A and *bin*B genes [90−92] were considered as an evidence of these high molecular-weight proteins to be composed of BinA and BinB components [37]. BLAST search against NCBI database with *L. sphaericus* SlpC query sequence revealed the presence of *slp*C-like genes in *B. thuringiensis* (GenBank IDs: WP_021727980.1, ADU04484.1, AAY28601.1) and in *Synechococcus* (GenBank Id; YP_001226520.1). However, a homologous *slp*C-like gene was not detected in the genome of *E. coli*.

The SlpC protein of *L. sphaericus* (Mr, 125.2 kDa) is predicted to be secreted with a leader peptide that is removed between Ala-30 and Ala-31 of the primary translation product and is a precursor of the 110 kDa peptide [89]. The molecular weight and hydrodynamic size estimations by molecular sieve column chromatography and dynamic light scattering suggest that a dimer of SlpC constitutes a major fraction in the solution form, though higher oligomers may also be possible. This contrasts with earlier studies which suggested that high-molecular-weight SDS–PAGE bands were mono-dispersive and had a molecular weight of 186 ± 38 kDa [37].

The mosquitocidal toxicity of SlpC protein from several *L. sphaericus* strains has been studied. The purified 110 kDa protein of *L. sphaericus* 2362 was found to be toxic to the larvae of *Culex pipiens* with LC_{50} values of 115 ng/mL [93]. The purified surface layer protein (10-80 µg) of wild type *L. sphaericus* C3-41 and recombinantly expressed in *E. coli* BL21 (100 µg) were not toxic to third-instar *Culex* larvae [94]. Whereas S-layer proteins from *L. sphaericus* strains OT4b25, OT4b26 and III (3)7 have recently been reported to exert mild toxicity towards third-instar *Culex* larvae (LC_{50}, 0.68 – 24 µg/mL) [95, 96]. The purified

SlpC protein in the present studies showed poor toxicity (20% mortality) at 150 µg/mL, as compared to control experiments without the SlpC protein. Intriguingly, the SlpC protein is highly conserved in different strains of *L. sphaericus* [94]. Whether the observed differences in SlpC toxicity are due to dose of the protein tested or due to processing of 125 kDa protein needs further investigation.

3.6. CONCLUSIONS

- The high molecular weight (110 and 125 kDa) proteins of *L. sphaericus* spore-crystals do not represent oligomeric binary toxin.
- These high molecular weight (110 and 125 kDa) proteins are products of highly conserved surface layer protein (SlpC) of *L. sphaericus*.
- The purified SlpC protein of ISPC-8 strain showed poor toxicity against the third instar *Culex* larvae.

Chapter 4

Receptor protein of *Lysinibacillus sphaericus* mosquito-larvicidal toxin displays amylomaltase activity

4.1. ABSTRACT

The activated BinA/B toxin from *Lysinibacillus sphaericus* binds to surface receptor protein (Cqm1) on the midgut cell membrane and kills *Culex quinquefasciatus* larvae on internalization. Cqm1 is attached to cells via a glycosyl-phosphatidylinositol (GPI) anchor. It has been classified as a member of glycoside hydrolase family 13 of the CAZy database. The present chapter reports characterization of the ordered domain (residues 23-560) of Cqm1. Gene expressing Cqm1 of BinAB susceptible mosquito was chemically synthesized and the protein was purified using *E. coli* expression system. Values for the Michaelis-Menten kinetics parameters towards 4-nitrophenyl α-D-glucopyranoside (α-pNPG) substrate were estimated to be 0.44 mM (*K*m) and 1.9 s^{-1} (*k*cat). Thin layer chromatography experiments established Cqm1 as α-glucosidase competent to cleave α-1,4-glycosidic bonds of maltose and maltotriose with high glycosyltransferase activity to form glucose-oligomers. The observed hydrolysis and synthesis of glucose-oligomers is consistent with open and accessible active-site in the structural model. The protein also hydrolyses glycogen and sucrose. These activities suggest that Cqm1 may be involved in carbohydrate metabolism in mosquitoes. Further, toxic BinA component does not inhibit α-glucosidase activity of Cqm1, while BinB reduced the activity by nearly 50%. The surface plasmon resonance study reveals strong binding of BinB with Cqm1 (*K*D, 9.8 nM). BinA interaction with Cqm1, however, are 1000-fold weaker. Notably the estimated *K*D values match well with dissociation constants reported earlier with larvae brush border membrane fractions. The Cqm1 protein forms a stable dimer that is consistent with its apical localization in lipid rafts. Its melting temperature (T_m) as observed by thermofluor-shift assay is 51.5 ºC and Ca^{2+} provides structural stability to the protein.

4.2. INTRODUCTION

The BinAB toxin binding receptor in *Culex quinquefasciatus* has been identified as Cqm1 (*C. quinquefasciatus* maltase 1) protein, which is anchored to the apical membrane of the midgut epithelium cells via glycosyl-phosphatidylinositol (GPI) tail [28, 29]. Cpm1 (*C. pipiens* maltase 1), ortholog of Cqm1, was firstly described and was found to display α-glucosidase activity [28, 29]. Homologous Agm3 and Aam1 proteins have also been characterized in *Anopheles gambiae* and *Aedes aegypti* species, respectively. Recently, enzymatic property of the receptor protein from *Culex* and *Aedes* has been investigated using substrate 4-nitrophenyl α-D-glucopyranoside (α-pNPG), and *Aedes* Aam1 protein was found to display remarkably higher catalytic efficiency as compared to Cqm1 [97]. Cqm1 and its orthologs have been classified into GH13 subfamily 17 (GH13_17) of the CAZy database, which includes α-amylases, α-glucosidases, cyclodextringlycosyltransferases (CGTase), branching enzymes and cyclomaltodextrinases (http://www.cazy.org/;) [98, 99].

α-glucosidases are essential for carbohydrate metabolism in insects and show diverse substrate specificities (against maltose, maltose-oligosaccharides, sucrose etc.). These preferentially hydrolyse 1,4-glycosidic bonds of α-D-glucosides at nonreducing ends. Three types of α-glucosidases have been defined broadly based on their substrate specificity [100]. Type I α-glucosidase hydrolyses heteroside substrate (sucrose and aryl-glucoside) more rapidly than holoside substrate (malto-oligosaccharides). Whereas, type II and type III enzymes prefer holoside substrate. Type III is also capable of hydrolysing polysaccharide substrate. Some α-glucosidases are competent to catalyse transglycosylation to produce α-1,4 linked maltose oligosaccharides [101].

BinAB does not display toxicity against *Aedes aegypti* mosquitoes. The recent crystal structure of BinAB [40] is expected to facilitate broadening of BinAB specificity against refractory mosquito species. Also, resistance to BinAB in susceptible mosquito population is

mostly reported due to mutation in the receptor [59−62, 102]. Characterizing biological activity of Cqm1 is important in determining its physiological role in mosquitoes, as this could provide novel leads for mosquito control programmes. Also, understanding the relationship between the physiological role of Cqm1 and its function as receptor of *L. sphaericus* BinA/B toxin might give a new insight in elucidating the mechanism of action of binary toxin [97].

The present chapter provides a detailed biochemical, biophysical and enzymatic characterization of homogeneously purified Cqm1 receptor protein from *Culex quinquefasciatus*. Cqm1 forms stable dimer and Ca^{2+} was found to enhance thermal stability of the protein. We observed that Cqm1 hydrolyses substrates with α-1,2- and α-1,4-glycosidic bonds and, interestingly, displays glycosyltransferase activity with glucose-oligomers. We also observed that BinB, and not BinA, can reduce catalytic activity of Cqm1 protein. BinB also binds Cqm1 with high affinity, whereas BinA interaction with the receptor is ~1000 fold weaker.

4.3. MATERIALS AND METHODS

4.3.1. Materials

Restriction enzymes and *Taq* polymerase were purchased from New England Biolabs. pNIC28-Bsa4 vector was from Opher Gileadi (Addgene plasmid # 26103). Luria Bertani (LB) broth and LB agar were procured from HiMedia Laboratories. *E. coli* XL-10 gold and BL21Star (DE3) bacterial strains were from Stratagene and Novagen, respectively. Superdex^{TM} 200 10/300 GL column, gel filtration protein molecular weight markers and Ni–IDA matrix were from GE Healthcare. High Q cartridge and Thin Layer Chromatography (TLC) plates (Silica gel 60 F254) were from Bio-Rad Laboratories and Merck, respectively. Isopropyl β-D-1-thiogalactopyranoside (IPTG), ethylene glycol-bis(β-amino ethyl ether)-N,N,N',N'-tetraacetic acid (EGTA), phenylmethylsulfonyl fluoride (PMSF), Trypsin (proteomics grade, from porcine pancreas), thrombin (from bovine plasma), SYPRO orange dye and α-pNPG were obtained from Sigma. Sugars and other chemicals were procured from SRL, India.

4.3.2. Mosquito colony

A BinAB susceptible *Culex quinquefasciatus* colony was established from local strain found in Trombay and Mankhurd regions, Mumbai, India, and were reared without exposure to *L. sphaericus*. The susceptibility was tested against equimolar concentrations of BinA/B proteins [33]. *Culex* nucleus culture was maintained at 28 ± 2°C and 85% relative humidity in a BARC laboratory.

4.3.3. Plasmid constructs

The *cqm*1 gene of BinAB susceptible *Culex quinquefasciatus* mosquito colony was PCR amplified and sequenced (GenBank accession number: KY929304). The local strain showed polymorphism at six nucleotide sites (GenBank accession number: KY929305). The gene is 1840 bp long with a 1743 bp long coding sequence (CDS), which encodes for 580

amino acid long Cqm1 protein. A *cqm*1 gene with CDS of 1743 bp was synthesized chemically with codons optimized for expression in *E. coli* (GenBank accession number: MG211157) through the facilities available with Biotech desk Pvt. Ltd (Bangalore, India) based on *cqm*1 sequence submitted to GenBank (accession number: KY929304).

For expressing soluble Cqm1 protein, truncated synthetic *cqm*1 gene (1614 bp) lacking 66 bases from the 5'-end and 60 bases from the 3'-end was PCR amplified using forward primer 5'-TAC TTC CAA TCC ATG CGT GAA CCG GAC TCA AAA G-3' and reverse primer 5'-TAT CCA CCT TTA CTG CGC GGA ACT CAC GGT CGC-3'. The PCR amplified truncated *cqm*1 gene was cloned into pET28a based pNIC28-Bsa4 expression vector with an N-terminus poly-His tag (pNIC28-*cqm*1 construct) using ligation independent cloning method and was transformed into the cloning host, *E. coli* XL-10 gold competent cells. The transformed cells were plated onto LB agar plates containing kanamycin (50 µg/mL) and incubated overnight at 37 ºC. Positively transformed colonies were selected through colony PCR using Taq DNA polymerase. The recombinant plasmid was isolated using plasmid purification kit (Qiagen) and was transformed into BL21 Star (DE3) competent cells for expression of recombinant Cqm1 protein (538 residues) lacking disordered signal and GPI-anchoring amino acid residues, and fused with an N-terminus poly-His tag. The tag consists of 6 histidine residues and 16 other residues including TEV protease site.

4.3.4. Protein purification

The *E. coli* BL21 Star (DE3) culture, harbouring pNIC28-*cqm*1 construct, was grown O/N at 37 ºC. The pre-culture was diluted (1:100) into fresh LB broth supplemented with kanamycin (50 µg/mL) and was grown at 37 ºC until the optical density at 600 nm reached approximately 0.7. The temperature for culturing was then reduced to 18 ºC, and the inducer IPTG was added to a final concentration of 0.5 mM. After overnight growth, cells were harvested by centrifugation at 10,000 x*g* for 5 min. The cell pellet was suspended in the lysis

buffer [25 mM Tris–HCl pH 8.0, 100 mM NaCl, 5% sucrose, 1 mM dithiothreitol, 1 mM PMSF, 1 mg/mL lysozyme and 1 tablet of Complete protease inhibitor cocktail (Roche)] and subsequently sonicated in pulse mode (3 sec ON, 3 sec. OFF) by using Vibra Cell sonicator. Crude whole cell lysate was centrifuged at 21,000 x*g* at 4 ºC for 30 min. Recombinant Cqm1 protein was isolated from the cell free extract by poly-His tag immobilized metal ion affinity (IMAC) chromatography using Ni–IDA matrix which was pre-equilibrated with buffer A [25 mM Tris-HCl (pH 8), 100 mM NaCl] containing 25 mM imidazole. The bound proteins were eluted using imidazole concentration gradient ranging from 25–1000 mM in the same buffer. Imidazole was removed by overnight dialysis at 4 ºC against buffer B [25 mM Tris-HCl (pH 8), 25 mM NaCl]. For further purification, the protein was loaded onto anion exchange chromatography (AEC) column pre-equilibrated with buffer B. The bound Cqm1 protein was eluted using 25–1000 mM NaCl gradient in the same buffer. The buffer of the purified Cqm1 protein was exchanged to buffer B by overnight dialysis at 4 ºC. The protein concentration was estimated using modified Folin-Lowry method [79] using bovine serum albumin (BSA, Sigma) as a standard. Total of 12 mg Cqm1 was purified from 1 L *E. coli* culture. Purified protein was analysed by 15% SDS–PAGE.

Recombinant BinA and BinB proteins were purified as described earlier [32, 33]. Briefly, recombinant BinA (active form; 15–366 amino acids) and BinB (residues 1–480) were expressed in BL21 (DE3) and pLysS host cells, respectively, as fusion proteins with N-terminal poly-His tag and a thrombin cleavage site between poly-His tag and target protein. The recombinant proteins were purified from soluble cell lysate using IMAC and AEC matrices. Purified BinA was treated with thrombin for removal of poly-His tag. Purified BinB protein was treated with trypsin (trypsin: BinB: 1:1000, w/w) at 20 ºC overnight. The reaction was stopped with 10 mM PMSF. Thrombin treated BinA and trypsin activated BinB (residues

17–448) were further purified by AEC. Buffer of BinA and BinB proteins was exchanged to buffer B by overnight dialysis.

4.3.5. Protein identification by MALDI-TOF/TOF

The protein band from 15% SDS-PAGE gel was excised and in-gel digestion with trypsin (proteomics grade solution, Sigma) was carried out using the protocol adapted from Shevchenko et al [103]. Peptides were analysed by peptide mass fingerprinting and MS analysis with a MALDI-TOF/TOF (Bruker Daltonics). External calibration was performed using the Peptide Standard Calibration II (Bruker Daltonics). Spectra were processed using the FlexAnalysis software (Bruker Daltonics). The detected peptides were matched against the putative peptides in Cqm1 sequence by FindPept tool [104] of ExPASY using monoisotopic masses as $[M+H]^+$ and tolerance of ±1 Da.

4.3.6. Oligomeric status of Cqm1 protein

The recombinant Cqm1 protein, treated with TEV protease, was loaded onto size-exclusion chromatography (SEC) Superdex^{TM} 200 10/300 GL column, pre-calibrated with gel filtration protein molecular weight markers (Ribonuclease A, 13.7 kDa; Carbonic anhydrase, 29 kDa; Ovalbumin, 44 kDa; Conalbumin, 75 kDa; Aldolase, 158 kDa; Ferritin, 440 kDa; Thyroglobulin, 669 kDa). The eluted peaks from gel filtration experiments were adjudged on 10% native-PAGE and 15% SDS–PAGE. The SEC purified protein was used for biochemical and biophysical analysis. The integrity and identity of the recombinant protein was confirmed by mass spectrometry and peptide mass fingerprinting.

The purified Cqm1 protein (21 μM) in buffer A was analysed by dynamic light scattering (DLS) using Malvern Zetasizer Nano ZS instrument, to determine its hydrodynamic size. Three sets of measurements with 12 acquisitions in each set were collected at 25 °C. The data analysis was done with the help Malvern Zetasizer software suite (http://www.malvern.com). To assess stability of Cqm1 oligomer, the protein was incubated

with 0.05% triton and with 500 mM sodium chloride in independent reactions at room temperature for 1 hr and the protein profile was observed through DLS analysis.

4.3.7. Folding status and thermal stability of Cqm1 protein

Circular dichroism (CD) spectrum for Cqm1 protein (0.75 mg/mL, 12 μM) in buffer C (25 mM potassium phosphate, pH 8.0) was measured on a JASCO spectrometer (J-815) in the far-UV region (200–260 nm) at 20 ºC in 1 mm pathlength quartz cuvette. Each spectrum was obtained by averaging three individual scans. The spectrum of the buffer blank was subtracted from the protein CD spectrum. Deconvolution and analysis of the CD spectra were carried out using K2D2 software (http://www.ogic.ca/projects/k2d2) [80]. Thermal denaturation profile of the protein was recorded using Circular dichroism ellipticity at 210 nm through variable-temperature scan in the temperature range 20–75 ºC with a scan rate of 1 ºC/min in two buffer conditions; buffer B and buffer C.

Intrinsic fluorescence of Cqm1 protein (20 μM) with or without 8 M urea was monitored on a JASCO spectrofluorometer (FP-8500) using a 1 cm cuvette, with excitation at 280 nm and emission spectrum recorded over 300–400 nm at 25 ºC. Each spectrum was obtained by averaging three individual scans. Intrinsic fluorescence measurements were also performed with Cqm1 protein (20 μM) in buffer D [20 mM HEPES buffer (pH 7.5) 100 mM NaCl] in the presence of 5 mM EGTA. Each fluorescence spectrum was obtained by averaging three individual scans.

Thermofluor-shift melting curve measurements were carried out to study thermal stability of Cqm1 protein as described by Niesen et al [81]. The measurements were performed in a 96 well plate in 25 μL reaction, in triplicates, containing recombinant Cqm1 protein (2 μM) in buffer C and freshly diluted SYPRO orange dye (1:1000, v/v). The plate was sealed and data was recorded on CFX96^{TM} Real time system (Bio-Rad) in FRET mode. The protein samples were heated from 20 to 95 ºC at ramp rate of 1.0 ºC/ min. Thermal

unfolding of Hen Egg White (HEW) lysozyme was also probed in the same 96 well plate as an internal control. The observed T_m of 71.5ºC for HEW lysozyme matches with the values reported in literature. For probing thermal stability due to binding of Ca^{2+}, the measurements were also performed with Cqm1 (2 µM) in buffer D and freshly diluted SYPRO orange (1:1000, v/v), in the presence of EGTA (5 mM) or EGTA (5 mM) and $CaCl_2$ (0.5, 1.0 and 10.0 mM). Experimental data were processed using software available with CFX96^{TM} Real time system and melting temperatures were determined from the first derivative of each of the thermofluor-shift melting curve.

4.3.8. Enzymatic assay

α-glucosidase activity of Cqm1 protein was assayed by Michaelis–Menten kinetics using α-pNPG as substrate. The enzyme activity was determined in 800 µL reaction mix with 65 nM Cqm1 in buffer E (50 mM sodium phosphate buffer, pH 6.8) and 100 µM glutathione (reduced) with α-pNPG concentration varying from 0.05 to 4 mM. The kinetic reaction was measured at 25 ºC over a period of 20 min for each substrate concentration with data recording at an interval of 2 min using 1 cm pathlength quartz cuvette. The amount of p-nitrophenol produced during the course of the reaction was measured at 405 nm using the Time Course Measurement Program available with JASCO UV-Vis spectrophotometer. The measured absorbance at 405 nm (OD_{405}) was plotted against time, and initial velocity (V_o) for each substrate concentration was estimated from the gradient. The obtained values of V_o were plotted against substrate concentration. The kinetic constants (*V*max, *K*m and *k*cat) were calculated by non-linear regression based on using Hill equation (Eq 1) with value of 1 for n using Origin software version 6, [S] being substrate concentration.

$$V = \frac{V\max\ [S]^n}{(K\mathrm{m})^n + [S]^n} \qquad \ldots\ldots(1$$

Thin layer chromatography (TLC) was performed to identify the specificity of Cqm1 protein against α-1,1-glycosidic, α-1,2-glycosidic, α-1,4-glycosidic and α-1,6-glycosidic bonds. Trehalose with α-1,1-glycosidic bond, sucrose with α-1,2-glycosidic bond, maltose and maltotriose with α-1,4-glycosidic bond, and isomaltose with α-1,6-glycosidic bond, were used as substrates. The polysaccharide glycogen was also used as substrate in an independent experiment. Cqm1 (5 µM) was incubated with individual sugars (1: 250, w/w) in 50 µL reaction volume at 37 ºC for about 2 and 24 hr in independent reactions and also for about 5 hr with glycogen and sucrose. Higher amounts of Cqm1, compared to kinetic analysis, were used to detect products in TLC experiments. The products from the reaction of Cqm1 with glucose, trehalose, sucrose, maltose, maltotriose, isomaltose and glycogen were resolved on silica gel based TLC plates using chloroform /acetic acid /water (3: 3.5: 0.5, v /v) as the solvent. The products were detected by incubation in an acetone solution of $AgNO_3$ and 40% NaOH/ethanol (1: 19, v⁄v).

4.3.9. Effect of temperature, pH, Ca^{2+} and BinA/B proteins on enzymatic activity

The effect of temperature on α-glucosidase activity of Cqm1 was investigated over a temperature range of 25–60 ºC using 65 nM Cqm1 in buffer E. The effect of pH on α-glucosidase activity of Cqm1 (65 nM) was analysed at 37 ºC over a pH range of 4.6–11 using sodium acetate buffer (pH 4.6 and 5.6), potassium phosphate buffer (pH 5, 8 and 8.5), sodium citrate buffer (pH 6), sodium phosphate buffer (pH 6.8), HEPES buffer (pH 7 and 7.5) and CAPS buffer (pH 10 and 11). The ionic strength of different buffer systems used in this experiment may vary and the present study cannot rule out effect of such variation on activity. The effect of Ca^{2+} on α-glucosidase activity of Cqm1 protein was examined using Cqm1 (65 nM) in buffer D. Cqm1 protein was incubated with EGTA (5 mM) for about 1 hr and the excess of EGTA was removed by dialysis. The dialysed protein was incubated with 10 mM Ca^{2+} at 37 ºC for about 1 hr prior to addition of substrate. For probing the effect of

BinA/BinB proteins on enzymatic activity against substrate α-pNPG, Cqm1 (65 nM) was incubated with BinA or BinB proteins (300 nM in buffer E) for about 1 hr at 37 ºC, prior to addition of the substrate. In the control experiments, BinA and BinB did not display observable hydrolytic activity against α-pNPG.

These activity assays were performed in 200 µL reaction volume using 100 µM glutathione (reduced) and substrate α-pNPG (1 mM). Incubation time for the enzymatic reaction was 20 min and the reaction was stopped by adding 800 µL of 100 mM sodium carbonate buffer. The absorbance of accumulated p-nitrophenol product was measured at 405 nm using JASCO UV-Vis spectrophotometer. Absorbance for each reaction was converted into activity units (µmol/min/mg). The optimum temperature and pH, and effect of Ca^{2+} and BinA/BinB, were determined by plotting the activity units.

4.3.10. Surface Plasmon Resonance analysis of BinA and BinB interactions with the receptor protein

The instrument used for surface plasmon resonance (SPR) analysis was Biacore T200 SPR system with CM5 Biacore sensor chip. For each experiment at 25 ºC nearly 1240 response units of Cqm1 (ligand) were immobilized onto the CM5 Chip via amine coupling followed by extensive washing with buffer F [50 mM potassium phosphate (pH 8) and 50 mM NaCl]. Different concentrations (0.039, 0.078, 0.156 and 0.3125 µM) of aqueous BinA and BinB (analytes) in buffer F were injected onto the Cqm1-bound sensor chip at 30 µL/min flow-rate in independent experiments. Final experiments for BinA were performed using higher concentrations of the protein (1.25, 2.5, 5, 10 and 20 µM), as lower BinA concentrations did not result in significant response units (RU). The analyte was allowed to interact with the immobilized Cqm1 for 60 sec before washing off with buffer F. The data were processed and association/dissociation rate constants were analyzed using Biacore kinetic evaluation software (V3.1) provided with the instrument.

4.3.11. Structure modelling, molecular dynamics and bioinformatics analysis

The online servers of Modeller (http://toolkit.tuebingen.mpg.de/modeller) [105] and SWISS-MODEL (http://swissmodel.expasy.org) [106] were used to generate homology models for Cqm1 protein. Several templates with sequence identity between 30–36% with Cqm1 sequence were selected by both the servers. The accepted structural model corresponding to residues 23–556 of Cqm1 and a Ca^{2+} were constructed using the template structure of trehalulose synthase (MutB) from *Pseudo mesoacidophile* (PDB code, 1ZJA) [107]. Sequence identity of 34% (similarity 38%) between Cqm1 and MutB was returned by the servers. Molecular dynamics (MD) simulations of the constructed Cqm1 model were performed using OPLS-AA/L force field and GROMACS package [108]. The overall electric charge of the system was compensated by adding Na^{+} ions. The structural models derived from SWISS-MODEL with or without Ca^{2+}were subjected to initial energy minimization, 100 ps temperature equilibration at 298 K using Berendsen thermostat and 100 ps pressure equilibrium at 1 atm using Berendsen barostat. The production run trajectory was simulated for 40 ns in 10,000,000 steps of 4 fs each and structures were written every 10 ps.

4.4. RESULTS

4.4.1. Purification of recombinant proteins

Attempts to purify recombinant Cqm1 protein of 580 amino acids using codon optimized synthetic gene and different expression systems, including GST and trigger factor translational fusion with Cqm1 sequence, were not successful. Three major problems were encountered: 1) protein expression was not observed, 2) total of expressed protein was in inclusion body, and 3) refolding of the protein from detergent solubilized inclusion body did not yield folded protein. The Cqm1 protein sequence was subjected to analysis for intrinsically disordered regions (http://bioinf.cs.ucl.ac.uk/psipred/) [109]. The N-terminus 22 residues and C-terminus 20 residues were identified as disordered (Fig. 4.1). The truncated *cqm*1 gene (1614 bp; Fig. 4.2A) lacking 66 bases from the 5'-end and 60 bases from the 3'-end was sub-cloned into pET28a based pNIC28-Bsa4 expression vector for purification of the protein using *E. coli* expression system. The N-terminal poly-His tagged recombinant Cqm1 protein, lacking N-terminus 22 signal peptide residues and 20 C-terminus GPI anchoring residues, was purified to homogeneity by three stage column chromatography methods; IMAC, AEC and SEC (Fig. 4.2B). The mass of 62378 Da determined by MALDI-TOF matches with estimated mass of 62329 Da deduced from the amino acid sequence encoded by the cloned gene. The peptide mass fingerprinting analysis confirms identity of the purified protein to be Cqm1 (Fig. 4.3). Recombinant BinA (active form; 15–366 amino acids) and BinB (residues 17–448) were also purified to homogeneity using IMAC and AEC chromatography methods (Fig. 4.2C). Western blot analysis using anti-His antibodies confirmed complete removal of N-terminus fusion tag from the purified Cqm1, BinA and BinB proteins (Fig. 4.2D).

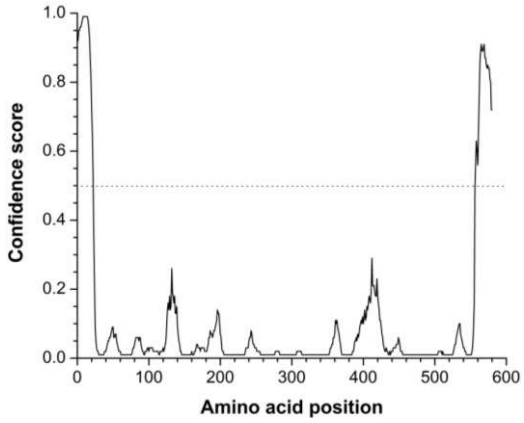


Figure 4.1. Prediction of disordered regions in the Cqm1 sequence (full length, 580 amino acid residues) using DISOPRED 3 tool. Amino acids in the input sequence are considered disordered when the black solid line is above the grey dotted line, that is the confidence score is higher than 0.5. Residues at the extreme N- and C- termini were predicted to be disordered by the server.

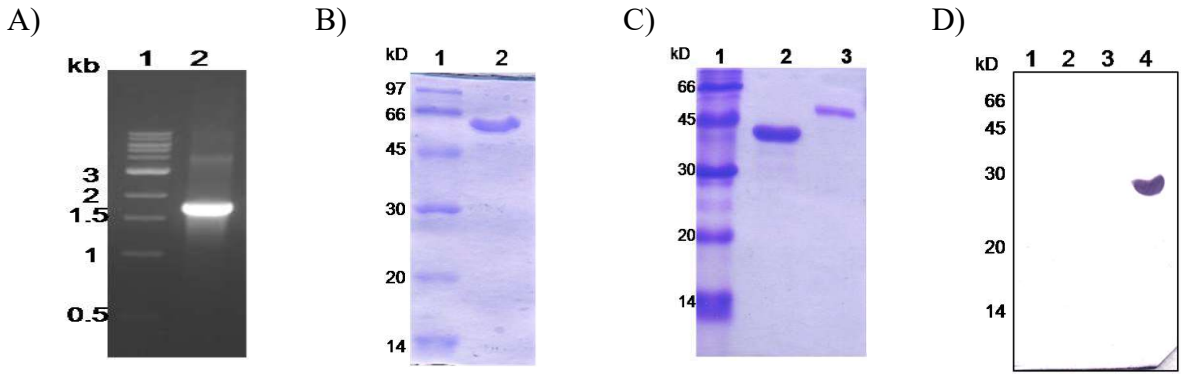
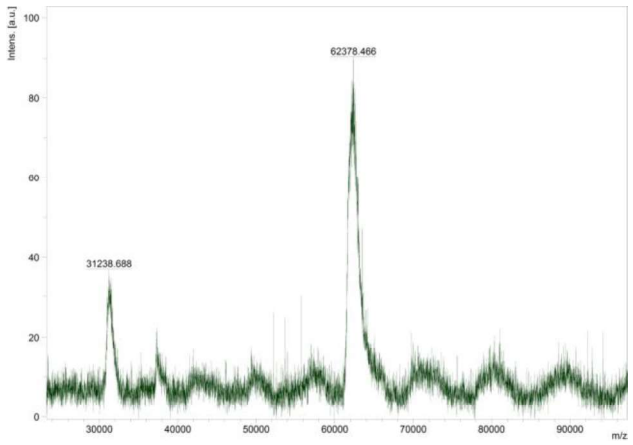


Figure 4.2. Cloning and expression of *cqm*1 gene. A) 1% agarose gel showing PCR product of amplified 1614 bp synthetic DNA from pUC57-*cqm*1 construct (lane 2). DNA ladder (1 kb) is also shown (lane 1). **B)** 15% SDS-PAGE showing purified Cqm1 (lane 2). Also shown are molecular weight markers (lane 1). **C)** 15% SDS-PAGE showing purified BinA (lane 2), BinB (lane 3). Also shown are molecular weight markers (lane 1). **D)** Western blot analysis of Cqm1 (lane 1), BinB (lane 2) and BinA (lane 3) proteins resolved on 15% SDS-PAGE using anti-His antibodies. Human translin (Mr, 27.5 kDa; lane 4) protein was used as an internal positive control for poly-His tag. The immune detection was visualized by secondary antibody and NBT/BCIP colour development reagents. Cqm1, BinB and BinA proteins without poly-His tags were not immunostained, while translin with intact poly-His tag was clearly visible.

A)



B)

SMREPDSKDWYQHATFYQIYPRSFLDSNGDGIGDLAGITSKMKYLADIGIDATWLSPPFK
SPLKDFGYDVSDFYAIQPEYGNLTDFDKLVEESHKNGIKLMLDFIPNHSSDQHEWFVKSV
VRDPEYSDFYVWRPPATGGGPPNNWISVFGGSAWTYNQARGEYYLHQFTPQQPDLNYRNP
KVLAEMTKMLFFWLDRGVDGFRLDAINHMFEDEQFRDEPLSGWGQPGEYDSLDHIYTKDI
PDVYDVVYNWRDQMDKYSAEKGRTIILMTEAYSSIEGTMLYYESADRKRQGAHMPFNFQL
IYDFKKEQNAVGLKSSIDWWMNNMPARHTPSWVAGSHDHSRVASRVGLDRVDQVMTLLHT
LPGTSITYYGEEVAMQDFKEAQQFDNRDPNRTPMQWDSSTSAGFSTNTNTWLRVHPDYAR
YNVDVMQKNPQSTFHHFQHLTKLRRHRTMQSGEYVHKTVGTKVYALLRELRGEDSFLTVL
NMAGAEDTVDLGDFVNLPQKMRVEVAQPNSKSKAGNEVDISKLTLGPYDSVVLRATVSSA
Q

Figure 4.3. MASS spectrometry analysis of recombinant Cqm1 protein. A) The MALDI-TOF analysis of Cqm1 proteins showed single large spectral peaks at m/z values of 62378 Da. Smaller peak at about 31239 Da correspond to doubly charged species of the protein.
B) The FindPept tool of ExPASY identified yellow shaded peptides from the observed MALDI peaks of the in-gel trypsin digested Cqm1.

4.4.2. Folding and oligomeric status of recombinant Cqm1 protein

The secondary structure conformation of recombinant Cqm1 protein was characterized by CD spectroscopy (Fig. 4.4A). The % secondary structure elements determined with the help of K2D2 web-tool reveals mixed α/β structure (α, ~8%; β, ~40%). Thermal denaturation of secondary structure of Cqm1 over a temperature range 20–75 ºC was carried out in two different buffer conditions; 10 mM Tris-HCl (pH 8), 10 mM NaCl, and 25 mM potassium phosphate buffer (pH 8). The protein does not show secondary structure denaturation in the probed temperature range in either of the buffer conditions.

Intrinsic protein fluorescence shows an emission maximum (λ_{EM}) at 335 nm, which is typical of partially buried tryptophan residues (Fig. 4.4B). Fluorescence spectral shift to red (λ_{EM} ~352 nm) and quenching of fluorescence intensity on incubation with 8M urea, owing to denaturation of the protein in the presence of chaotropic agent, suggest that the purified recombinant protein has folded tertiary structure. Also, protein fluorescence reduced marginally in the presence of EGTA that specifically chelates bound Ca^{2+}. Several members of GH13 have been reported to bind Ca^{2+}.

The apparent melting temperature of the tertiary structure (T_m) of Cqm1 is estimated to be 51.5 ºC by thermofluor-shift analysis in the presence of HEPES buffer (pH, 7.5) (Fig. 4.4C & D). Decrease in T_m (to 48 ºC) is observed in the presence of EGTA. The addition of 10 mM Ca^{2+} rescued T_m to 52 ºC suggesting that the divalent cation stabilizes the protein. A similar trend has also been observed for dextran glucosidase from *Streptococcus* mutant [110]. Influence of the buffer was also observed on T_m of Cqm1. The protein displays lower stability (T_m ~48 ºC) in the presence of phosphate buffer (pH, 8.0).

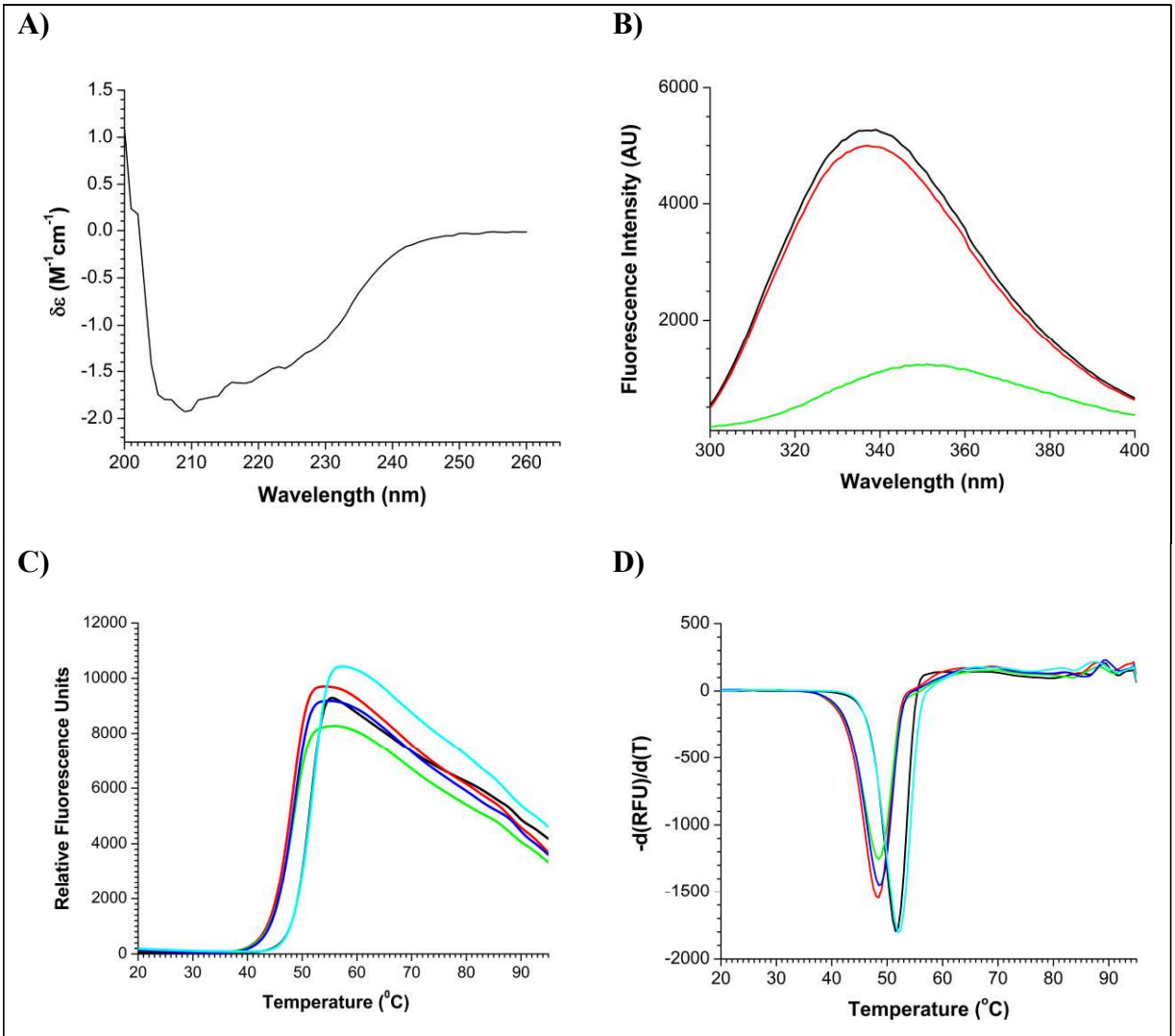


Figure 4.4. Recombinant Cqm1 adopts native structure. **A**) CD spectrum of Cqm1 protein (0.75 mg/mL protein in 25 mM potassium phosphate, pH 8.0). Deconvolution of the spectrum reveals mixed α/β structure (α, ~8%; β, ~40%). **B**) An emission maximum at 335 nm in intrinsic fluorescence of Cqm1 (black) is shifted to 352 nm with loss of intensity on denaturation with 8M urea (green). Fluorescence intensity also decreases marginally in the presence of EGTA (5 mM; red). **C**) Thermofluor-shift melting curve of Cqm1 (black), Cqm1 in the presence of 5 mM EGTA (red), and Cqm1 in the presence of EGTA (5 mM) and Ca^{2+} (0.5 mM, green; 1mM, blue; 10 mM, cyan). **D**) Derivative plot showing thermofluor-shift melting curve of Cqm1 (black), Cqm1 in the presence of 5 mM EGTA (red), and Cqm1 in the presence of EGTA (5 mM) and Ca^{2+} (0.5 mM, green; 1mM, blue; 10 mM, cyan).

The recombinant Cqm1 protein elutes as dimer on SEC column with an estimated molecular weight of 120 kDa suggesting homodimer status of the protein (Fig. 4.5A). The protein migrates as a single band on 10% native-PAGE with migration close to that of 140 kDa marker protein (Fig. 4.5B). A hydrodynamic diameter of about 7.6 ± 1.8 nm corresponding to a molecular weight of about 106 kDa is estimated in dynamic light scattering analysis. Marginal increase in the hydrodynamic radius of the protein is observed in the presence of 0.05% triton and 500 mM sodium chloride (Fig. 4.5C). However, there was no significant change in the polydispersity in the presence of triton and salt (Table 4.1). Data taken together suggests that Cqm1 exists in solution as a stable homodimer, and both hydrophobic and ionic residues may contribute towards dimeric interface.

Table 4.1: Dynamic Light Scattering analysis of Cqm1 protein. Parameter values were taken from the analyses using Malvern Zetasizer software suite.

Protein	Polydispersity (%)	Size, Diameter (nm)	Estimated MW (kDa)
Cqm1	22.2	7.6 ± 1.8	105.6 ± 23.4
Cqm1 + 500mM NaCl	22.3	8.0 ± 1.9	122 ± 27.2
Cqm1 + 0.05% Triton	21.4	8.1 ± 1.9	117.5 ± 25.1

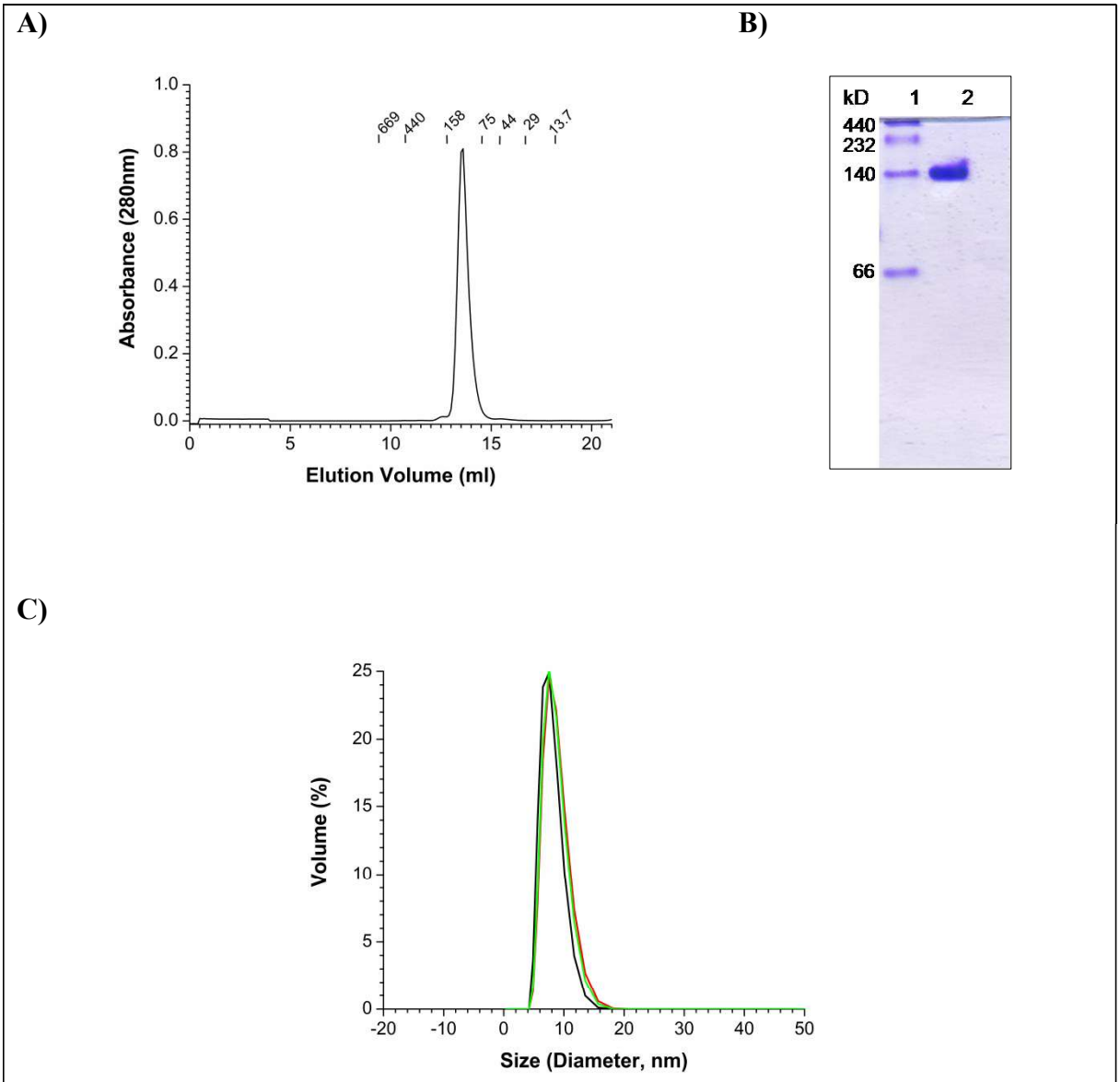


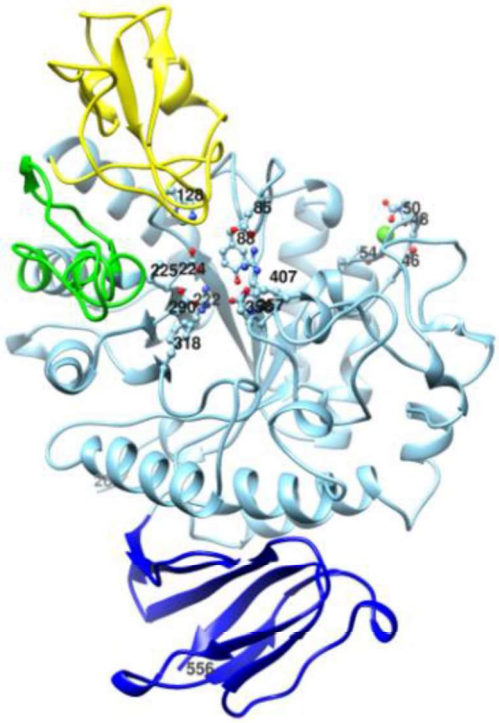
Figure 4.5. Oligomeric status of Cqm1. **A**) Size exclusion chromatography elution profile on Superdex 200 column in buffer A (25 mM Tris-HCl pH 8, 100 mM NaCl). Elution volume for the standard protein molecular weight markers (Ribonuclease A, 13.7 kDa; Carbonic anhydrase, 29 kDa; Ovalbumin, 44 kDa; Conalbumin 75 kDa; Aldolase, 158 kDa; Ferritin, 440 kDa; Thyroglobulin, 669 kDa) are also shown. **B**) 10% native-PAGE analysis; Lane 1, protein markers for native-PAGE; Lane 2, Cqm1 protein eluted from Superdex 200 column. The positions of Cqm1 band and molecular weight markers are identified. **C**) Dynamic light scattering analysis of Cqm1 (black), in the presence of 500 mM sodium chloride (red), and in the presence of 0.05% triton (green).

4.4.3. Structure modelling and molecular dynamics simulations

No structure of GH13_17 subfamily members is available in the Protein Data Bank. The homology models returned by MODELLER and SWISS-PDB were based on MutB template and were used for structure-activity analysis. The two models are similar with an RMSD of 0.49Å for 436 pair of atoms on superposition in Chimera suite [12]. The Cqm1 model (MODELLER scores: zDOPE score, -0.83; GA341 score, 1.0) is reliable, as active-site and Ca^{2+} binding residues of MutB align with the invariant residues of CAZy GH13_17 subfamily, and as stereochemical analysis of phi-psi dihedral angles indicates nearly 87.3% of residues in the most favored regions with none in the disallowed regions of the Ramachandran plot.

The structural models of Cqm1 reveals three independent domains (A–C), conserved Ca^{2+} binding residues (Asp-46, Asn-48, Asp-50 and Asp-54) and different conformations of loop regions (residues 241–245, 273–283, 306–311, 322–327, 369–373, 399–406, 509–517, 531–533). The domain A is a $(\beta/\alpha)_8$-barrel (residues 23–128 and 197–471), domain B is a loop between the β3 strand and α3 helix of domain A (residues 129–196), and domain C is an extension at the C-terminus characterized by a Greek key conformation constituted by residues 472–556 (Fig. 4.6A). The active site scaffold is constituted by Asp-85, Tyr-88, His-128, Arg-222, Asp-224, Ala-225, Glu-290, Phe-318, His-357, Asp-358 and Arg-407 of domain A (catalytic domain) (Fig. 4.7). These residues are invariant in the characterized proteins of CAZy GH13_17 subfamily. Catalytic triad (Asp-224, Glu-290 and Asp-358) can be expected to perform catalysis.

A)



B)

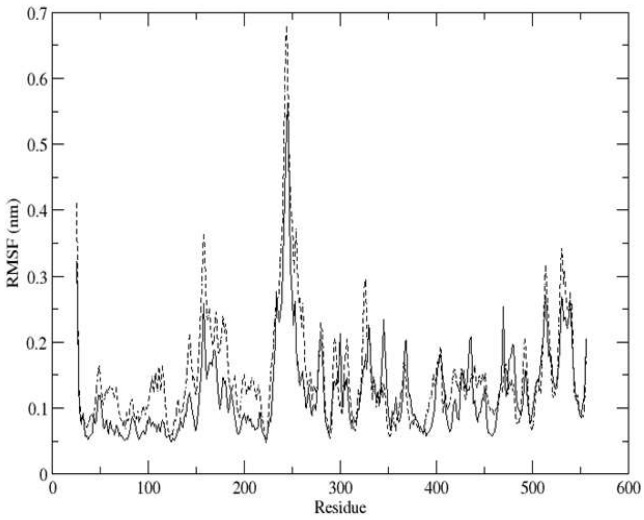


Figure 4.6. Structural model and MD simulations. **A**) Ribbon model of Cqm1 protein (domain A, cyan; domain B, yellow and domain C, blue) showing also active-site and Ca^{2+} binding residues in ball-and-stick representation. Ca^{2+} is shown as green sphere. The loop-4 between β4 and α4 elements is shown in green color. **B**) The root mean square fluctuation (RMSF) of 23-558 residues about the average position during the 40 ns of MD simulations is shown. RMSF values for Cqm1 without Ca^{2+} is shown in solid line and with bound Ca^{2+} is given in dash line.

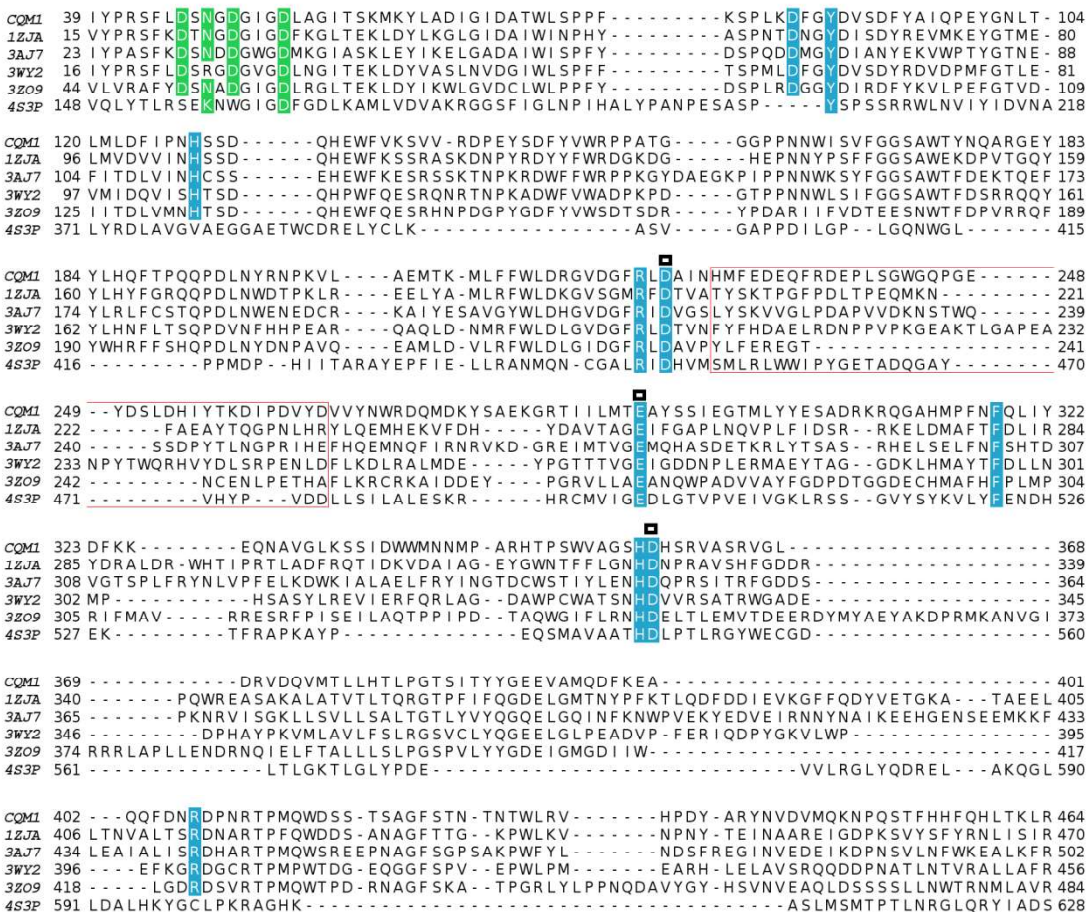


Figure 4.7. Alignment of glycoside hydrolases. Structures of sucrase (PDB Id, 1ZJA), isomaltase (PDB Id, 3AJ7), α-glucosidase (PDB Id, 3WY2) and trehalose synthase (PDB Id, 1ZO9) were identified by MODELLER to be used for construction of Cqm1 structural model. Alignment of these sequences along with the sequence of *E. coli* amylomaltase (PDB Id, 4S3P) was achieved using PROMALS3D [38]. Alignment is shown for rows with selected features. Individual protein sequences are numbered. Conserved residues of active-site scaffold are shaded blue, Ca^{2+} binding residues are shaded green, and loop-4 between β4 and α4 elements is marked as red box. Catalytic triad (Asp-224, Glu-290 and Asp-358) of Cqm1 is marked with black squares.

4.4.4. Interaction of BinA/B proteins with Cqm1

Interaction between Cqm1 and BinA/B proteins were evaluated in real-time with surface plasmon resonance. The Cqm1 (ligand) was immobilized on a CM5 Biacore chip and BinA/B proteins were used in mobile phase. The interaction of the two toxin components with receptor Cqm1 was confirmed by the increase in SPR signal as a function of concentration. The results showed increased mass density at the sensor surface due to BinB binding (Fig. 4.8A). The binding curves could be explained with model that corresponds to 1:1 interaction of Cqm1 and BinB with excellent fit between experimental data and the fitted curve (Chi^2, 0.0284). The association and dissociation rate constants (*k*a and *k*d) are $7.12x10^4$ M^{-1} s^{-1} and 0.0007 s^{-1}, respectively, that suggests strong binding with equilibrium dissociation constant (*K*D) value of 9.8 nM. In comparison, BinA does not show interaction with Cqm1 at lower concentrations. The interaction of Cqm1 and BinA are detected at higher concentration of BinA (Fig. 4.8B) with estimated *k*a and *k*d values of $1.49x10^4$ M^{-1} s^{-1} and 0.032 s^{-1} that results in *K*D value of 2.2 µM (1:1 interaction of Cqm1 and BinA with Chi^2 value of 0.237). The SPR analysis thus suggests specific and strong interaction of BinB with Cqm1, while interaction of BinA with Cqm1 may be non-specific and transient. Significantly retarded mobility of Cqm1 in the presence of BinB on 10% native-PAGE and independent migration of BinA and Cqm1 supports these results (Fig. 4.9).

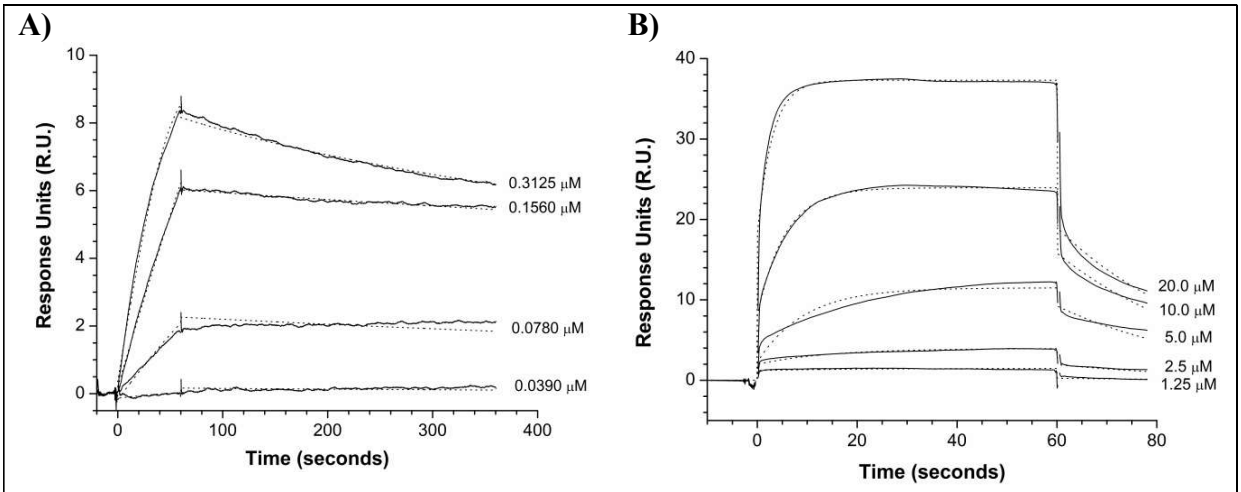


Figure 4.8. Interaction of Cqm1 with BinAB toxin components. Nearly 1240 response units of Cqm1 were immobilized on to CM5 Biacore chip that was subsequently washed thoroughly with buffer F, and different concentrations of the analytes (BinA or BinB) were injected over Cqm1 surface in independent experiments. The response was monitored for 60 sec before washing with buffer F. **A**) BinB as analyte. **B**) BinA as analyte. The black trace (solid line) at each concentration indicates experimental determination. The interactions could best be modeled by a 1:1 interaction model (dotted traces) in both the studies.

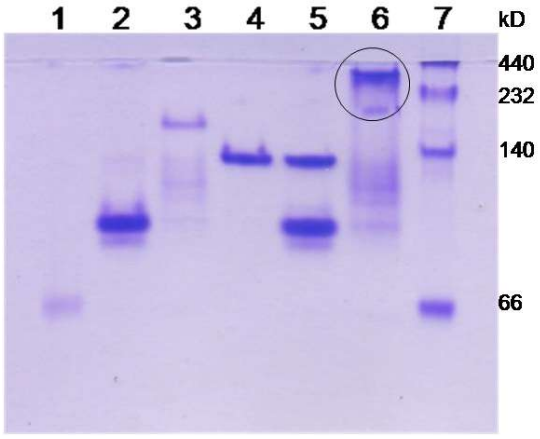


Figure 4.9. **Interaction of BinA and BinB proteins with receptor protein Cqm1; 10% native-PAGE analysis.** Lane 1, BSA as protein marker; Lane 2, BinA; Lane 3, BinB; Lane 4, Cqm1; Lane 5, BinA incubated with Cqm1; Lane 6, BinB incubated with Cqm1; Lane 7, native-PAGE standard protein molecular weight marker. The encircled protein band in lane 6 is expected to be stable complex of Cqm1 and BinB proteins. BinA however, does not co-migrate with Cqm1 (lane 5) under the conditions of experiments.

4.4.5. Enzymatic activity of Cqm1 protein

The Michaelis-Menten kinetics parameters towards α-pNPG were estimated from nonlinear regression fit to the velocity-substrate profile (Fig. 4.10A). The maximum velocity (*V*max), affinity constant (*K*m) and catalytic constant (*k*cat) for α-pNPG hydrolysis are estimated to be 1.81 μmol.min^{-1}.mg^{-1}, 0.44 mM and 1.9 s^{-1}, respectively. The resulting specific activity and catalytic efficiency (*k*cat/*K*m) of Cqm1 are calculated to be 3.2 Units /mg, and 4318 M^{-1}s^{-1}, respectively.

The optimum temperature for hydrolytic activity against α-pNPG is estimated to be about 37 ºC. The activity increases steadily from 70% at 25 ºC to the maximum at 37 ºC, but declines at higher temperatures above 50 ºC (Fig. 4.10B). The Cqm1 protein displays significant hydrolytic activity at pH values above 5 with the maximum activity at pH value of about 7.2. Better than 70% activity is observed within the pH range of 6.0 to 8.0. However, enzyme activity rapidly decreases at pH value above 8.0 and Cqm1 is completely inactive at alkaline pH value of 11.0 (Fig. 4.10C).

The enzymatic activity of Cqm1 against α-pNPG is not affected in the presence of BinA. BinA displays larvicidal activity [33, 88] and displays transient interaction with Cqm1. However, BinB, which binds with high affinity to Cqm1, reduces α-glucosidase activity by nearly 50% (Fig. 4.11C). Since BinA is known to be the toxic factor, it rules out the possibility discussed earlier [28] that larvicidal activity may be due to inhibition of receptor biochemical activity.

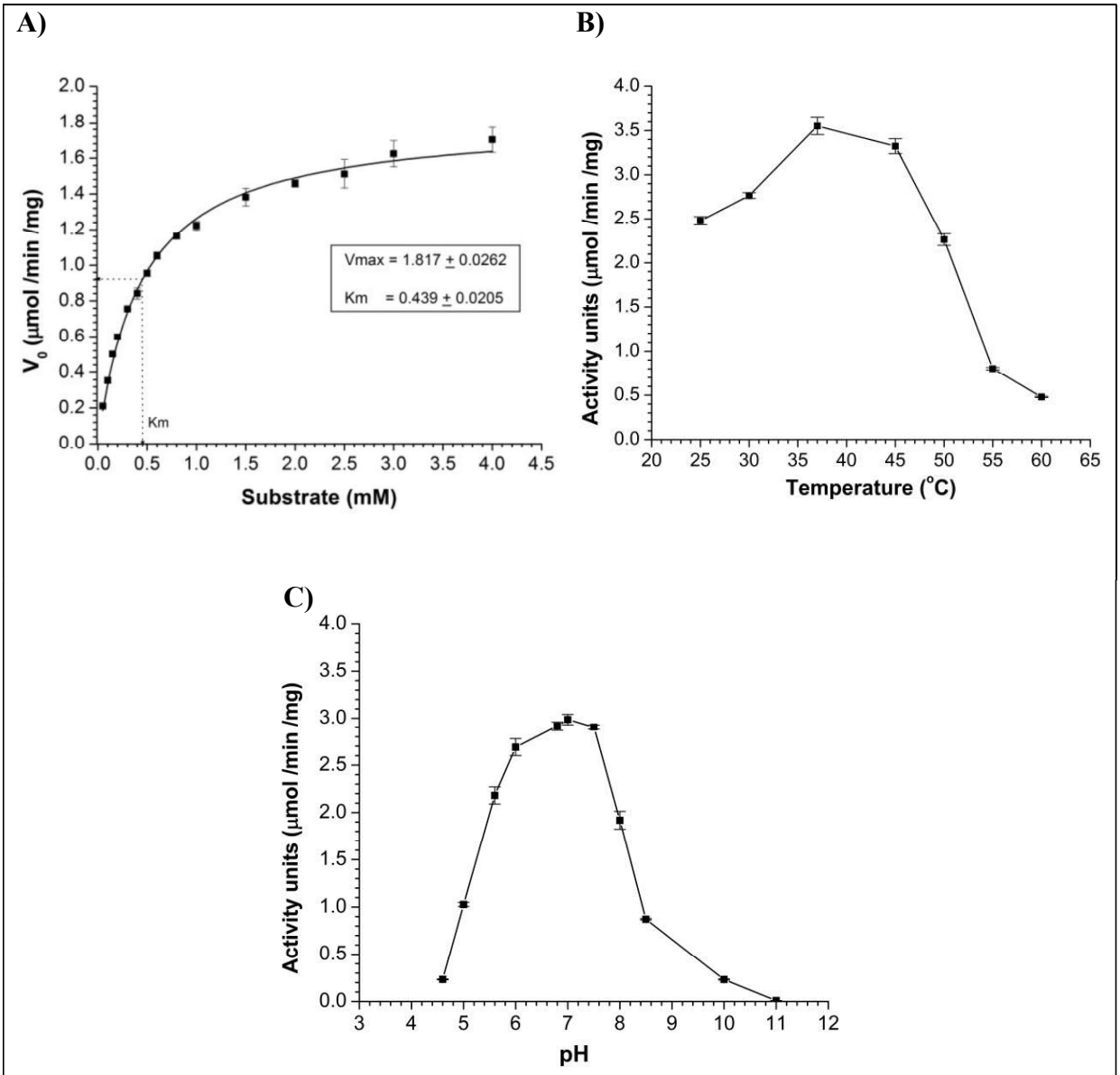


Figure 4.10. Kinetic analysis of α-glucosidase activity of Cqm1 protein. A) The Michaelis-Menten plot using Cqm1 (65 nM) at pH 6.8 and temperature 25 ºC is shown. The kinetic parameters were obtained by nonlinear regression fit to the Michaelis-Menten equation using Origin software. **B**) Plot depicting variation of enzymatic activity at pH 6.8 with varying temperature (25 to 60 ºC) using Cqm1 (65 nM) and substrate (1 mM). **C**) Plot displaying variation of enzymatic activity with varying pH (4.6 to 11) at 37 ºC using Cqm1 (65 nM) and substrate (1 mM). The vertical bars represent standard errors observed in three independent experiments. The substrate α-pNPG was used for all the experiments.

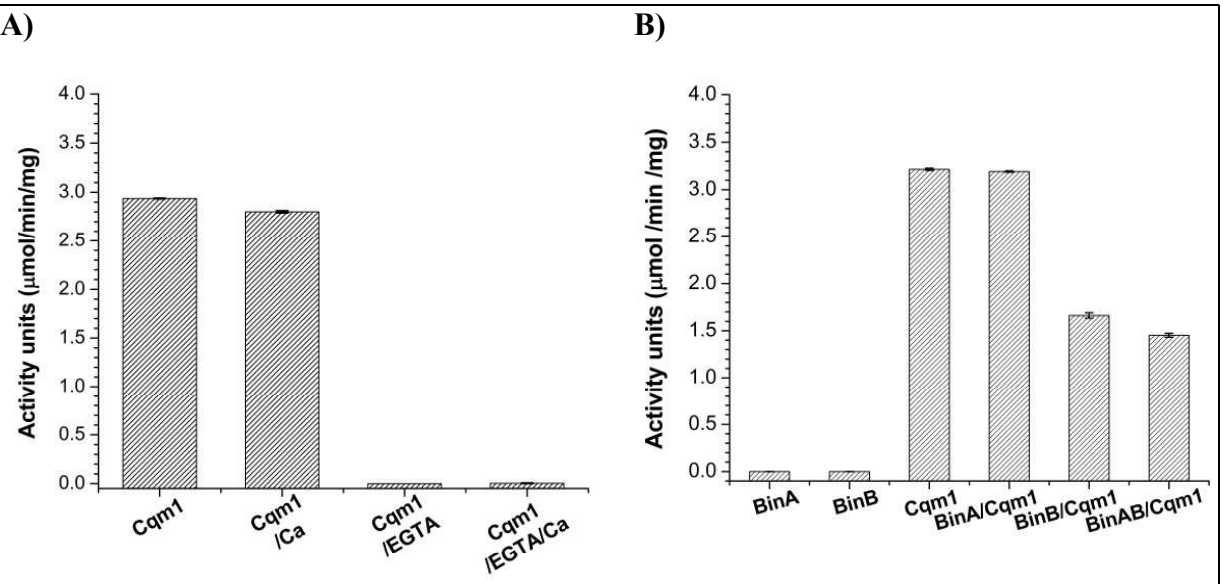


Figure 4.11. Inhibition of Cqm1 α-glucosidase activity. A) Enhancement of Cqm1 (65 nM) enzymatic activity was not observed with the addition of Ca^{2+} (Cqm1/Ca). However, EGTA (5 mM) totally inhibited α-glucosidase activity (Cqm1/EGTA). Addition of Ca^{2+} (10 mM) to the dialysed fraction of Cqm1/EGTA did not rescue the activity (Cqm1/EGTA/Ca). **B**) α-glucosidase activity of Cqm1 was probed in the presence of BinA (BinA/Cqm1), BinB (BinB/Cqm1), and equimolar mixture of BinA/BinB (BinAB/Cqm1) proteins. BinA and BinB individually were also used in control experiments without Cqm1. α-glucosidase activity was examined at 37 ºC and pH 7.5 using α-pNPG (1 mM) as a substrate in both the experiments. The error bars show standard errors from three independent experiments.

Recombinant Cqm1 enzyme was incubated with different disaccharides and glycogen at 37 °C for different time intervals, and reaction products were analysed by TLC to probe substrate specificity. Cqm1 fails to act on trehalose and isomaltose (Fig. 4.12A). Instead, Cqm1 shows efficient transglycosylation activity with maltose and maltotriose, as is evident from the accumulation of maltose-oligomers (Fig. 4.12B). Cqm1 also displays weak hydrolytic activity against sucrose, and higher molecular mass oligomers of sucrose are not observed even after 24 hr incubation. However, higher accumulation of fructose is observed when glucose was added to Cqm1/sucrose reaction mixture (Fig. 4.13A). Low molecular weight saccharides, such as glucose, are observed in reaction with glycogen after about 5 hr (Fig. 4.13B). The data taken together suggests that Cqm1 is competent to cleave α-1,4- and α-1,2-glycosidic bonds and synthesizes new α-1,4-linkages with shorter maltodextrins to form maltose-oligomers.

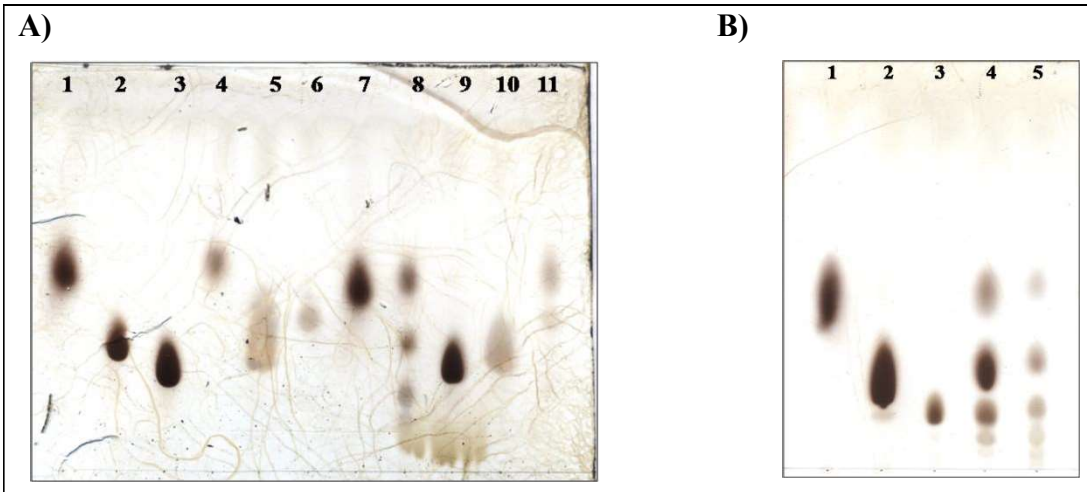


Figure 4.12. Thin layer chromatography analysis. A) Lane 1, glucose; Lane 2; maltose; Lane 3, isomaltose; Lane 4, fructose; Lane 5, trehalose; Lane 6, sucrose. Glucose, maltose, isomaltose, trehalose and sucrose were also incubated individually with Cqm1 (5 μM) for overnight at 37 ºC and reaction mixtures were resolved on TLC in lanes 7, 8, 9, 10 and 11, respectively. **B**) Lane 1, glucose; Lane 2; maltose; Lane 3, maltotriose. Maltose and maltotriose were individually incubated with Cqm1 (5 μM) at 37 ºC for 2 hr and mixtures were resolved on TLC in lanes 4 and 5, respectively.

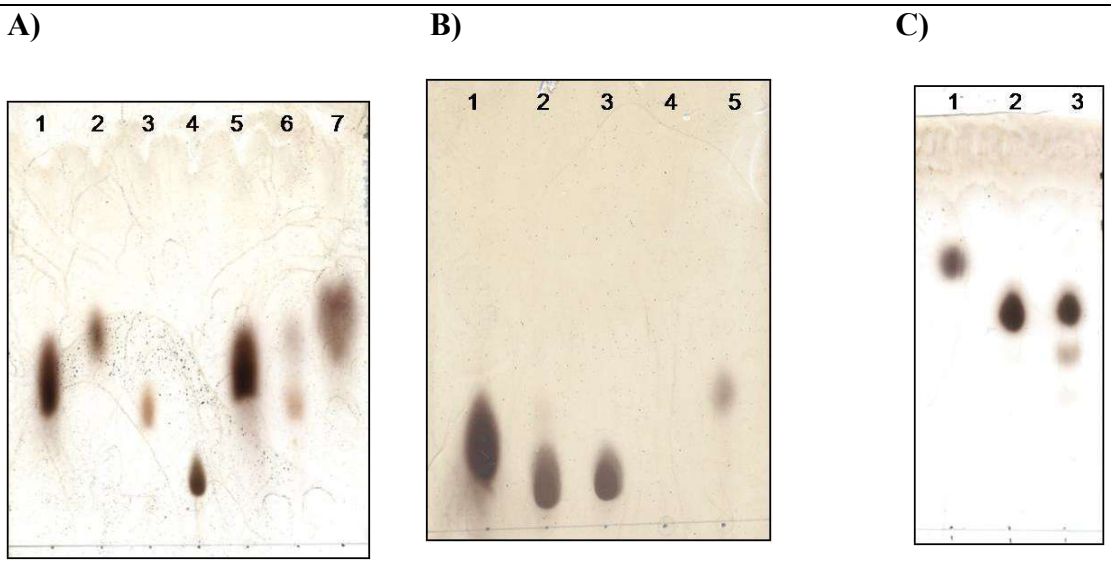


Figure 4.13. Thin layer chromatography analysis. A) Lane 1, glucose; lane 2; fructose; lane 3, sucrose; lane 4, maltotriose. Glucose and sucrose were incubated with Cqm1 at 37 °C for 5 hr in independent reaction and were resolved on TLC in lanes 5 and 6, respectively. Sucrose hydrolysis by Cqm1 in the presence of glucose was also carried out at 37 °C for 5 hr and the reaction products were resolved in lane 7. Rate of sucrose hydrolysis was enhanced when glucose was included in the reaction mix, releasing more fructose. **B**) Lane 1, glucose; lane 2; maltose; lane 3, isomaltose; lane 4, glycogen. Glycogen was incubated with Cqm1 at 37°C for 5 hr and was resolved on TLC in lane 5. A low molecular mass saccharide is observed due to hydrolysis of glycogen. Glycogen itself is not stained by the used procedure. **C**) Lane 1, glucose; lane 2; maltose. Maltose was incubated with Cqm1 at 37 °C for 20 min and was resolved on TLC in lane 3. High molecular mass dextrins, and not glucose, are observed in short incubation time intervals.

4.5. DISCUSSION

The active BinAB toxin from *L. sphaericus* binds to surface receptor protein (Cqm1), which is bound to the midgut epithelial cells via a glycosyl-phosphatidylinositol (GPI) anchor. The Cqm1 protein has been classified as a member of glycoside hydrolase family 13 (GH 13) of the CAZy database. It shares 74% identity with Aam1 receptor protein from *Aedes* species (Fig. 1.8). The *cqm*1 gene (1840 bp) from BinAB susceptible *Culex quinquefasciatus* mosquito colony contains a 1740 bp coding DNA sequence (CDS), which encodes for 580 amino acid long Cqm1 protein. Due to difficulties encountered in cDNA preparation, *cqm*1 gene with CDS of 1740 bp was synthesized chemically with codons optimized for expression in *E. coli* host. Codon optimization was required as the *cqm*1 gene had nearly 12% rare codons for *E. coli*. The disordered regions were removed and the active core of *cqm*1 was PCR amplified and cloned into pNIC28-Bsa4 vector by LIC method, and the encoded protein was purified using *E. coli* expression system (Fig. 4.2). CD spectroscopy revealed protein to be comprised of mixed α/β structure (α, ~8%; β, ~40%) (Fig. 4.4A). Intrinsic protein fluorescence measurement in the absence and presence of 8M urea suggest that the purified recombinant protein has folded tertiary structure. The identity of the purified protein was confirmed by peptide mass fingerprinting analysis to be Cqm1 (Fig. 4.3).

The size exclusion chromatography, dynamic light scattering and Native-PAGE suggests that Cqm1 exists in solution as a stable homodimer, and both hydrophobic and ionic residues may contribute towards dimeric interface. The stable oligomer of Cqm1 may be required for its apical localization in lipid rafts, as observed for other GPI-anchored proteins [111]. It has been suggested earlier that oligomer of Cpm1 (Cqm1 ortholog from *Culex pipiens*) promotes the opening of ionic pores by interacting with BinAB [112].

The SPR analysis suggests specific and strong interaction of BinB with Cqm1 with an estimated *K*D value of 9.8 nM. This matches well with the earlier reported *K*D of 7—20 nM

for BinB binding to larvae midgut preparations [56, 113]. In comparison, BinA and Cqm1 interaction were detected only at higher concentration of BinA (Fig. 4.8B) with estimated *K*D value of 2.2 μM. This also agrees well with the IC_{50} of about 2 μM observed for BinA with brush border membrane fractions [113]. It also suggests that interaction of BinA with Cqm1 may be non-specific and transient. Significantly retarded mobility of Cqm1 in the presence of BinB on 10% native-PAGE and independent migration of BinA and Cqm1 supports these results (Fig. 4.9).

The protein displayed hydrolytic activity against 4-nitrophenyl α-D-glucopyranoside (α-pNPG) and not against 4-nitrophenyl β-D-glucopyranoside. The low *k*cat value of 1.9 s^{-1} is in agreement with the values generally observed for maltases [114]. The observed catalytic efficiency of Cqm1 (4318 $M^{-1}s^{-1}$) compares favorably with the homologous protein of *Aedes aegypti*, but is much higher than the value of about 40 $M^{-1}s^{-1}$ reported for the *Culex* protein expressed along with un-structured signal peptide residues [97].

The optimum temperature and pH for hydrolytic activity against α-pNPG were observed to be about 37 ºC and pH 7.2, respectively (Fig. 4.10B & C). The observed pH optimum relates well to the physiological conditions found in the posterior half of larval midgut (pH of 7—8 in *Culex*), where fluorescently labelled BinAB were observed to co-localize [115, 116].

The Ca^{2+} binding residues reside in the loop-1 connecting β1 and α1 elements of the $(\beta/\alpha)_8$-barrel in the structural model and are strictly conserved in GH13_17 subfamily. We observe that EGTA inhibits α-glucosidase activity of Cqm1 protein completely (Fig. 4.11B) and Ca^{2+} also enhanced thermal stability of the tertiary structure, as adjudged from thermofluor-shift analysis (Fig. 4.4C). During the 40 ns molecular dynamics (MD) simulation, loop defining Ca^{2+} binding site is found to maintain its conformation when divalent cation is not bound to it. The peptide fragments showing diversity in the two homology models also

show higher flexibility in MD simulations. Notably, loop-4 between β4 and α4 elements in the catalytic domain of Cqm1 and consisting of residues 236–253 shows high RMSF of about 6Å (Fig. 3B). However, despite higher flexibility it does not shield the active-site. Compared to *Halomonas* sp. H11 α-glucosidase, Cqm1 sequence harbours 6 residue deletion in the loop-4 (Fig. 4.6B) [117].

Further, toxic BinA component did not inhibit α-glucosidase activity of Cqm1 against α-pNPG, BinA displays larvicidal activity [33, 88] and displays transient interaction with Cqm1. However, BinB which binds with high affinity to Cqm1 reduced the activity by nearly 50%. Since BinB alone does not exert toxic effect, it rules out the possibility discussed earlier [28] that larvicidal activity may be due to inhibition of receptor biochemical activity.

Notably, thin layer chromatography experiments established Cqm1 as α-glucosidase competent to cleave α-1,4- and α-1,2-glycosidic bonds and synthesizes new α-1,4-linkages with shorter maltodextrins to form maltose-oligomers. Conservation of Ala-225 and Phe-318 residues is consistent with the hydrolytic activity against α-1,4-linkages. The α-1,6-glucosidases have Val at position equivalent to Ala-225 of Cqm1. The higher activity against holoside substrates (maltose and maltotriose), and poor activity against sucrose and α-aryl glucoside (α-pNPG), suggest that Cqm1 α-glucosidase should be characterized as type II or type III-glucosidase. Type II and III enzymes are usually catalogued in GH31 family of CAZy database, however.

The formation of glucosyl-enzyme intermediate resulting from the hydrolysis of glucose-oligomers precedes transglycosylation step, which seems to be preferred over hydrolysis of glucosyl-enzyme intermediate. Higher accumulation of maltose-oligomers is observed in reaction mixture with shorter incubation time of 2 hr (Fig. 4.12B), compared to longer incubation period of 24 hr (Fig. 4.12A). Also, maltose-oligomers, and not glucose, are observed in 20 min incubation reaction of Cqm1 with maltose (Fig. 4.13C). This may be due

to depletion of maltose-oligomer substrates of Cqm1 over longer time spans. Hydrolysis of glycogen and transglycosylation of maltose-oligomers are consistent with open and accessible active site, consisting of a catalytic triad (Asp-224, Glu-290 and Asp-358), for binding of large substrates. The short loop-4 between β4 and α4 elements in the catalytic domain permits binding of maltose-oligomers and glycogen in the active site. In comparison, longer loop-4 in *Halomonas* sp. H11 α-glucosidase sterically inhibits binding of maltose-oligomers in the active site during transglycosylation step [117].

Cqm1 displays amylomaltase activity, which catalyses transglycosylation by successive transfer of glucose moiety of maltose onto the malto-oligosaccharides, and sucrase activity, which hydrolyses sucrose to glucose and fructose. Similar activities have also been reported for α-glucosidase from *Acyrthosiphon pisum* and *Apiscerana japonica* which also belong to GH13_17 subfamily [118, 119]. The appearance of low-molecular weight saccharides in Cqm1/glycogen reaction mixture also suggests that the protein may be involved in hydrolysis of high-molecular-mass maltose-oligomers. Thus, Cqm1 or its orthologs in *Aedes* and *Anopheles* species can be expected to participate in carbohydrate metabolism in mosquitoes to meet the high-energy demand for reproduction and for harbouring disease-causing vectors. Also, α-glucosidases seem to be essential for the partitioning of carbohydrates from the diet into carbon nutrition and osmo-regulation in insects [120, 121]. Further studies can reveal if Cqm1 or its orthologs in *Aedes* and *Anopheles* can be targeted for mosquito control programs, as *Culex* mosquitoes, incompetent to present Cqm1 on cell surface are known to be viable.

4.6. CONCLUSIONS

- Interaction of BinAB with mosquito gut receptor protein (Cqm1) is essential for *L. sphaericus* larvicidal activity.
- Recombinant Cqm1 protein exists as a stable dimer which supports its apical localization in lipid rafts.
- Cqm1 unexpectedly displays amylomaltase activity and is also competent to hydrolyse sucrose and glycogen.
- Ca(II) provides thermal stability to the protein.
- BinB toxin component reduces catalytic activity of Cqm1 protein.
- Cqm1 can be expected to participate in carbohydrate metabolism in mosquitoes, which makes it a valuable target for world-wide mosquito control programmes.

Chapter 5

Mosquito-larvicidal BinA toxin displays affinity for glycoconjugates: biophysical basis of BinA-mediated cytotoxicity

5.1. ABSTRACT

Lysinibacillus sphaericus parasporal BinAB toxin displays mosquito larvicidal activity against *Culex* and *Anopheles,* but *Aedes* species are refractory to it. Recently reported crystal structure of BinAB revealed the presence of N-terminal carbohydrate binding domain (β-trefoil fold) in BinA. Earlier competitive studies reported inhibitory effects of sugars on BinA toxicity. Hemagglutination and hemolytic activities were not observed for BinA in the present studies. We attempted to characterize carbohydrate specificity of BinA by high-throughput approaches using extrinsic fluorescence and thermofluor shift assay. A total of 34 saccharides (mono-, di- and polysaccharides, and glycoproteins) known to bind to most of the lectins were used for initial high-throughput screening. The promising glycans were identified based on significant change (>10%) in the fluorescence intensity. High-throughput thermofluor shift assay, however, failed to detect ligand binding. Further, binding of simple sugars could not be confirmed by Isothermal Titration Calorimetry. Surface plasmon resonance (SPR) revealed differential binding of BinA with glycoproteins (fetuin, asialofetuin and thyroglobulin) and affinity for simple sugars, L-fucose and L-arabinose. In the carbohydrate competition assay, arabinose, fucose and fetuin inhibited BinA toxicity towards *Culex* larvae. This study for the first time demonstrates that BinA is competent to bind diverse and structurally different glycosylated proteins that may be linked to its cytotoxic activity, as protein N-glycosylation has been reported to be essential for development and survival of insect larvae. Further, BinA did not show specific interaction with the *Culex* receptor protein (purified without GPI anchor). We hypothesize that recognition of the extracellular receptor (Cqm1 and Agm1 proteins of *Culex* and *Anopheles* mosquitoes, respectively) may be mediated by binding with glycan core of their GPI anchors. The strong binding of BinA to glycosylated Aam1 receptor of *Aedes* may block its intracellular translocation resulting in non-toxic accumulation of BinA in the posterior mid gut region of mosquito larvae.

5.2. INTRODUCTION

Mosquito borne diseases (like, Zika, Chikungunya, Malaria etc.) have been recognized as world health problem since long time. The *Lysinibacillus sphaericus* binary (BinAB) toxin is the naturally occurring paracrystalline larvicide that has been found to be effective in controlling mosquito population of *Culex* and *Anopheles* species world-wide [122]. These crystals are composed of homologous molecules of protoxins; BinA and BinB polypeptides. The mature BinB (~42 kDa) binds to a maltase1 receptor on the midgut cell membrane, and mature BinA (~39 kDa) causes cell death on internalization [31, 113, 123]. The receptor (Cqm1 protein of *Culex,* Agm1 protein of *Anopheles* and Aam1 protein of *Aedes*) is attached to cells via a glycosyl-phosphatidylinositol (GPI) anchor.

BinA and BinB (BinA/B) proteins act synergistically and a heterotetramer (BinA2.BinB2) is expected to be formed on the receptor, prior to cellular internalization. The pre-formed covalent complex of BinA/B proteins, and PEGylated BinA alone, without BinB, also show high larvicidal activity [32, 33]. The recent crystallographic analysis reveals that BinA and BinB proteins are composed of N-terminal β-trefoil (carbohydrate binding domain; CBD) and pore-forming aerolysin domains (Fig. 5.1) [40]. The heteromeric complex of BinA/B translocates inside the cell through pore formation and BinA is thought to be responsible for cytotoxicity. However, the cause of larval cell death is not known. Another important issue being discussed, since the availability of BinAB crystal structure, is to broaden specificity of BinAB towards *Aedes aegypti*, which displays poor or no susceptibility towards BinAB toxin and is carrier of dreaded Dengue and Zika viral diseases.

Earlier some glycans were observed to reduce BinA toxicity against *Culex* cells [27]. Glycans or glycoconjugates are being recognized more recently for their major metabolic, structural and biological roles [124, 125]. Carbohydrate-protein interactions are generally weak with the affinity constants varying from micromolar to millimolar range. The available

biophysical methods (like, fluorescence, thermofluor shift assay, surface plasmon resonance, isothermal titration calorimetry) are sensitive to detect and characterize weak-to-moderate binding affinities and each one of the currently available technologies offer a characteristic type of biophysical information about the ligand-binding event [126, 127].

The present study reports biophysical analysis of BinA specificity towards simple sugars and glycosylated proteins. The initial high-throughput screening was based on fluorescence spectroscopy and thermofluor shift assay. Promising ligands were analysed for their binding affinities through isothermal titration calorimetry (ITC), surface plasmon resonance (SPR) and carbohydrate inhibitory assay. Based on these results we propose that BinA cytotoxicity may be linked to its high affinity for glycosylated proteins inside the larval gut cells. The suggestion also rationalizes different responses of *Culex* and *Aedes* mosquito species towards BinAB toxin.

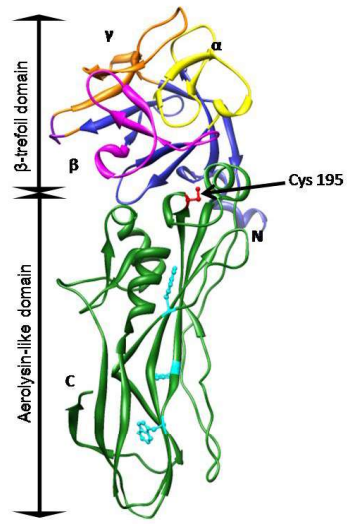


Figure 5.1. Ribbon model of BinA protein structure (PDB ID: 5FOY) showing the location of a free thiol group (marked with an arrow), Cys195 (red; ball and stick), located just beneath the β-trefoil (carbohydrate binding) domain. The three carbohydrate binding modules (α, β, γ) of CBD are shown in yellow, magenta and orange, respectively. The three Trp residues (cyan; ball-and-stick) in BinA sequence are localized in aerolysin-like domain (green ribbon). Ribbon model was drawn by using Chimera suite [12].

5.3. MATERIALS & METHODS

5.3.1. Materials

Thrombin (from bovine plasma), Phenylmethylsulfonylfluoride (PMSF), Trypsin (proteomics grade, from porcine pancreas), Bradford reagent, D-fructose, L-sorbose, *L-fucose,* L-rhamnose, N-acetyl-D-mannosamine, methyl-α-D-glucopyranoside, methyl-β-D-glucopyranoside, methyl- α-D-mannopyranoside, methyl-α-D-galactopyranoside, methyl-β-D-galactopyranoside, N-acetyl-D-lactosamine, thyroglobulin (bovine), asialofetuin (bovine) and fetuin (bovine), were procured from Sigma. D-arabinose, L-arabinose, D-ribose, D-galactose, Dextrose, D-mannose, N-acetyl-D-galactosamine, N-acetyl-D-glucosamine, adonitol, myo-inositol, N-acetyl-neuraminic acid, β-gentobiose, D-lactose, D-maltose, D-trehalose, D-melezitose, D-melibiose, D-raffinose, inulin, glycogen and β-phenyl-D-glucopyranoside were procured from SRL. 6-acryloyl-(dimethyl amino) naphthalene (acrylodan) was from Molecular Probes. 12 kDa MW cutoff dialysis tubing (cellulose membrane) was from Spectra/Por. All other chemicals were of analytical grade.

5.3.2. Expression and purification of BinA, BinB and Cqm1 proteins

Recombinant BinA, BinB and Cqm1 proteins were purified as described earlier (Chapter 4) [32]. Briefly, recombinant BinA (active form; residues 15-366), BinB (residues 1-448) and Cqm1 (residues 23-560) were expressed individually in BL21 (DE3), pLysS and BL21 Star (DE3) expression host cells, respectively, as fusion proteins with N-terminal poly-His tag and a thrombin (for BinA and BinB) and TEV (for Cqm1) protease cleavage site between poly-His tag and target protein. The recombinant proteins were purified from soluble cell lysate using Immobilized Metal Ion Affinity Chromatography (IMAC) and Anion Exchange Chromatography (AEC) methods. The poly-His tag was removed from BinA and Cqm1 by thrombin and TEV treatment, respectively. The active BinA and Cqm1 proteins were further purified by Size Exclusion Chromatography (SEC) using Superdex 75 and 200

columns, respectively, in Buffer A (50 mM Tris-HCl pH 8.0 and 100 mM NaCl). Purified recombinant BinB protein was treated with trypsin (trypsin:BinB::1:1000, w/w) at 20 ºC overnight. The reaction was stopped with 10 mM PMSF. Trypsin activated BinB (residues 17-448) was further purified by AEC. The purified proteins were adjudged on 12% SDS-PAGE, and were dialyzed to buffer B (50 mM Tris-HCl pH 8.0 and 25 mM NaCl) and stored at 4 ºC. Protein concentration was determined using Bradford assay [78].

5.3.3. Hemagglutination and Hemolytic assay

2 mL of rabbit blood was mixed with phosphate buffered saline (1X PBS, pH 7.2) and centrifuged at 2000 rpm for 10 min. The red blood cells (RBCs) were washed thrice and suspended in the 1X PBS buffer, pH 7.2. The isolated RBCs were trypsinized at 37 ºC for 1 hr followed by washing the cells twice with PBS buffer. The cells were resuspended in 1X PBS buffer, pH 7.2 to get a final suspension of ~10^8 cells/mL.

The hemagglutination assay was carried out in a 96-well microtiter plate following a two-fold serial dilution method [128]. 50 µL of the BinA protein, (0.4 mg/mL to 0.00625 mg/mL) and equimolar mixture of recombinant BinA/B proteins (0.4 mg/mL to 0.00625 mg/mL), were incubated with an equal volume of 2% suspension of rabbit erythrocytes (~ 4 × 10^8 cells/mL) in independent reactions at 37 ºC for 1 hr. Following incubation, the reaction was stabilized by keeping the plate at 4 ºC for additional 1 hr. Agglutination was checked visually. Positive hemagglutination was recorded if the cells formed a fuzzy mat, while formation of button at the bottom of the well indicated absence of agglutination. Lectin from *Phaseolus vulgaris* was used as a positive control and bovine serum albumin (BSA) was used as a negative control for hemagglutination assay. The hemagglutination assay was carried out both in the absence and the presence of Ca^{2+} (1 mM) to study its effect on the hemagglutination activity of BinA and BinA/B proteins.

Hemolytic activity was measured spectrophotometrically. Rabbit red blood cells (100 µL; 5X 10^7 cells/mL) in 1X PBS buffer (pH 7.2) were mixed with serial two-fold dilutions of equal volume of BinA and BinA/B proteins in independent reactions and incubated for 4 hr at 37 ºC. After incubation, the amount of hemoglobin released into the medium was measured spectroscopically at 405, 540 and 630 nm with a UV-VIS spectrophotometer (V-630, JASCO) in 1 cm cuvette. The erythrocyte hemolysis is monitored generally at these three wavelengths. The percent red blood cell disruption was then quantified relative to positive control samples lysed with 1% triton-X 100. The effect of Ca^{2+} on the hemolytic activity of BinA and BinA/B proteins was also studied in an independent experiment. Percentage hemolysis was calculated using Eq. 1, where A_{max}, A_{min} and A_t represent the absorbance values for completely hemolyzed RBCs by triton-X 100, untreated RBCs, and each test reaction after 4 hr of incubation, respectively.

$$\% \text{ Hemolysis} = \left(\frac{A_t - A_{min}}{A_{max} - A_{min}} \right) * 100 \qquad \text{.....(1}$$

5.3.4. BinA labeling with acrylodan

Acrylodan is a hydrophobic fluorophore which is extremely sensitive to its local environment. It binds only to free thiol groups in the protein [129]. The BinA Cys195 was observed in the reduced form in the crystal structure. The accessibility of free sulfhydryl group of Cys195 was determined by Ellman reagent, 5, 5'-dithio-bis-(2-nitrobenzoic acid) [DTNB]. The DTNB assay was performed by adding 950 µL of 0.1 mM DTNB (diluted in buffer B from 10 mM stock in DMSO) to 50 µL of different dilutions of BinA protein. Samples were mixed by brief vortexing followed by incubation for 30 min at room temperature. Following incubation, the absorbance of the samples was measured at a wavelength of 412 nm using UV-Vis spectrophotometer (JASCO). The reaction results in the formation of 2-nitro-thiobenezoic acid (TNB), a measurable yellow colored product, with molar extinction coefficient of 14.15 $mM^{-1}cm^{-1}$ at 412 nm. The concentration of the free

sulfhydryls was calculated from the molar extinction coefficient of TNB using Beer-Lambert law.

The active BinA protein was labeled with acrylodan (Ac-BinA) as described earlier [129]. BinA protein was dialyzed overnight at 4 ºC against buffer C (20 mM sodium phosphate pH 7.2, 20 mM NaCl). It was then mixed with acrylodan (stock solution made in acetonitrile) in 1:2 molar ratio with continuous stirring in the dark for about 6 hr at 4 ºC. Excess of acrylodan was dialyzed out against buffer C over a period of 48 hr with an exchange of buffer at regular intervals. Labeled BinA protein was further purified through size exclusion chromatography using Superdex 75 column in buffer C to ensure complete removal of excess acrylodan. The purity of protein sample was adjudged from 12% SDS-PAGE gel. Stoichiometry of acrylodan labeling of BinA protein was estimated by comparing acrylodan concentration (by absorption at 372 nm, molar extinction coefficient 16,400 M^{-1} cm^{-1}) and protein concentration (by absorption at 280 nm, molar extinction coefficient 49,280 M^{-1} cm^{-1}) The labeling efficiency (%) was estimated by Eq. 2.

$$\%\ \text{Labelling efficiency} = \left(\frac{\text{concentration of Acrylodan}}{\text{concentration of protein}}\right) * 100 \qquad \text{…. (2}$$

The fluorophore labeling of BinA protein was further confirmed through spectrofluorometry by measuring the emission spectra of unlabeled BinA and Ac-BinA proteins following excitation at 391 nm.

5.3.5. High throughput screening of carbohydrate binding by fluorescence

The protein-glycan interactions were probed using intrinsic fluorescence from Trp residues (three Trp residues in BinA sequence) and extrinsic fluorescence of acrylodan covalently attached to BinA. 34 different sugars and their derivatives comprising of monosaccharides, oligosaccharides, polysaccharides and glycoproteins (Table 5.1) known for binding to lectins and lectin-like proteins were used. BinA (20 μM) or Ac-BinA (20 μM) in

buffer D (20 mM Tris pH 8.0, 20 mM NaCl) was incubated with different sugars (1:10 molar ratios) overnight at room temperature. The emission spectra with and without different sugars were recorded at an excitation wavelength of 280 nm for BinA and at 391 nm for Ac-BinA. For BinA emission spectra were recorded for each protein-sugar mixture in a 1 cm path length cell using JASCO spectrofluorometer (FP-8500) at room temperature, while Ac-BinA emission spectra were recorded in a high-throughput mode using 96 well flat-bottom brown-colored plate using INFINITE M200 (Teccan) multimode plate reader. The effect of Ca^{2+} on protein-ligand interaction was also investigated in the presence of 1 mM $CaCl_2$. The fluorescence spectra of the buffer and individual sugar were subtracted from their respective test spectra. The inner filter correction, owing to absorbance of the fluorescence intensity due to the ligands, was also employed to correct the observed fluorescence using Eq 3, where F_{corr}, Fluorescence corrected; F_{obs}, Fluorescence observed; OD_{ex}, Absorbance at excitation wavelength; OD_{em}, absorbance at emission wavelength. For intrinsic fluorescence experiments, excitation and emission wavelengths were 280 nm and 324 nm, respectively. For experiments with Ac-BinA, excitation and emission wavelengths were 391 nm and 492 nm, respectively.

$$F_{corr} = F_{obs} \text{ X antilog } \left(\frac{OD_{ex} + OD_{em}}{2}\right) \qquad \ldots\ldots(3$$

5.3.6. High throughput screening of carbohydrate binding by thermofluor shift assay

Thermofluor shift assay is a rapid and inexpensive screening method to identify low molecular weight ligands that bind and stabilize the protein [81]. The melting curve measurements, to assess the thermal stability of BinA protein in the absence and presence of simple sugars, were carried out as described by Niesen et al [81]. Each reaction was carried out in duplicate in a 96 well plate with 25 µL reaction volume in buffer C and consisted of recombinant BinA protein (2 µM) with different sugars (20 µM) and freshly diluted SYPRO

orange dye (1:1000, v/v). The plate was sealed and data was recorded on CFX96^{TM} Real time system (BioRad) in FRET mode. The plate was heated from 20 to 95 ºC at ramp rate of 1.0 ºC/ min. Experimental data were processed using software available with CFX96^{TM} Real time system and melting temperatures (T_m) were determined from the first derivative of each of the thermofluor-shift melting curve.

5.3.7. Interaction of simple sugars with BinA

Isothermal titration calorimetry was carried out using a MicroCal iTC200 System (Malvern, UK) at 25 °C in buffer C to investigate the binding of D-arabinose, L-fucose, D-trehalose and N-acetyl-D-lactosamine to *L. sphaericus* BinA protein. For these experiments, the syringe was filled with sugars and the titrations were performed by 20 injections, with a first injection of 0.4 µL followed by 2 µL each thereafter, into the (0.2 mL) sample cell containing BinA protein (40 µM) with constant stirring at 1000 rpm to ensure equilibrium. Data were analyzed using the MicroCal Origin software. The titrations against the protein were performed with the standard ligand concentration (1 mM) and also with high ligand concentration (20 mM).

The interactions between BinA protein and the promising ligands were also characterized by FTIR spectroscopy. FTIR spectra of BinA (1.5 mg/mL) in buffer E (50 mM potassium phosphate pH 8.0 and 50 mM NaCl), and with simple sugars D-arabinose, D-trehalose and L-fucose, were recorded at 4 cm^{-1} resolution on a Vertex 70 FTIR spectrometer (Bruker, Germany) equipped with BioATR system and a HgCdTe (MCT) detector. The protein and individual sugars were mixed, in increasing molar ratios (from 1:1 to 1:1000), prior to experiment and then analyzed through FTIR at room temperature. The FTIR spectra with 120 scans were recorded in the range of 900–4000 cm^{-1}. The spectra analysis was done using the OPUS version 6.5 software provided with the instrument. Difference spectra

(absorbance spectra of protein + ligand minus absorbance of protein) were determined to show the absorbance change upon ligand addition.

Surface plasmon resonance studies were carried out using Biacore T200 SPR system at 25 ^{o}C to investigate binding of D-arabinose, L-fucose, D-trehalose and N-acetyl-D-lactosamine to BinA protein. For each experiment, nearly 1014 response units of BinA (ligand) were immobilized onto the CM5 Biacore sensor chip via amine coupling followed by extensive washing with buffer E. Different concentrations (2.5, 5, 20, 40 and 80 μM) of each sugar (analytes) in buffer E were injected onto the BinA-bound sensor chip at 30 μL/min flow-rate in independent experiments. The analyte was allowed to interact with the immobilized protein for 60 sec before washing off with buffer E. The data were analyzed by kinetic and steady state affinity analysis using Biacore evaluation software (V3.1) provided with the instrument. Kinetic analysis for ligand binding was done using the 1:1 and two state models of interaction. Ligand binding was adjudged by affinity analysis from the curve fitted to a plot of Req (Response Units at equilibrium) against ligand concentration to a model representing equilibrium 1:1 binding. The Req was estimated for each concentration at 15 $\pm$ 5 sec (after the start of injection) time-point of the response curve.

5.3.8. Determining the binding affinities of glycoproteins with BinA by SPR

BinA-glycoproteins interactions were probed using Autolab Esprit SPR system with bare gold sensor chip. Nearly 400 response units of BinA (ligand) was immobilized onto the Bare gold sensor chip followed by extensive washing with buffer C. The concentration of BinA and stock solution of glycoproteins were estimated by Bradford assay. Different concentrations of the three glycoproteins (analytes) namely, fetuin (2.5, 5, 10, 20 and 40 μM), asialofetuin (3.75, 7.5, 15, 30 and 60 μM) and thyroglobulin (1.37, 2.75, 5.5, 11 and 22 μM), were injected onto the BinA-bound sensor chip in independent experiments. The analytes (mobile phase) were allowed to interact with the immobilized BinA for 300 s before washing

off with buffer C. The association kinetics analysis was done using Autolab kinetic evaluation software (V5.4) provided with the instrument. The interaction of analytes alone with bare gold chip was also monitored as a control. The association and dissociation rate constants were estimated from the fit against experimental SPR data.

5.3.9. BinA interaction with the receptor and BinB proteins

Size exclusion column chromatography was carried out to probe the interaction between BinA and Cqm1 (receptor) proteins. Active BinA and Cqm1 (purified without GPI anchor residues) proteins were mixed in 2:1 stoichiometry ratio. Following incubation for 1hr at 4 ºC, the mixture was analyzed with Superdex-200 SEC column, pre-equilibrated with buffer A. The elution peaks were adjudged by comparing with the elution profiles of both BinA and Cqm1 proteins individually and by 12% SDS-PAGE.

Chemical crosslinking of BinA and Cqm1 proteins was attempted using glutaraldehyde. Prior to reaction, the buffer of the constituent proteins was exchanged to buffer E. The reaction mixtures with 100 µg of each, BinA and Cqm1 proteins, in a total volume of 100 µL were treated with 5 µL of 2.5% freshly prepared solution of glutaraldehyde for 5 min at 37 ºC. BinA and Cqm1 proteins were also treated individually with 2.5% glutaraldehyde as internal controls. The reaction was terminated by the addition of 20 µL of 1 M Tris-HCl (pH, 8.8). Resulting products from chemical crosslinking were adjudged on 10% SDS-PAGE.

BinA and BinB interaction was probed through SPR using Biacore T200 SPR system. The experiment was carried out at 25 ºC. Nearly, 1014 response units of BinA protein (ligand) were immobilized onto the CM5 Biacore sensor chip via amine coupling followed by extensive washing with buffer E. Different concentrations (0.156, 0.3125, 0.625, 1.25, 2.5, 5 and 10 µM) of BinB protein (analyte) in buffer E were injected onto the BinA-bound sensor chip at 30 µL/min flow-rate. The analyte was allowed to interact with the immobilized

protein for 60 sec before washing off with buffer E. Kinetic analysis of the data was done using Biacore evaluation software (V3.1) provided with the instrument.

5.3.10. Carbohydrate inhibitory assay for BinA toxicity

Toxicity assay was performed for BinA, BinA pre-incubated with arabinose, fucose, mannose and glycoprotein fetuin. BinA concentration of 20 ng/mL was tested against 3rd instar larvae of *Culex quinquefasciatus*. Arabinose, fucose, mannose, and fetuin were used in control experiments. Simple sugars with 1000-fold higher concentration and fetuin with 10-fold higher concentration, compared to BinA, were incubated with BinA for 1 hr in buffer C. The toxicity was tested in 10 mL of sterile distilled water containing ten 3rd instar larvae with two replicates for each sample in two independent experiments. The larval mortality was recorded after 48 h and was corrected by Abbott method [85] using mortality observed in water and buffer only control experiments, each having 10 larvae respectively, with two replicates.

5.4. RESULTS

5.4.1. Purification of proteins and labelling of BinA with acrylodan

Recombinant BinA (active form; 15−366 amino acids), BinB (residues 17−448) and Cqm1 (residues 23−560) protein were purified to homogeneity using IMAC and AEC chromatography methods (Fig. 5.2A). BinA protein was successfully labeled with acrylodan and excess of the unbound fluorophore was removed through extensive dialysis. Further, the labeled protein was purified to homogeneity using size exclusion chromatography and adjudged on a 12% SDS-PAGE gel (Fig. 5.2B). The percentage labeling efficiency was found to be 94%. Spectrofluorometric analysis of BinA following fluorophore labelling displayed Ac-BinA emission spectrum with λem ~480 nm (expected for acrylodan emission) upon excitation at 391 nm, which confirmed positive labelling of BinA protein with acrylodan (Fig. 5.2C).

5.4.2. Hemagglutination and hemolytic activity

BinA did not show any hemagglutination even in the presence of Ca^{2+} (Fig. 5.3). However, weak hemolysis was observed in the presence of Ca^{2+} (Table 5.1).

Table 5.1: **Hemolytic assay result**

Protein	Concentration (mg/mL)	% hemolysis at A405		% hemolysis at A540		% hemolysis at A630	
		Without Ca^{2+}	With Ca^{2+}	Without Ca^{2+}	With Ca^{2+}	Without Ca^{2+}	With Ca^{2+}
BinA	0.4	0	11.5	0	12.2	0	8.6
	0.2	0	16.5	0	16.6	0	10.9
	0.1	0	35	0	34.7	0	24.8
	0.05	0	20.6	0	21.3	0	15.3
	0.025	0	31.2	0	26.7	0	18.5
	0.0125	0	19.8	0	17.2	0	12.9
	0.00625	19	25.9	8	17.2	0	11.2
BinAB	0.4	2.2	5.8	0.77	8.4	0	3.8
	0.2	0	28.3	0	29.6	0	22.6
	0.1	0	5.4	0	7.0	0	8.8
	0.05	0	32.5	0	30.1	0	17.6
	0.025	0	7.1	0	0.7	0	0
	0.0125	0	22.5	0	14.0	0	4.1
	0.00625	0	0	0	0	0	0

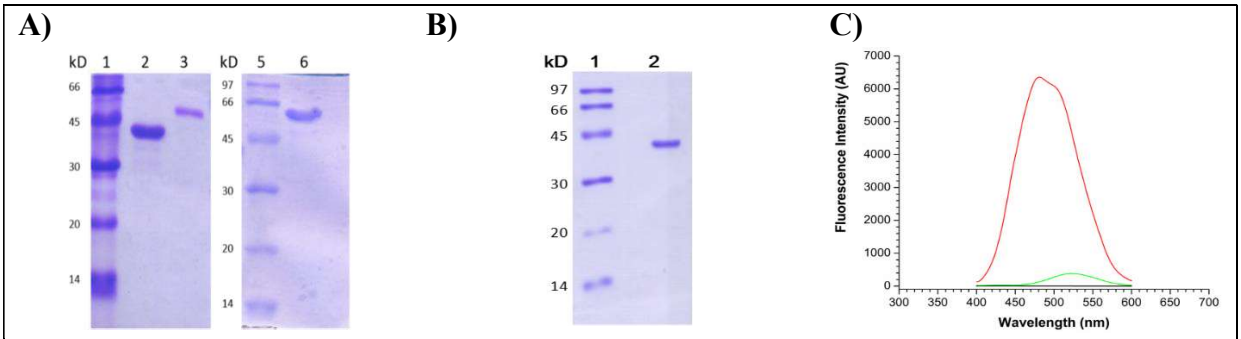


Figure 5.2. Purification and acrylodan labelling of BinA. A) 15% SDS-PAGE showing recombinantly purified BinA (lane 2), BinB (lane 3) and Cqm1 (lane 6) proteins. Lanes 1 and 5 show the protein molecular weight markers. **B)** 15% SDS-PAGE gel to show homogeneously purified acrylodan labeled BinA (lane 2). Lanes 1 shows the protein molecular weight markers. **C)** Spectrofluorometric analysis confirming positive labeling of BinA with acrylodan. Acrylodan labeled BinA (Ac-BinA, red line); Free acrylodan (green line) and unlabeled BinA (black line). Ac-BinA emission spectrum showed λem ~ 480 nm (expected for acrylodan emission) upon excitation at 391 nm.

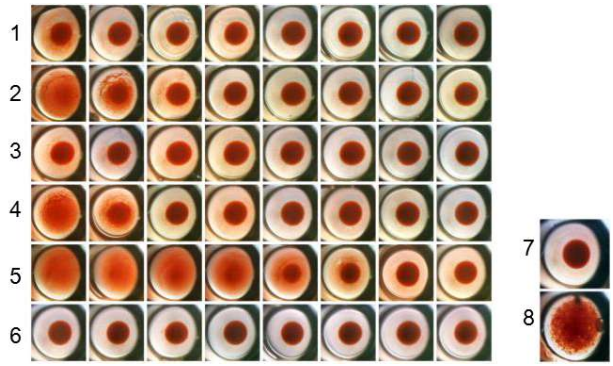


Figure 5.3. Hemagglutination activity of BinA with rabbit RBCs. Lane 1, BinA, 0.4 mg/mL (serial two-fold dilutions); Lane 2, BinA, 0.4 mg/mL (serial two-fold dilutions) in the presence of 1 mM Ca^{2+}; Lane 3, BinAB, 0.4 mg/mL (serial two-fold dilutions); Lane 4, BinAB, 0.4 mg/mL (serial two-fold dilutions) in the presence of 1 mM Ca^{2+}; Lane 5, Phaseolus vulgaris lectin, 1 mg/mL (serial two-fold dilutions) used as positive control; Lane 6, BSA, 1 mg/mL (serial two-fold dilutions) used as negative control; Lane 7, PBS buffer; Lane 8, PBS buffer in the presence of 1 mM Ca^{2+}.

5.4.3. Screening the potential glycans through fluorescence studies

The three Trp residues contributed to the intrinsic fluorescence of BinA protein upon excitation at 280 nm with an emission maximum at 321 nm. The tested glycans did not result in significant change in the emission spectrum (wavelength or intensity) of the protein, except for glycoproteins (fetuin, asialofetuin and thyroglobulin). The glycoproteins showed quenching of intrinsic fluorescence (Fig. 5.4). Structural analysis of the BinA protein revealed that the three Trp residues are positioned at >20 Å from the CBD (Fig. 5.1). Intrinsic fluorescence might not be influenced by the changes brought upon by ligand binding to the CBD due to large distance. No effect of Ca^{2+} on the interaction of sugars with BinA protein was observed.

DTNB assay revealed the presence of one free and accessible thiol group in BinA protein. BinA structure confirmed this group to be Cys195 residue of BinA that is placed close to the CBD (Fig. 5.1) [40]. Acrylodan was used as a site-specific fluorescence probe, to overcome the limitations of intrinsic fluorescence. Acrylodan was expected to bind to this Cys195 residue of BinA. Among simple sugars, arabinose, trehalose, fucose and N-acetyl-D-lactosamine, and polysaccharides, glycogen and inulin changed the fluorescence intensity (>10%) of the Ac-BinA (Fig. 5.5, Table 5.2). Glycoproteins (fetuin, asialofetuin and thyroglobulin) enhanced the fluorescence intensity significantly, accompanied with a blue-shift in the emission maximum of Ac-BinA (Fig. 5.5). These alterations in the fluorescence indicated that the fluorophore attached to BinA experienced a relatively more non-polar environment upon binding to the glycoproteins. The presence of divalent Ca(II) ion did not effect the extrinsic fluorescence of Ac-BinA in the presence of different sugars.

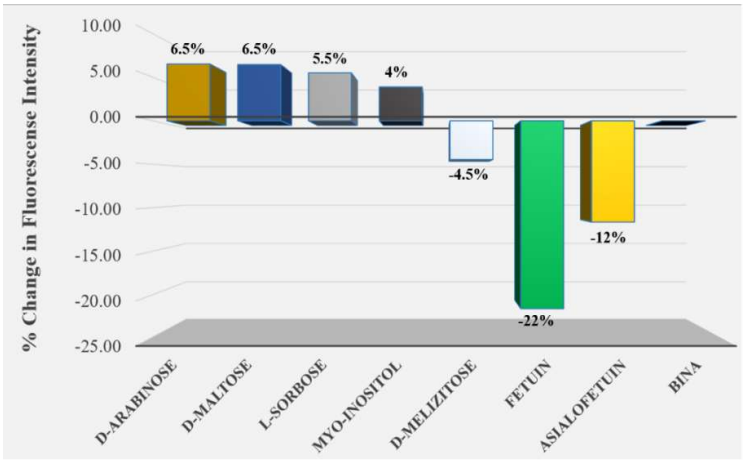


Figure 5.4. Intrinsic fluorescence analysis of BinA interaction with saccharides. Effect (% decrease or increase of intensity) of different glycans on the intrinsic fluorescence of BinA. *Thyroglobulin saturated the lamp, hence not shown.

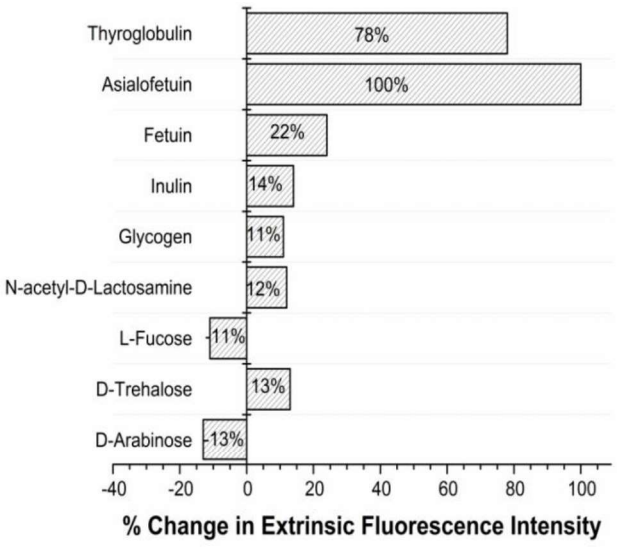


Figure 5.5. Extrinsic fluorescence analysis of BinA interaction with saccharides. The figure shows the comparative fluorescence of promising ligands which showed significant (>10%) increase in the emission intensity. *Asialofetuin showed increase in intensity which saturated the lamp.

5.4.4. Effect of glycans on the thermal stability of BinA

The stability of most proteins decreases with temperature [130]. Ligand binding has been shown to have a stabilizing effect on the protein resulting in an increase in its melting temperature [131, 132]. A 96-well plate format and SYPRO orange dye were used to study melting curve of BinA with or without the glycan ligands (Fig. 5.6A). SYPRO orange was used to measure increase in fluorescence due to binding of the dye to hydrophobic regions exposed due to thermal denaturation. The melting temperature (T_m) of 54 ^oC was determined from the first derivative of thermofluor-shift melting curve for BinA. Significant changes were not observed in the T_m value in the presence of used ligands and values vary in the narrow range of 54 /- 0.5 ^oC (Fig. 5.6B, Table 5.2). As larger T_m shifts are observed due to hydrophobic (entropy driven) binding of protein-ligand complexes [81], it may be inferred that either used sugars do not interact with BinA or binding process is enthalpy driven. The earlier studies suggested that free-energy of binding is predominantly dominated by enthalpy in carbohydrate interaction [133].

A)

	1	2	3	4	5	6	7	8	9	10	11	12
A	Buffer	BinA	BinA / sugar 1	BinA / sugar 2	BinA / sugar 3	BinA / sugar 4	BinA / sugar 5	BinA / sugar 6	BinA / sugar 7	BinA / sugar 8	BinA / sugar 9	BinA / sugar 10
B	BinA / sugar 11	BinA / sugar 12	BinA / sugar 13	BinA / sugar 14	BinA / sugar 15	BinA / sugar 16	BinA / sugar 17	BinA / sugar 18	BinA / sugar 19	BinA / sugar 20	BinA / sugar 21	BinA / sugar 22
C	BinA / sugar 23	BinA / sugar 24	BinA / sugar 25	BinA / sugar 26	BinA / sugar 27	BinA / sugar 28	BinA / sugar 29	BinA / sugar 30	BinA / sugar 31	BinA / sugar 32	BinA / sugar 33	BinA / sugar 34
D	Sugar 1	Sugar 2	Sugar 3	Sugar 4	Sugar 5	Sugar 6	Sugar 7	Sugar 8	Sugar 9	Sugar 10	Sugar 11	Sugar 12
E	Sugar 13	Sugar 14	Sugar 15	Sugar 16	Sugar 17	Sugar 18	Sugar 19	Sugar 20	Sugar 21	Sugar 22	Sugar 23	Sugar 24
F	Sugar 25	Sugar 26	Sugar 27	Sugar 28	Sugar 29	Sugar 30	Sugar 31	Sugar 32	Sugar 33	Sugar 34		
G												
H												

B)

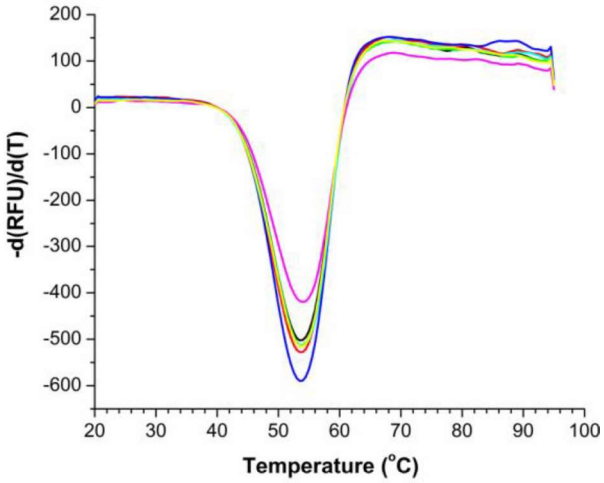


Figure 5.6. Thermofluor shift assay to determine the effect of sugars on the thermal stability of BinA protein. **A**) Ninety-six well plate format followed for thermofluor shift assay. The used sugars correspond to serial numbers given in Table 5.2. **B**) Derivative of the change in fluorescence due to SYPRO orange binding. The figure shows the relative thermal stability of BinA in the presence of sugars screened through extrinsic fluorescence analysis; BinA alone, black line; BinA and arabinose, red line; BinA and trehalose, green line; BinA-fucose, blue line; BinA and N-acetyl-D-lactosamine, cyan line; BinA and glycogen, magenta line and BinA and inulin, yellow line.

Table 5.2: High throughput screening of BinA against different sugars.

S.no.	Sugars and their derivatives	Effect on BinA melting temperature (°C)	% Change in extrinsic fluorescence (#)	Effect of Sugars on BinA toxicity (*)
	BinA	54		
1	D-arabinose	53.5	**- 12.9**	ND
2	L-arabinose	54	- 2.9	ND
3	D-ribose	54	0.2	ND
4	D-fructose	53.5	0.5	ND
5	D-maltose	54	0.9	ND
6	D-mannose	53.5	0.1	No effect
7	D-galactose	54	7.2	No effect
8	D-glucose	53.5	7.3	No effect
9	L-sorbose	53.5	7.2	ND
10	D-raffinose	53.5	- 4.0	ND
11	L-fucose	53.5	**- 11.4**	No effect
12	L-rhamnose	54	6.7	ND
13	Methyl-α-D-mannopyranoside	53.5	- 5.1	ND
14	Methyl-β-D-glucopyranoside	53.5	2.4	ND
15	Methyl-α-D-glucopyranoside	53.5	- 0.7	ND
16	Myo-Inositol	54	- 6.0	ND
17	Adonitol	53.5	- 1.3	ND
18	D-trehalose	53.5	**13.1**	ND
19	D-melibiose	53.5	- 2.1	ND
20	D-melizitose	54	7.7	ND
21	β-D- gentiobiose	53.5	- 0.1	ND
22	D-lactose	53.5	6.8	ND
23	Methyl-α-D-galactopyranoside	53.5	8.3	ND
24	Methyl-β-D-galactopyranoside	53.5	- 0.7	ND
25	N-acetyl-D-glucosamine	53.5	2.1	Reduced by 50% at 42 mM
26	N-acetyl-D-galactosamine	54	1.7	Reduced by 50% at 60 mM
27	N-acetyl-D-neuraminic acid	54	1.9	Reduced by 50% at 10 mM
28	N-acetyl-D-lactosamine	54	**11.7**	ND
29	N-acetyl-D-mannosamine	53.5	- 6.5	ND
30	Glycogen	54	**10.7**	ND
31	Inulin	53.5	**13.9**	ND
32	Fetuin	ND	**23.9**	ND
33	Asialofetuin	ND	**Saturation**	ND
34	Thyroglobulin	ND	**78.3**	ND
35	D-glucosamine	ND	ND	No effect
36	D-galactosamine	ND	ND	No effect
37	Muramic acid	ND	ND	Reduced by 50% at 25 mM
38	N-acetyl-muramic acid	ND	ND	Reduced by 50% at 10 mM
39	N, N'-diacetylchitobiose	ND	ND	Reduced by 50% at 10 mM
40	N, N', N"-triacetylchitotriose	ND	ND	Reduced by 50% at 10 mM

*) Values reported by Broadwell & Baumann, 1987.
#) >10% change in the extrinsic fluorescence of acrylodan labeled BinA was considered significant.
ND: Not determined; NS: Not significant

5.4.5. Validating the potential simple sugars through ITC, FTIR and SPR

Simple sugars (arabinose, fucose, trehalose and N-acetyl-D-lactosamine), which showed promising interaction with BinA during high-throughput screening based on extrinsic fluorescence, were analyzed further for their affinity by ITC, FTIR and SPR. It is generally believed that ITC can detect protein-ligand interactions in the micromolar ranges. In the ITC analysis, none of the four sugars showed any significant interaction with BinA at the standard ligand concentration of 1 mM. At high N-acetyl-D-lactosamine concentration of 20 mM, the binding reaction showed significant but identical heat change at all injections (Fig. 5.7). We anticipate that this behavior is an artifact due to excess ligand. The ITC data thus suggests weak ligand binding with values of dissociation constant outside the detection limits of ITC.

Water-mediated interactions are known to play key role in carbohydrate-lectin binding [134]. FTIR spectroscopy is sensitive to changes in the hydrogen bonding patterns of the protein upon ligand binding. The OH stretching band of water (2900−3000 cm^{-1}) represents the changes in water absorption by the protein upon ligand binding [83]. FTIR spectroscopy was done for a qualitative analysis of ligand binding with BinA. Figure 5.8 shows the difference spectra revealing the absorbance changes in the OH stretching band (2900−3000 cm^{-1}) due to ligand binding for simple sugars like arabinose, fucose and trehalose. The difference spectra show some indications of changes in the hydrogen bonding pattern in 2900−3000 cm^{-1}. However, significance of these changes in relation to ligand binding is difficult to interpret.

Kinetic evaluation of the interaction between BinA and simple sugars was then probed in real-time with SPR. BinA protein (ligand) was immobilized on a Biacore CM5 sensor chip and simple sugars (analytes) were introduced through mobile phase in independent experiments. The increase in SPR signal was observed for each sugar. However, a reliable kinetic model could not be fitted for the observed SPR profiles. The interaction of

each sugar was adjudged from the curve fitted to a plot of Req vs ligand concentration by steady state affinity analysis, which showed that simple sugars, namely, arabinose and fucose interact weakly with BinA with *K*D values in micromolar range (19.5 ± 2.9 and 5.7 ± 2.4 μM, respectively), while trehalose interaction was insignificant (Fig. 5.9). Also, interaction for N-acetyl-D-lactosamine could not be confirmed since it showed non-specific binding, as adjudged from binding-curve profiles (Fig. 5.9). The estimated *K*D values may not be treated as highly reliable, since affinity constants were estimated for 1:1 interaction in the steady state affinity model, whereas structural analysis of BinA has revealed the presence of more than one carbohydrate binding site [40].

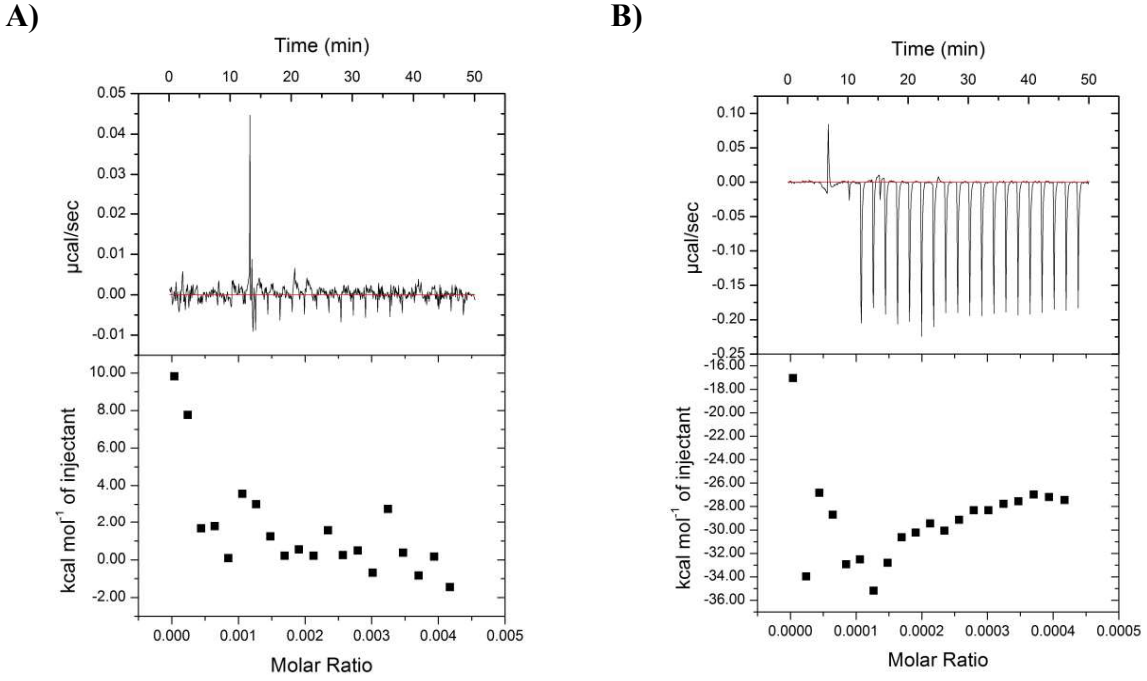


Figure 5.7. ITC analysis of BinA binding with simple sugars. The figures depict the profile obtained at two different ligand concentrations, **(A)** normal concentration of 1 mM **(B)** high concentration of 20 mM.

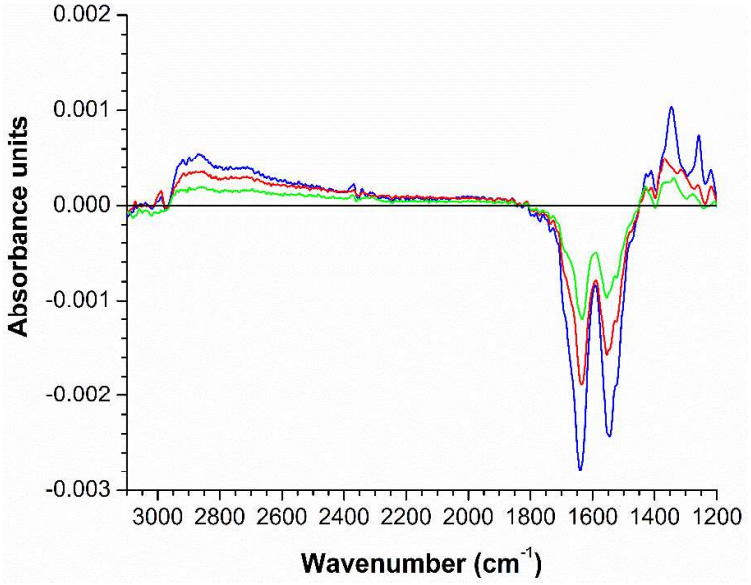


Figure 5.8. FTIR Difference spectra for BinA with arabinose (blue), fucose (red) and trehalose (green).

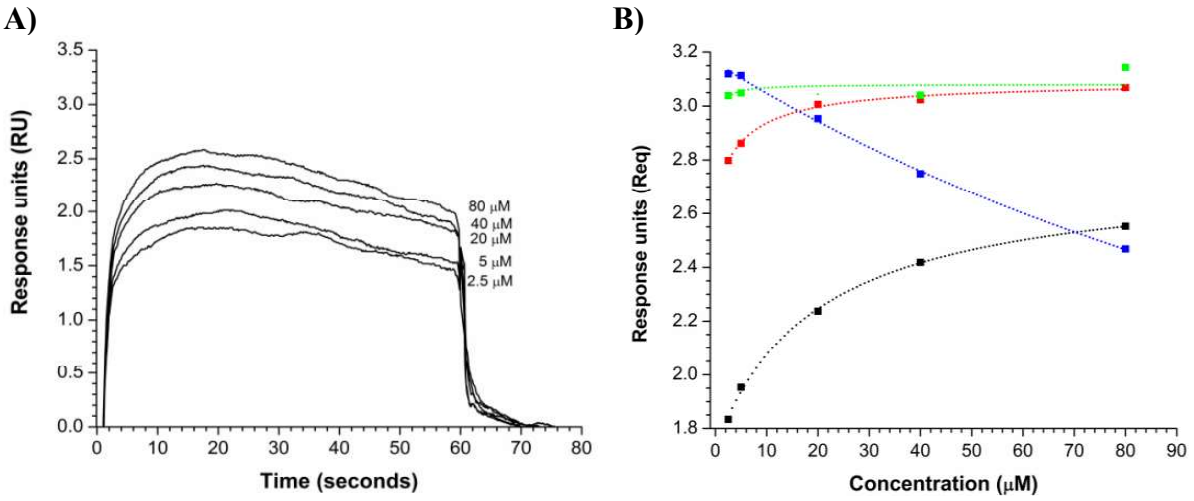


Figure 5.9. Probing interaction of simple sugars with BinA by surface plasmon resonance. **A**) A typical response curve is shown for BinA interaction with arabinose. **B**) BinA interaction with arabinose (black line), BinA interaction with fucose (red line), BinA interaction with trehalose (green line) and BinA interaction with N-acetyl-D-lactosamine (blue line) were adjudged by steady state affinity analysis. Solid square dotes represent the experimental points obtained at different concentrations; dotted line represents the theoretical fit.

5.4.6. BinA displays affinity for glycoproteins

The interaction between BinA and three glycoproteins (fetuin, asialofetuin and thyroglobulin) was quantified independently in real-time with SPR from the increase in SPR signal as a function of analyte concentration (Fig. 5.10). The best profile fitting with the integrated biphasic model of association was achieved for the association kinetics of all the three glycoproteins with BinA that describes interactions of two independent phases. The association and dissociation rate constants for asialofetuin for the first phase (*k*a1 and *k*d1) were estimated to be 4416.3 M^{-1} s^{-1} and 0.1 s^{-1}, while the rate constants for the second phase (*k*a2 and *k*d2) were estimated to be 94.2 M^{-1} s^{-1} and 0.011 s^{-1}, respectively. Thyroglobulin displayed higher affinity with *k*a1 and *k*d1 values of 7.4 x10^4 M^{-1} s^{-1} and 0.1 s^{-1}, *k*a2 and *k*d2 values of 227.4 M^{-1} s^{-1} and 0.015 s^{-1}, respectively. And fetuin interactions with BinA were observed to be weak with *k*a1 and *k*d1 values of 3375.7 M^{-1} s^{-1} and 0.058 s^{-1}, *k*a2 and *k*d2 values of 276.8 M^{-1} s^{-1} and 0.007 s^{-1}, respectively.

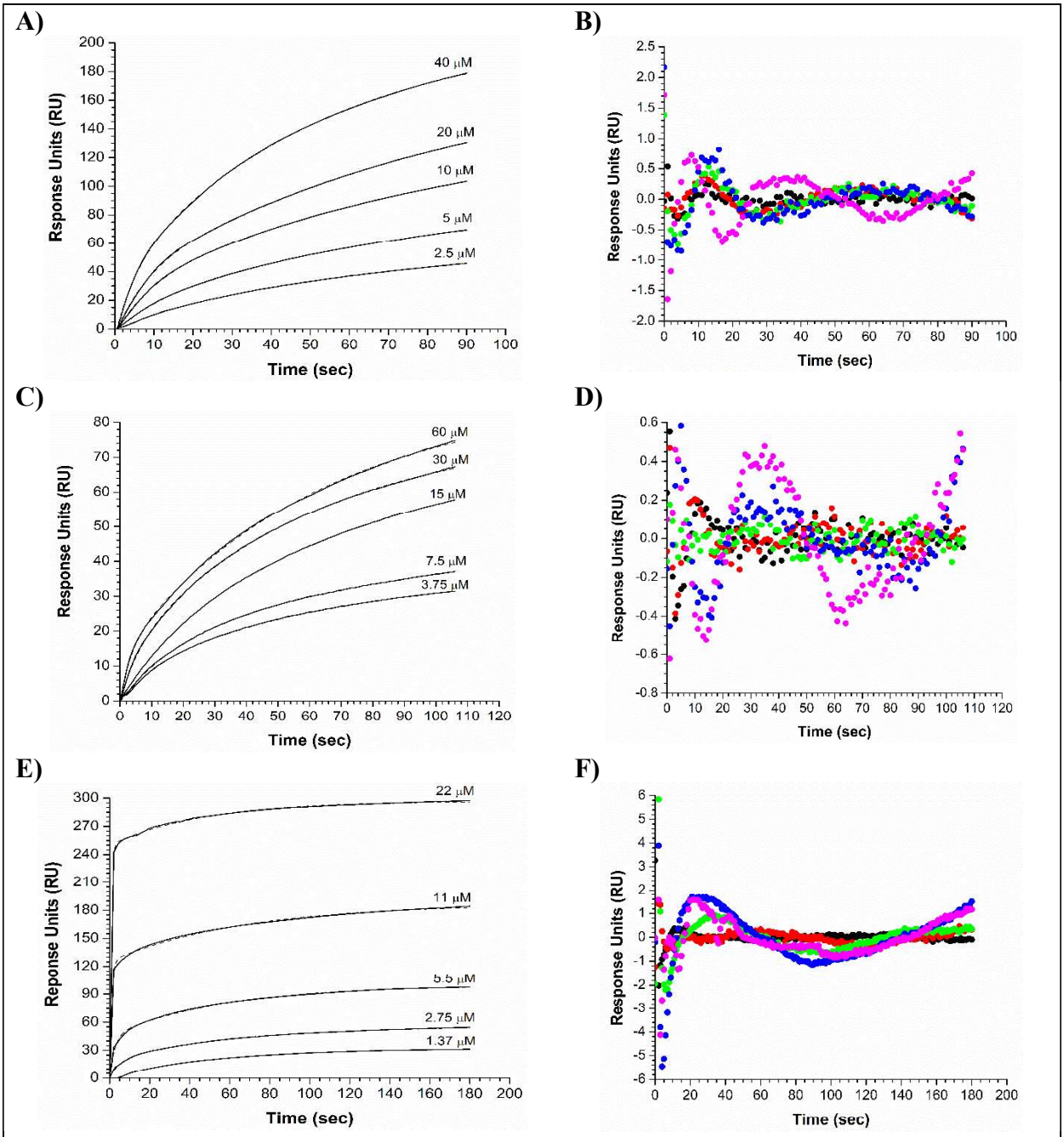


Figure 5.10. SPR analysis of BinA interaction with glycoproteins. Response curve for BinA (ligand) interaction with analytes **(A)** fetuin; **(C)** asialofetuin and **(E)** thyroglobulin. The dash lines are the fitted curves against the experimental SPR data (solid lines) using the association kinetics analysis module. Analyte (glycoprotein) concentration is shown alongside. Residual plots for each analyte concentration: **B)** fetuin (black, 2.5 µM; red, 5 µM; green, 10 µM; blue, 20 µM; magenta, 40 µM); **D)** asialofetuin (black, 3.75 µM; red, 7.5 µM; green, 15 µM; blue, 30 µM; magenta, 60 µM); and **F)** thyroglobulin (black, 1.375 µM; red, 2.75 µM; green, 5.5 µM; blue, 11 µM; magenta, 22 µM).

5.4.7. BinA recognizes the receptor and BinB proteins with similar affinities

The elution profile from the SEC showed two main peaks. Each of these peaks corresponded to the elution volumes of the individual proteins, Cqm1 and BinA, respectively. Further resolution of these peaks on 12% SDS-PAGE indicated presence of only one protein in each peak, suggesting that stable BinA/Cqm1 complex was not formed under the conditions of the experiments (Fig. 5.11A). The earlier SPR experiments had suggested weak and transient interactions between Cqm1 and BinA proteins with slow-association and fast dissociation rate constants and *K*D value of 2.2 μM (Chapter 4). To probe further, we attempted indiscriminate crosslinking of these proteins by glutaraldehyde, which is known to stabilize weak and transient interactions by forming covalent bonds using primary amines and sulfhydryls groups of the two protein molecules in close proximity. However, analysis of the chemically crosslinked reaction mixture of BinA and Cqm1 on 10% SDS-PAGE gel did not reveal formation of any unique high molecular weight band (Fig. 5.11B). Taken together, the data may suggest weak and non-specific interaction between BinA and Cqm1 proteins.

Kinetic analysis of BinA interaction with BinB was carried out in real time using SPR. Increased mass density was observed at the sensor surface with increasing BinB (analyte) concentration. The best profile fitting, between the experimental data and the fitted curves, was achieved with the two-state reaction kinetic model (Chi^2 0.76). The association and dissociation rate constants for the first phase of interaction (*k*a1and *k*d1) were 1.98 x 10^4 $M^{-1}s^{-1}$ and 0.07 s^{-1} respectively, while for the second phase of interaction (*k*a2 and *k*d2) were 0.01905 s^{-1} and 0.03054 s^{-1} respectively, with equilibrium dissociation constant (*K*D) value of 2.18 μM. The two-state reaction kinetic model describes a 1:1 binding of analyte to the immobilized ligand followed by a conformational change that stabilizes the complex (Biacore, GE Healthcare).

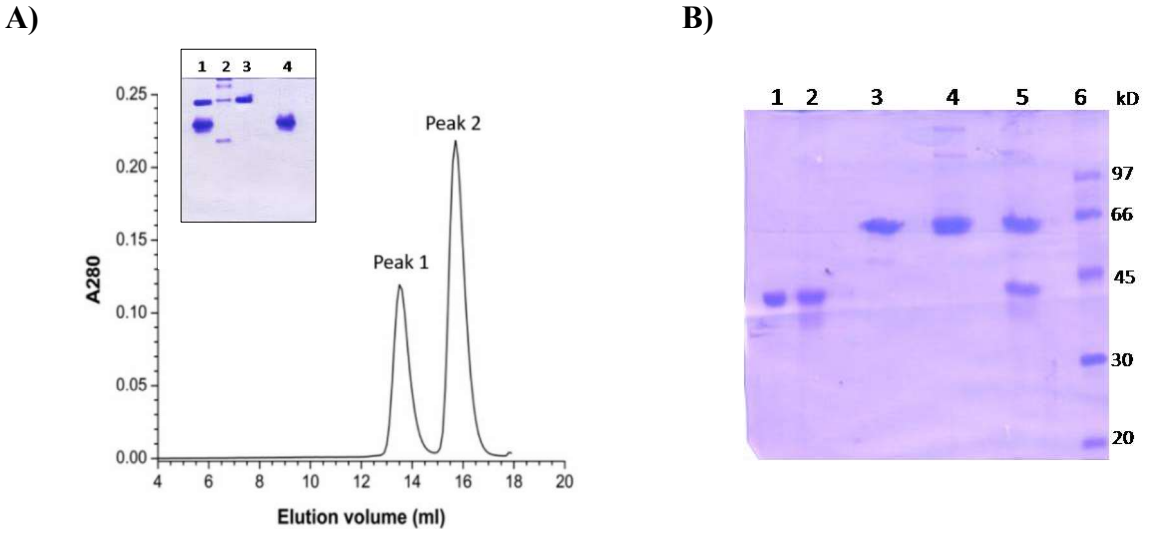


Figure 5.11. Interaction of BinA with Cqm1 protein. **A**) Size exclusion chromatography elution profile of mixture of BinA and Cqm1 proteins. The position of proteins as eluted when loaded individually on the same column is shown. In the inset 12% SDS-PAGE analysis of the elution peaks from size exclusion chromatography is shown. Lane 1, Cqm1 BinA mixture; Lane 2, standard molecular weight protein markers (97, 66, 30 kDa); Lane 3, Elution peak 1; Elution peak 2. **B**) 10% SDS-PAGE gel showing the results of chemical crosslinking with glutaraldehyde. Lane 1, BinA; Lane 2, BinA treated with glutaraldehyde; lane 3, Cqm1; Lane 4, Cqm1 incubated with glutaraldehyde; Lane 5, BinA and Cqm1 incubated with glutaraldehyde; Lane 6, protein molecular weight markers.

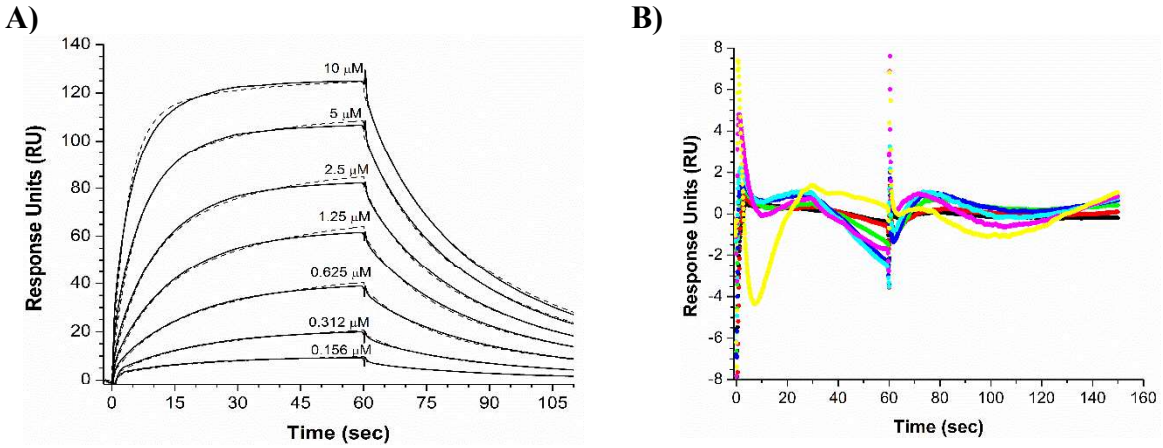


Figure 5.12. Interaction of BinA with BinB protein. SPR Response curve for BinA (ligand) and BinB (analyte). Analyte concentrations are indicated alongside. Solid line is the experimental curve, dash line is theoretical curve for two-state reaction. **B)** Residual plot for each analyte concentration. Black, 0.156 µM; red, 0.312 µM; green, 0.625 µM; blue, 1.25 µM; cyan, 2.5 µM; magenta, 5 µM; yellow, 10 µM.

5.4.8. Effect of glycans on BinA toxicity

The recombinant BinA alone showed toxicity against the Culex larvae with an LC50 value of 21.1 ng/mL [33]. In the present experiments 20 ng/mL of BinA led to nearly 32% mortality of 3rd instar Culex larvae. Lowering of BinA toxicity was observed in the presence of tested glycans, arabinose, fucose and fetuin (Fig. 5.13; Table 5.3). No change in mortality was observed in control experiment with mannose. Surprisingly, fetuin alone, without BinA, also showed some toxicity by an unknown mechanism (Table 5.3). These data suggest that antagonist effect of the tested glycans is due to their inhibition of carbohydrate binding activity of BinA.

Table 5.3: Toxicity analysis for determining the effect of glycans on BinA larvicidal activity.

Protein	% Abbott mortality
BinA	32
BinA/ arabinose	0
arabinose alone	0
BinA/ fucose	3
fucose alone	3
BinA/ fetuin	10
Fetuin alone	12
BinA/ mannose	36
mannose alone	2

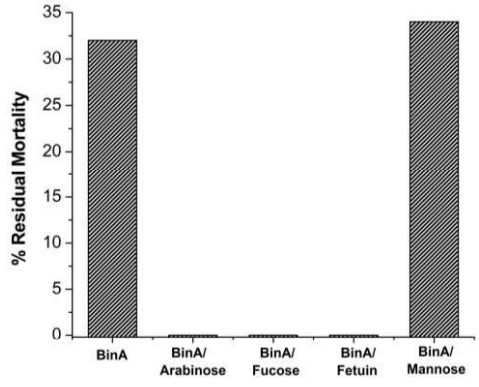


Figure 5.13. Mortality (%) representing the effect of glycans on BinA toxicity. Highly reduced residual larval mortality was observed in the presence of BinA/arabinose, BinA/fucose and BinA/fetuin mixtures as compared to BinA alone. The larval mortality was calculated by subtracting mortality due to control (buffer/sugar) from the % mortality observed for BinA or BinA/sugar combinations. Mannose, used as negative control, did not reduce BinA toxicity.

5.5. DISCUSSION

Carbohydrates, as free oligosaccharides or as glycoconjugates, interact with carbohydrate-binding proteins, like lectins, and play an important role in many biological events [125]. Several simple sugars [chitobiose, chitotriose, muramic acid, sialic acid (N-acetyl-neuraminic acid), N-acetylmuramic acid, N-acetylgalactosamine, and N-acetylglucosamine] were found to decrease toxicity of BinA against *Culex* cells [27]. Crystal structure analysis also revealed that BinA is composed of lectin-like and aerolysin-like domains [40].

BinA protein did not display hemagglutination and hemolytic activities against rabbit RBCs (Fig. 5.3; Table 5.1). Fluorescence based approaches were employed in the present study for high-throughput screening of possible lectin substrates for binding specificity towards BinA. Extrinsic fluorescence analysis identified simple sugars (arabinose, fucose, trehalose, N-acetyl-D-lactosamine, glycogen and inulin) and glycosylated proteins (fetuin, asialofetuin and thyroglobulin) as promising glycans with affinity towards BinA. Thermofluor shift analysis and ITC experiments failed to detect interactions between BinA and simple sugars. FTIR analysis provided indications for interaction of arabinose and fucose with BinA. The SPR analysis, however, suggested binding of arabinose and fucose sugars with *K*D values in micromolar ranges. Affinity of BinA towards three N-glycosylated proteins (fetuin, asialofetuin and thyroglobulin) was also confirmed by SPR analysis. Highest affinity was observed for thyroglobulin with first phase association and dissociation rate constants of 7.4 $x10^4$ M^{-1} s^{-1} and 0.1 s^{-1}, respectively, while fetuin interactions with BinA were observed to be weak with *k*a1 and *k*d1 values of 3375.7 M^{-1} s^{-1} and 0.058 s^{-1}, respectively. The large observed differences in association/dissociation rate constants may suggest specificity of BinA towards the carbohydrate structure of glycoproteins. The complex oligosaccharides of fetuin are composed of galactose, mannose, sialic acid and N-acetyl-D-

glucosamine in nearly identical molar ratio [135]. The oligosaccharide composition of asialofetuin is similar to fetuin with probably little or no sialic acid [136]. The oligosaccharide fraction of thyroglobulin contains fucose in addition to galactose, mannose, sialic acid and N-acetyl-D-glucosamine [137, 138]. The stronger interaction between BinA and thyroglobulin thus appear to involve fucose residue.

The present study, through high-throughput *in vitro* biophysical approaches, has been successful in characterizing the carbohydrate binding specificity of BinA. It clearly demonstrates that BinA is competent to bind diverse and structurally different glycosylated proteins with differential affinity. BinA displayed affinity towards Cqm1 (Chapter 4) and BinB proteins with *K*D ~2 μM. Similar affinity constants were observed for simple sugars in the present study. Also, the competitive binding assay showed loss of BinA toxicity in the presence of the tested glycans (arabinose, fucose and fetuin). These results are suggestive of the mechanism for cytotoxic effects of BinA. BinA may bind diverse glycosylated proteins inside the cell, while recognition of the receptor may be mediated by dual interaction of BinA with glycan core of GPI anchor of the membrane bound receptor and with the receptor polypeptide. It also rationalizes non-toxic effect of BinA on *Aedes aegypti*, as BinA binding to glycosylated Aam1 [53] may block internalization of BinA inside the cell.

5.6. CONCLUSION

- BinA protein does not display hemagglutination or hemolytic activity.
- BinA has differential affinity for structurally diverse glycosylated proteins and also interacts with simple sugars.
- Ligands screened for positive interaction with BinA, like arabinose, fucose and fetuin, inhibited the toxicity of BinA protein.
- As glycosylated proteins are known to be essential for larval development, binding of BinA with these inside the cell may be responsible for cell death.
- The glycan specificity of BinA may also be responsible for recognition and accumulation of toxic component at the epithelial cell membrane inside larval gut in *Aedes*, which are refractory to cytotoxic effects of BinAB toxin.

Chapter 6

PEGylation enhances mosquito-larvicidal activity of *Lysinibacillus sphaericus* Binary toxin

6.1. ABSTRACT

Toxic strains of *Lysinibacillus sphaericus* have been used in field for larval control of mosquito vector diseases. The high toxicity of *L. sphaericus* is attributed to the binary (BinAB) toxin produced as parasporal crystalline inclusions during the early stages of sporulation. BinA and BinB, the primary components of these spore-crystals, exert high toxicity when administered together. But instability, short half-lives and rapid proteolytic digestion can limit their use as an effective insecticide. BinA alone displays larvicidal toxicity, in the absence of BinB, *albeit* with much reduced activity. The present work for the first time demonstrates the beneficial effect of PEGylation (covalent attachment of polyethylene glycol) on mosquito-larvicidal activity of BinA. Polymer conjugation was achieved using 750 Da polyethylene glycol (PEG) at two different pH values (pH 7.2 and 8.5). Two different isoforms of the biopolymers, purified to homogeneity, were highly water soluble and resistant to trypsin and proteinase K. The mono-PEGylated BinA isoforms also displayed preservation of the toxin structure with improved thermal stability by about 3–5 ^{o}C, as evident from thermal denaturation studies by circular dichroism and thermofluor shift analysis. Notably, PEGylation enhanced BinA toxicity by nearly 6-fold. The PEGylated BinA isoforms alone displayed high larvicidal activity (LC_{50} value of ~3.4 ng/mL) against 3rd instar *Culex* larvae, which compares favourably against LC_{50} reported for the combination of BinA and BinB proteins. Since BinA can be synthesized easily through recombinant technology and easily PEGylated, the conjugated biopolymers offer a promising opportunity for mosquito control programs.

6.2. INTRODUCTION

Improvements in the efficacy of protein based biopesticides have been achieved through recombinant DNA technology by increasing their synthesis and/or enabling the production of combination of toxins from different bacteria in a single strain. For instance, combination of BinA/B and Bti toxins in the same bacterial species showed improved mosquito-larvicidal activity [34, 139]. However, instability, short half-lives and rapid digestion by the proteases of peptide based biopolymers, calls for their frequent administration at the breeding sites and limit their use as insecticide [140]. Modification of the useful biopesticides enhancing their stability and toxicity could provide viable means to use these more effectively in mosquito control programs. PEGylation of therapeutic proteins and peptides, by covalent linking of polyethylene glycol chains to the target molecules, is proving to be highly potent tool in their real-life applications [70, 141]. The most relevant advantages of PEGylation are known to be as, prolonged bio-availability, increase in stability, and reduced immunogenicity [68, 69, 73, 142, 143]. There are currently over 10 different PEGylated products approved by the US FDA, including PEG-adenosine deaminase, PEG-asparaginase etc [73]. Also, use of aliphatic PEG polymer in enhancing insecticidal activity of proteins was suggested by Jeffers et al., who observed an enhancement of nearly 10-fold in toxicity of a modified decapeptide against *Aedes aegypti* [144].

The present study reports conjugation of methoxy polyethylene glycol isocyanate to the toxic BinA component of mosquito-larvicidal BinA/B toxin. The resulting bioconjugate was characterized using biophysical approaches, like MALDI-TOF, dynamic light scattering (DLS), differential scanning fluorimetry (DSF) and circular dichroism (CD), for assessing extent of modification and thermal stability of the biopolymer. The modification was found to improve larvicidal efficacy by nearly 6-fold and enhanced thermal stability of secondary and tertiary structure of the protein. The PEG conjugated BinA protein was also found to be

resistant to broad-spectrum proteases. This is the first study probing the effect of PEGylation on the activity of a safe and world-wide used *L. sphaericus* BinAB toxin. The modification of recombinantly produced BinA can be easily achieved and can be helpful in reducing dependency on less stable BinB protein to achieve maximal mosquito-larvicidal activity. This is in-line with WHO recommendations on improvement and efficient implementation of mosquito control interventions.

6.3. MATERIALS AND METHODS

6.3.1. Materials

The methoxy polyethylene glycol isocyanate of MW 750 Da (mPEG-ISC-750) for PEGylation reaction was from NANOCS, USA. Lysozyme, barium chloride and IPTG were from SRL, Mumbai, India. The Ni-IDA matrix was from GE Healthcare, India. Others chemical reagents and culture media used were from Himedia, India. SYPRO® Orange protein gel stain, and trypsin and chymotrypsin proteases were from Sigma, and proteinase K was from New England BioLabs (NEB). Anti-polyethylene glycol and anti-PEG (methoxy group) antibodies were procured from Genetex. Alkaline phosphatase conjugated anti-rabbit IgG and NBT/BCIP were procured from Sigma and Roche, respectively. Bovine serum albumin and anti-His antibodies were obtained from Sigma.

6.3.2. Purification of BinA/B proteins

The active BinA (residues 15−366) with fused N-terminus poly-His tag was expressed and purified from BL21 (DE3) cells as described by Kale et al [32]. Briefly, BL21(DE3) cells transformed with pET28a-*bin*A construct, carrying truncated *bin*A gene (1056 bp) coding for residues 15-366 of BinA that constitutes active protein, were grown in LB medium supplemented with kanamycin (50 µg/mL) at 37 ºC. The culture was transferred to 18 ºC when it achieved an OD600 ~0.2 and was grown further. The protein expression was induced with 0.25 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at OD600 ~0.7 and cultured further for 4 hrs. The cells were harvested by centrifugation at 10000 x*g* for 10 min at 4 ºC and were re-suspended in the lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl, 1 mM PMSF, 15% sucrose, 5% glycerol, 2 mM DTT) containing lysozyme (2 mg/mL). The suspension was subsequently sonicated in pulse mode (5s ON-10s OFF, 15 s each) followed by centrifugation at 21000 x*g* for 30 min at 4 ºC. The supernatant was loaded onto the pre-equilibrated Ni-IDA

(Nickel-iminodiacetic acid) matrix. The matrix was washed with buffer (50 mM Tris pH 8.0, 100 mM NaCl, and 25 mM Imidazole). The bound protein was eluted using imidazole linear gradient (25–1000 mM) in buffer (50 mM Tris pH 8.0, 100 mM NaCl) over six column volumes. The buffer of the eluted protein was exchanged to buffer A (50 mM Tris pH 8.0, 20 mM NaCl) by dialysis. The recombinant protein was treated with thrombin at 37 ºC overnight to remove the poly-His tag. Removal of the tag was confirmed by western blotting using anti-His antibodies. The protein was further purified by anion exchange chromatography (High Q column, BioRad). The column was pre-equilibrated with buffer A. The bound proteins were eluted using a linear gradient of NaCl (20–1000 mM) in buffer (50 mM Tris pH 8.0) over six column volumes. The eluted fractions were adjudged on 12% SDS-PAGE. The purified recombinant BinA protein was dialyzed against 1X PBS at pH 7.2 and 8.5 for subsequent reactions and stored at 4 ºC.

The purification of BinA/B mixture (pro-BinA and pro-BinB proteins) from bacterial spore-crystals was achieved using highly toxic *L. sphaericus* ISPC-8 having serotype 5a5b and phage type 3. The ISPC-8 strain was grown till sporulation and the spore pellet collected by centrifugation. The purification of BinA/B from the spore pellet was carried out using ion-exchange and size exclusion chromatography as described by Hire et al [88]. The final purification of BinA/B mixture was achieved by size exclusion chromatography using Superdex-200 column (GE healthcare).

6.3.3. PEGylation of recombinant BinA

Methoxy polyethylene glycol isocyanate of MW 750 Da (PEGylation reagent) was used for BinA modification. Isocyanate derivatives of PEGs react readily at pH range of 7.0 to 9.0 to form covalent bond with free amino groups on the protein. The PEGylation reagent (reconstituted in DMSO) was added in excess to the purified recombinant BinA in the molar ratio of 1:50 at pH 7.2 and 8.5 in 1X PBS buffer in independent reactions. The reactions were

allowed to proceed overnight at 4 ºC with continuous stirring. The PEGylated BinA proteins, m1BinA (BinA modification at pH 7.2) and m2BinA (BinA modification at pH 8.5), were stored at 4 ºC.

6.3.4. Purification and oligomeric status of PEGylated BinA proteins

The PEGylated BinA proteins were purified through size exclusion chromatography using Superdex-200 column in order to separate differentially modified BinA molecules and for complete removal of unreacted excess of PEGylation reagent. The column was calibrated using standard molecular marker proteins (Bovine thyroglobulin, 669 kDa; Apoferritin, 440 kDa; Bovine serum albumin, 66 kDa; Ovalbumin, 44 kDa; Carbonic anhydrase, 29 kDa). The column was pre-equilibrated with 1X PBS buffer (pH 7.2) for purification of m1BinA and with 1X PBS buffer (pH 8.5) for m2BinA purification. The protein was loaded with the help of a syringe and the eluted fractions were collected and adjudged on 12% SDS-PAGE. The protein fractions eluted corresponding to monomeric state of recombinant BinA (~42 kDa) were collected and were dialyzed against buffer A. The dialyzed proteins were loaded onto pre-equilibrated High Q AEC column. The bound proteins were eluted with linear gradient (20–1000 mM NaCl) over six column volumes. The eluted PEGylated BinA proteins were dialyzed and stored in the buffer B (10 mM Tris-HCl, pH 8.0, 10 mM NaCl) at 4 ºC.

The recombinant and PEGylated BinA proteins in buffer B were analysed by dynamic light scattering using about 0.3 mg/mL protein concentration at 25 ºC with Malvern Zetasizer Nano ZS instrument. Three sets of measurement with 12 acquisitions in each set were collected. The data analysis was done with the help Malvern Zetasizer software suite (http://www.malvern.com).

6.3.5. Electrophoretic and MALDI-TOF analyses

The PEGylated BinA proteins, after final purification through AEC, were loaded onto 15% reducing SDS-PAGE gel and 12% native PAGE gel along with the purified recombinant BinA protein. After electrophoresis, the gels were stained with Coomassie solution. To visualize PEGylated proteins, SDS-PAGE was completely de-stained overnight and stained using the Barium iodide method from Kurfurst [145]. Briefly, the gel was soaked in 5% glutaraldehyde solution for 15 min at room temperature and was subsequently stained for the PEG moiety. First, the gel was maintained in 20 mL perchloric acid (0.1 M) for 15 min. Then 5 mL of 5% barium chloride solution and 2 mL of 0.1 M iodine solution were added according to the procedure as described by Skoog [146]. After about 15 min, the staining solution was replaced and gel was washed with water for another 15 min.

An unstained 15% SDS-PAGE gel containing recombinant BinA, m1BinA and m2BinA proteins was also electro-blotted on nitrocellulose membrane and was challenged with anti-polyethylene glycol and anti-PEG (methoxy group) antibodies using the antibody concentrations recommended by the manufacturer. Alkaline phosphatase-conjugated anti-rabbit IgG was employed as a second antibody, and color development was performed using 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and nitro-blue tetrazolium (NBT).

The PEGylated and recombinant BinA proteins were also analysed by MALDI-TOF to determine their molecular weights and extent of modification. The protein masses were estimated from the highest observed peak intensity in each spectrum.

6.3.6. Toxicity assays

The toxicity assays were performed for BinA/B purified from *L. sphaericus*, recombinant BinA and PEGylated BinA proteins (m1BinA and m2BinA). Prior to the toxicity assay, protein concentrations were determined by modified Bradford assay [78]. Different concentrations of the protein samples were tested against 3rd instar larvae of *Culex*

quinquefasciatus. The nucleus culture of *C. quinquefasciatus* was obtained from National Institute of Virology, Pune and was maintained at 28 ± 2 ºC and 85% relative humidity. Eggs were allowed to hatch in plastic bowls containing 1L tap water supplemented with 0.13g sterilized larval food (13:6:1, wheat flour: chickpea flour: yeast extract) [84]. Five concentrations each for the BinA/B, recombinant BinA and PEGylated BinA proteins were tested in 10 mL of sterile distilled water containing ten 3rd instar larvae with three replicates for each concentration. The experiments were done twice for each concentration on different days with the same batch of larvae to minimize variation between the larval batches, due to origin and seasons. As a control for the effect of PEG, bovine serum albumin (BSA) was modified by mPEG-ISC-750 and PEGylated BSA was assayed for toxicity against *Culex* larvae using concentration of 200 ng/mL. No toxicity was observed in the control experiments. The toxicity of m2BinA was also assayed against 3rd instar *Aedes* larvae using protein concentration of 100 ng/mL. All the proteins tested for toxicity were constituted in a low salt buffer (10 mM Tris-HCl, pH 8, and 10 mM NaCl).

The larval mortality was recorded after 48 h and was corrected by Abbott method [85] using mortality observed in water and buffer only control experiments, each having 10 larvae respectively, with three replicates. Mortality data was analysed using the Probit analysis [86] and the median lethal concentration (LC_{50}) along with the 95% confidence interval were calculated with the help of the R project for statistical computing (https://www.r-project.org/).

6.3.7. Estimation of secondary structure and thermal stability by Circular Dichroism

The effect on the secondary structure of BinA protein upon PEGylation was investigated using far-UV circular dichroism (CD). The spectra were recorded for the PEGylated BinA proteins (10 µM) and recombinant BinA (10 µM) in buffer B on a JASCO spectrometer (J-815) equipped with peltier thermostatic cell holder. The data were recorded in 200–260 nm range at 20 ºC in 1 mm Quartz cuvette. Each spectrum was obtained by

averaging three individual scans. The spectrum of the buffer blank was subtracted from the sample CD spectrum. Deconvolution and analysis of the CD spectra were carried out using K2D2 software available online (http://www.ogic.ca/projects/k2d2) [80].

Further, thermal denaturation measurements of recombinant and PEGylated BinA were performed using ellipticity at 218 nm from temperature-wavelength scan in the temperature range 20–75 ^{o}C with a scan rate of 1 ^{o}C/min. The CD spectra during temperature-wavelength scans were recorded in 200–260 nm wavelength range at regular temperature intervals (every increase of 5 ^{o}C between 20–50 ^{o}C and 2 ^{o}C thereafter). The transition temperature (T_m) of secondary structures was determined from the first derivative of temperature-wavelength scan spectra at 218 nm, as suggested by Greenfield [147]. For recording thermal denaturation in 20 mM phosphate buffer (including 20 mM NaCl), ellipticity at 218 nm was recorded in simple temperature scans, without acquisition over the complete wavelength range.

6.3.8. Estimation of thermal stability by Thermofluor shift assay

Thermofluor shift assay was carried out to study thermal denaturation of tertiary structure of the proteins as described by Niesen et al [81]. The recombinant and PEGylated BinA proteins (0.1 mg/mL in 1X PBS buffer) were assayed in a 96 well plate in duplicates. SYPRO orange dye solution (5000X concentrate in 100% DMSO) was added directly to the protein samples to a final concentration of 5X. The plate was properly sealed and data was recorded on Light Cycler® 480 Instrument II (Roche). The protein samples were heated from 25 ^{o}C to 95 ^{o}C at ramp rate of 1.2 ^{o}C/ min with 25 acquisitions at each point. Thermal denaturation of Hen Egg White (HEW) lysozyme was also probed in the same 96 well plate as an internal control. The fluorescence signal (λ_{ex}, 498 nm; λ_{em}, 580 nm) was recorded as a function of temperature. Experimental data were processed using the in-built software and the melting temperature for each protein was determined from the first derivative.

6.3.9. Assessment of stability against proteolytic enzymes

Recombinant BinA, m1BinA, m2BinA and BinA/B mixture proteins were assayed in independent experiments for their stability against proteolytic enzymes, like trypsin, proteinase K and chymotrypsin. The proteins were mixed with proteases in protease: protein ratio of 1:10 (w/w) for trypsin, and 1:50 for proteinase K and chymotrypsin. Lower concentrations were used for proteinase K and chymotrypsin as 185 and 232 cleavage sites, respectively, in BinA sequence were predicted using Peptide Cutter tool (http://web.expasy.org/peptide_cutter/) [77] for these proteases, compared to only 30 sites for trypsin enzyme. The protein-protease reaction mixtures were incubated at 37 ºC. Aliquots were collected at two different time points of 1 hr and overnight incubation. The reaction was terminated with 1 mM phenylmethylsulfonyl fluoride (PMSF). The collected aliquots were adjudged on 12% SDS-PAGE gel for assaying proteolytic stability of the proteins.

6.3.10. Assessment of the affinity of PEGylated BinA towards glycans using SPR

PEGylated BinA-glycan interaction was probed using Autolab Esprit SPR system with bare gold sensor chip. Thyroglobulin was the chosen glycan as it showed highest affinity with BinA (as described earlier in chapter 5). Nearly 400 response units of BinA (ligand) were immobilized onto the Bare gold sensor chip followed by extensive washing with buffer C (20 mM sodium phosphate pH 7.2, 20 mM NaCl). The concentration of BinA and stock solution of glycoprotein were estimated by Bradford assay [78]. Different concentrations (1.375, 2.75, 5.5 and 11 μM) of thyroglobulin (analyte), were injected onto the BinA-bound sensor chip in independent experiments. The analyte (mobile phase) was allowed to interact with the immobilized BinA for 300 sec before washing off with buffer C. The association kinetics analysis was done using Autolab kinetic evaluation software (V5.4) provided with the instrument. The interaction of analyte alone with bare gold chip was also monitored as a

control. The kinetic rate constants were determined from the fit against experimental SPR data.

6.4. RESULTS

6.4.1. Purification of BinA/B proteins

The recombinant BinA (active form; 15-366 amino acids) was expressed in BL21 (DE3) host cells as fusion protein with N-terminal poly-His tag and a thrombin cleavage site between poly-His tag and BinA using *p*ET28a-*bin*A construct [32]. The recombinant protein was purified from soluble cell lysate using IMAC affinity chromatography and High Q anion exchange chromatography (AEC). The BinA protein was found to exist as monomer, as was evident from its elution profile on size exclusion chromatography (SEC) column. The mixture of pro-BinA (41.9 kDa) and pro-BinB (51.4 kDa) proteins was also purified to homogeneity from spore-crystals of highly toxic local strain of *L. sphaericus* ISPC-8 using ion-exchange and size-exclusion chromatography (Fig. 3.1, lane 4) [88].

6.4.2. Synthesis and purification of PEGylated BinA

The purified recombinant BinA was used for the PEGylation reaction with mPEG-ISC-750 at two different pH values, pH 7.2 (m1BinA) and pH 8.5 (m2BinA). PEG-isocyanate derivative is useful for amine group conjugation, yielding stable urethane linkage. The reaction depends upon the nucleophilicity of the amine group. At pH 7.2, the N-terminus amine group can be expected to react with isocyanate due to its low pKa (7.7 ± 0.5) generally found in proteins [148]. At pH 8.5, however, amine group of accessible lysine residues (pKa 10.4 ± 1.1) can also form covalent linkage. After overnight reaction, the PEG conjugated proteins were purified through SEC using Superdex-200 column. Monomeric PEGylated BinA constituted the largest fraction (Fig. 6.1) for each of the PEGylation reaction. Highly

pure PEG conjugated BinA proteins were obtained using anion exchange chromatography and were used for further analysis.

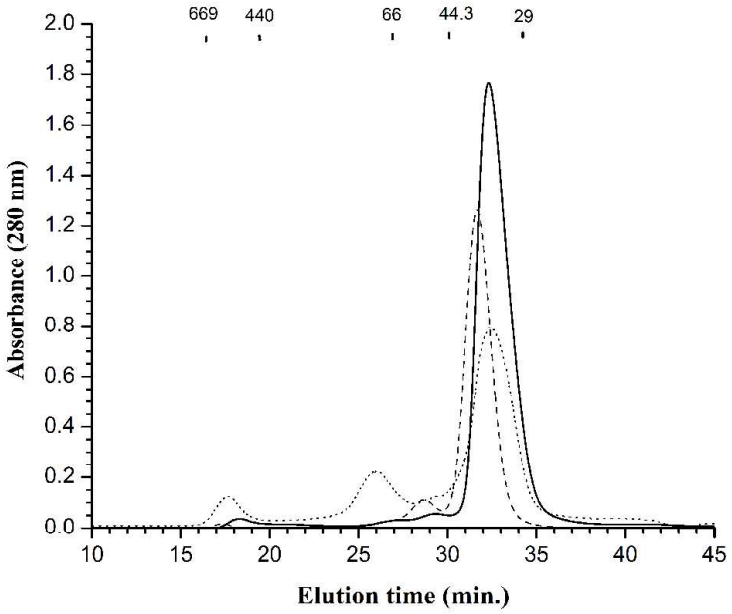


Figure 6.1. Purification of PEGylated BinA by SEC. Elution profiles of m1BinA (dotted line) and m2BinA (dash line) and recombinant BinA (solid line) from Superdex-200 column are overlaid. In each elution profile, the major peak corresponds to monomer of BinA and contains most of the BinA protein, as adjudged from SDS-PAGE analysis. The elution time of standard molecular markers are also shown. The minor peaks contain small amounts of oligomeric BinA, as adjudged by 12% SDS-PAGE.

6.4.3. Electrophoretic analysis

The purified recombinant BinA and PEGylated BinA (m1BinA and m2BinA) proteins were electrophoresed under native as well as denaturating conditions. In the Coomassie stained 15% SDS-PAGE the PEG conjugates showed marginal shift in the molecular weight towards the higher side (Fig. 6.2A). The presence of PEG moiety in the PEGylated BinA isoforms was, however, confirmed with the western blot analysis using anti-polyethylene glycol and anti-PEG (methoxy group) antibodies (Fig. 6.2B & C). The three proteins showed similar mobility on the native PAGE (Fig. 6.2D). The single bands of m1BinA and m2BinA in the native PAGE suggest homogenous-polymer modification of BinA at both the pH values.

The SDS-PAGE gel was also stained with barium iodide by the I_2-$BaCl_2$ method [145, 146]. The PEGylated BinA, however, did not stain differentially using barium-iodide method, as staining of unmodified BinA was also observed each time (Fig. 6.2E). We assume that the small size of PEG moiety used in the study could not be easily detected by the traditional PEG staining method.

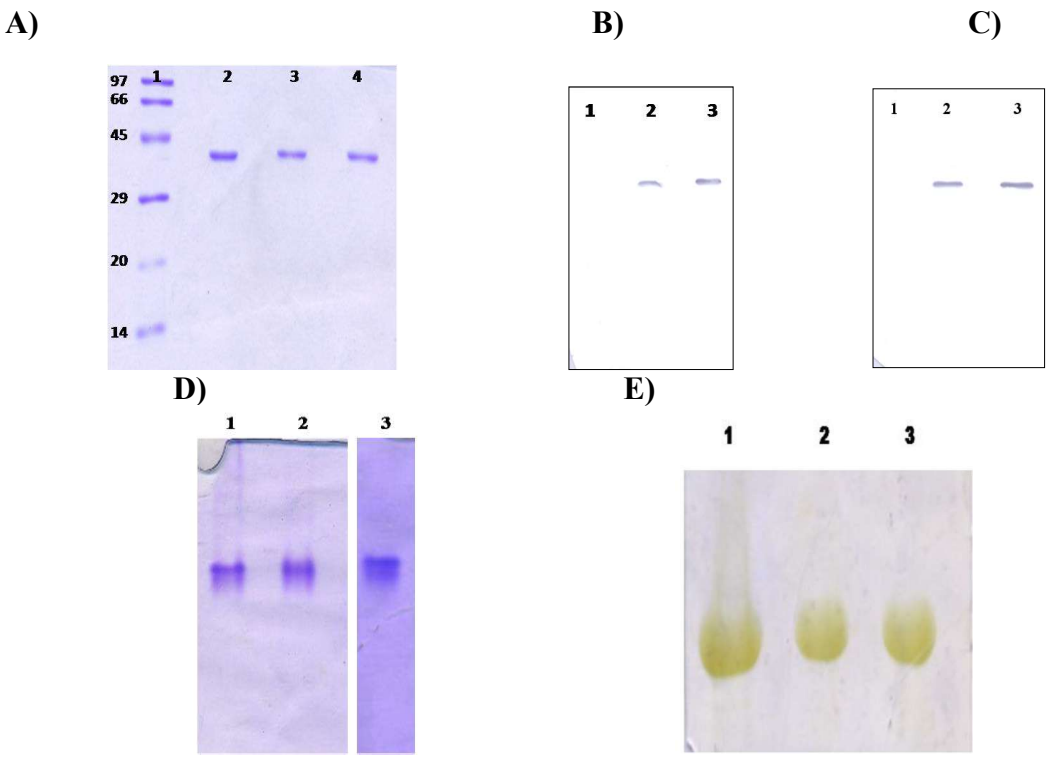


Figure 6.2. Electrophoretic analysis of PEGylated BinA after anion exchange chromatography. A) 15% SDS-PAGE. Lane 1, Protein molecular weight markers; Lane 2, purified recombinant BinA; Lane 3, purified m1BinA isoform; Lane 4, purified m2BinA isoform. **B & C**) Western blot analysis of m1BinA and m2BinA proteins resolved on 15% SDS-PAGE using anti-polyethylene glycol antibodies and anti-PEG (methoxy group) antibodies, respectively. The immune detection was visualized by secondary antibody and BCIP/NBT color development reagents. Unmodified BinA (lane 1) was not immunostained, while PEGylated m1BinA (Lane 2) and m2BinA (Lane 3) were clearly visible. **D**) 12% native PAGE of recombinant and PEG conjugated BinA proteins. Lane 1, BinA; Lane 2, m1BinA; Lane 3, m2BinA. **E**) **Barium iodide staining of PEG-BinA isoforms**. Lane 1, recombinant BinA; Lane 2, Isoform1 of PEGylated BinA synthesized at pH 7.2 (m1BinA); Lane 3, Isoform2 of PEGylated BinA synthesized at pH 8.5 (m2BinA). The gel was initially stained with Coomassie blue, completely destained overnight and stained using the Barium iodide method. Unexpectedly, both PEGylated and unmodified protein bands were observed to be stained brown, though marginally slower mobility of PEGylated BinA (due to conjugation of one PEG moiety of 750 Da) is clearly visible.

6.4.4. MALDI-TOF and DLS analysis

The PEGylated BinA isoforms and recombinant BinA were further analysed by MALDI-TOF to determine the state of PEGylation. The analysis showed single large spectral peak at m/z of 41,044 Da and 41,053 Da for the PEG conjugated BinA proteins at pH 7.2 and 8.5, respectively, whereas unmodified recombinant BinA showed a spectral peak at m/z of 40,152 Da (Fig. 6.3). The observed mass of 40,152 Da for BinA matches well with the expected mass of 40,142 Da estimated the amino acid sequence of the protein that is composed of BinA residues 15-366 and four residues (GSHM) at the N-terminus from the poly-His tag after thrombin treatment of the expressed protein. Further, the observed mass differences between PEGylated and unmodified BinA suggest single site PEGylation of BinA at both the pH values. The observation of single peak for both the PEG-BinA conjugates also suggested that the used purification protocol was effective in obtaining homogeneous PEG polymer of modified BinA.

The DLS experiments also confirmed PEG conjugation resulting in marginal increase in the hydrodynamic size of the modified proteins, compared to recombinant BinA. The hydrodynamic diameters for the unmodified, m1BinA and m2BinA proteins were estimated to be 5.8 ± 0.37, 5.98 ± 0.98 and 6.1 ± 1.3 nm, respectively (Fig. 6.4). The proteins were monodispersive with the estimated polydispersity of 5, 12.5 and 24%, respectively, suggesting that the oligomeric status of BinA did not change on modification.

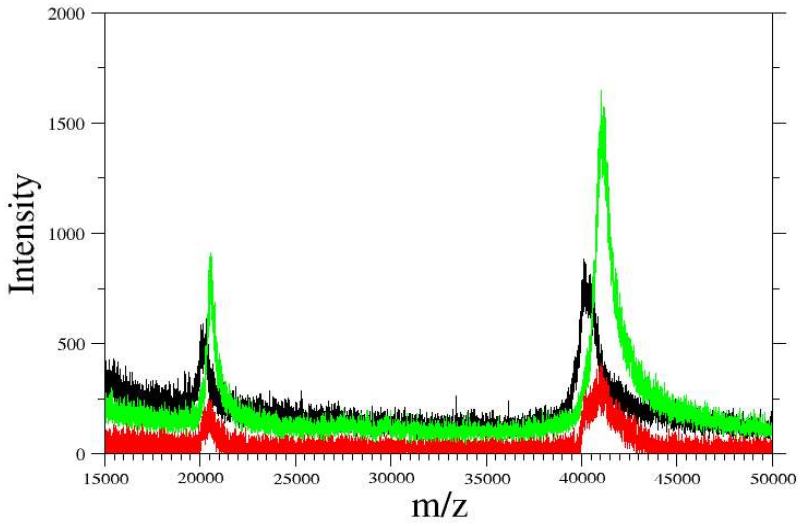


Figure 6.3. MALDI-TOF analysis of native and PEG conjugated BinA proteins. The MALDI-TOF analysis of recombinant BinA (black) and PEGylated BinA proteins, m1BinA (red) and m2BinA (green), showed large spectral peaks at m/z values of 40152 Da for BinA, 41044 Da for m1BinA and 41053 Da for m2BinA. Smaller peaks at about 20500 Da correspond to doubly charged species of these proteins.

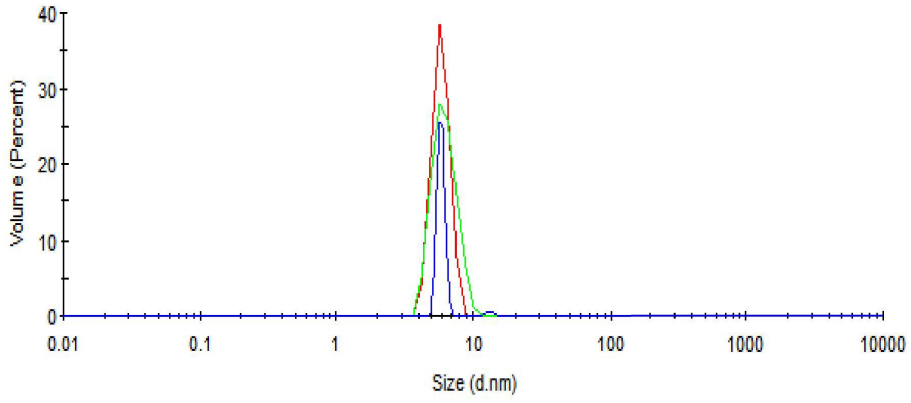


Figure 6.4. Dynamic scattering profiles of recombinant BinA (blue) and PEGylated BinA isoforms (m1BinA, red; m2BinA, green). The hydrodynamic diameters for the native BinA and PEGylated BinA isoforms were estimated to be 5.8 ±0.37, 5.98± 0.98 and 6.1± 1.3 nm, respectively, and polydispersity indices were 5, 12.5 and 24%, respectively.

6.4.5. Toxicity analysis

Toxicity of recombinant BinA, PEGylated BinA proteins and BinA/B mixture purified from *L. sphaericus* spore-crystals was assayed against 3rd instar larvae of *Culex quinquefasciatus* from the same generation and was performed on the same day. The recombinant BinA showed toxicity against the larvae with an LC_{50} value of 21.1 ng/mL. However, interestingly, PEGylated BinA isoforms, without BinB, displayed superior insecticidal characteristics with an LC_{50} values of about 3.4 ng/mL (Table 6.1; Fig. 6.5). In comparison, an LC_{50} value of 6.5 ng/mL was observed for BinA/B mixture. However, toxicity was not observed in the control experiments using PEGylated BSA against *Culex* and for m2BinA against *Aedes* Larvae.

Table 6.1: LC_{50} values of recombinant BinA (BinA), PEGylated BinA isoforms (m1BinA and m2BinA) and BinA/B mixture purified from *L. sphaericus* spore-crystals against 3rd instar *Culex quinquefasciatus* larvae.

Protein	LC_{50} value (ng/mL)	95% confidence values		χ^2-value	P-value
		Lower	Higher		
BinA	21.1	16.1	27.5	1.65	0.65
m1BinA	3.1	2.0	4.6	12.6	0.005
m2BinA	3.7	2.2	6.1	4.8	0.08
BinA/B	6.5	4.5	9.5	32.2	< 0.005

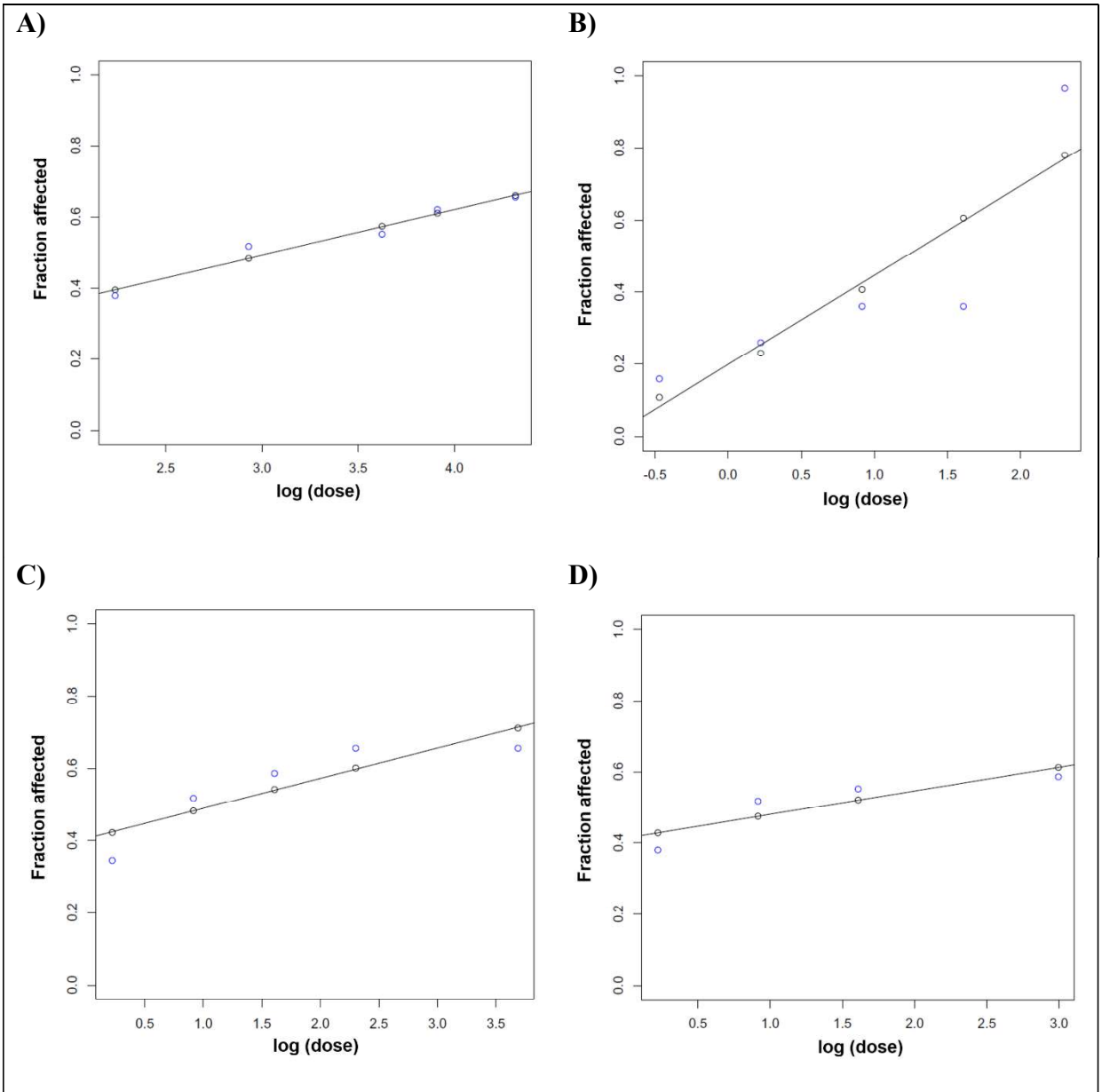


Figure 6.5. Dose response curve for the effect of PEGylation on the larvicidal activity of BinA. A) BinA **B**) BinA/B **C**) m1BinA **D**) m2BinA. Blue circles, experimentally determined values; black circles, theoretical fit using probit values estimated by R suite. The regression line is fitted against the estimated probit values.

6.4.6. PEGylation enhances thermal stability of BinA protein

To probe if PEG conjugation with BinA induced conformational changes in the protein structure, secondary structure was estimated for the recombinant and PEGylated proteins using CD spectroscopy (Fig. 6.6A). The deconvolution of CD data points, computed using the K2D2 software [80] revealed BinA and its PEGylated isoforms to be mainly comprised of β-structure with nearly 48% residues in β-strand conformation and nearly 2% residues in α-helices. The similarity of CD spectra of the three proteins suggested that the secondary structure of BinA was maintained on PEGylation.

Also, positive ellipticity for the three proteins was observed in the wavelength range 245-225 nm, which probably is due to tryptophan residues as observed for HCAII enzyme [149].

The thermal denaturation of the recombinant and PEGylated BinA proteins was studied using CD by monitoring ellipticity at 218 nm during temperature-wavelength scans. The loss of secondary structures was adjudged from the complete CD spectrum at different temperatures. The temperature range showing loss of secondary structure was used to estimate apparent transition temperature from the first derivative of CD values at 218 nm. The unmodified BinA protein showed a T_m of 55 ^oC, whereas the m1BinA isoform exhibited a T_m of 61 ^oC and m2BinA isoform showed a T_m of 55 ^oC (Fig. 6.6B). The denaturation profiles were found to be buffer sensitive, which is not unexpected. For instance, in the phosphate buffer recombinant BinA and m2BinA proteins showed T_m of ~65 ^oC, and m1BinA did not show loss of secondary structure till the temperature of about 75 ^oC in the temperature scan at 218 nm.

The influence of PEG conjugation on the stability of tertiary structure of BinA protein was studied using thermofluor shift assay. The thermal denaturation was monitored in the presence of SYPRO orange dye, which is highly fluorescent in non-polar environment. Fluorescence was recorded as a function of temperature (Fig. 6.7A) and the first derivative

was used to estimate apparent melting temperature of the tertiary structure (Fig. 6.7B). The melting temperatures determined for recombinant BinA and PEG-BinA conjugates at pH 7.2 and 8.5, were 55.5, 58.2 and 58.8 ºC, respectively. These values can be treated with high confidence as the observed T_m of 69.5 ºC for HEW lysozyme, used as an internal control in this experiment, matched exactly with the literature value.

The significant increase of about 6 ºC and 3 ºC in the midpoint of denaturation of BinA secondary and tertiary structures after PEGylation, respectively, clearly indicated enhanced thermal stability of BinA protein on PEG conjugation.

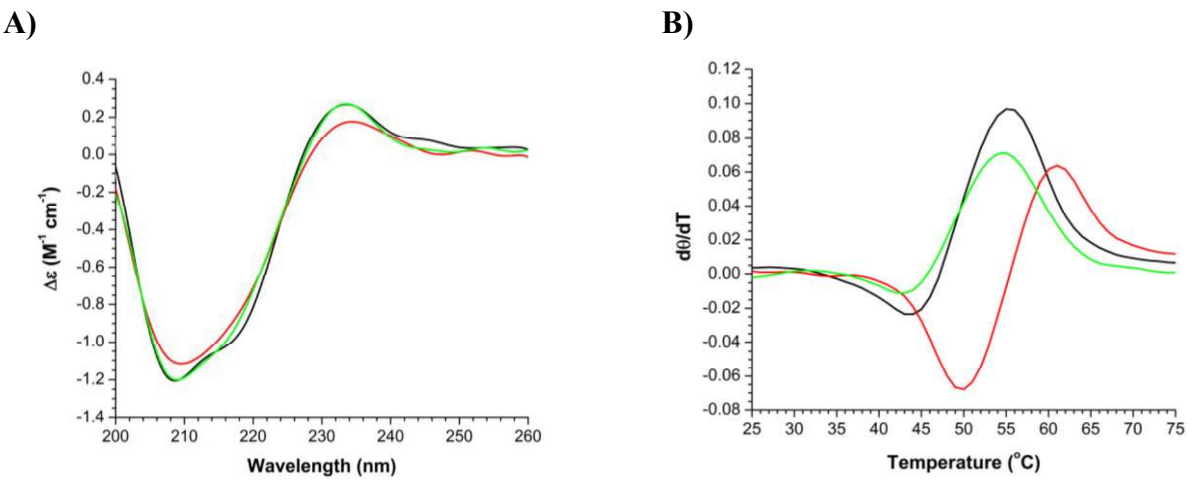


Figure 6.6. Analysis by CD spectroscopy. **A)** CD spectra of BinA (black), m1BinA (red) and m2BinA (green) proteins at 25 ºC. The contribution due to buffer was subtracted and data were smoothed with FFT algorithm with 15 points using Origin® software. **B)** Thermal denaturation of BinA (black) and PEGylated BinA isoforms (m1BinA, red; m2BinA, green) was measured from the first derivative of the observed ellipticity at 218 nm in temperature-wavelength scans.

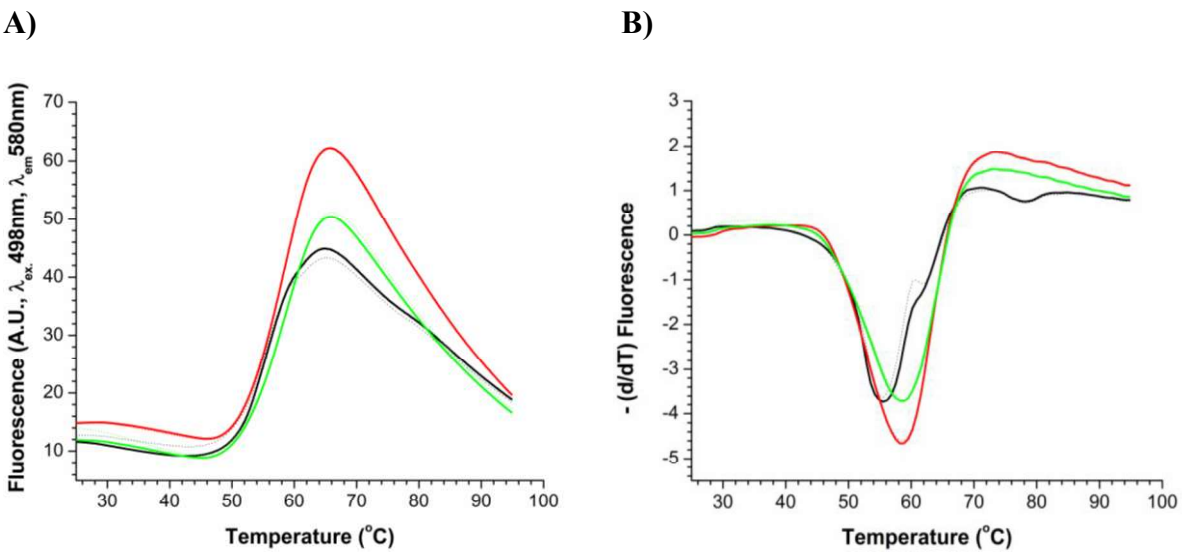


Figure 6.7. Thermal denaturation of the BinA proteins by Thermofluor shift assay. A) Variation in fluorescence intensity of SYPRO orange as a function of temperature was monitored in duplicate in the presence of recombinant BinA (black lines) and PEG conjugated BinA isoforms (m1BinA, red; m2BinA, green). The λ_{ex} and λ_{em} were 498 nm and 580 nm, respectively. **B**) The first derivative of observed fluorescence was plotted as a function of temperature to determine the apparent transition temperature, T_m, which were estimated to be 55.5 ºC, 58.2 ºC and 58.8 ºC for recombinant BinA, m1BinA isoform and m2BinA isoform, respectively.

6.4.7. BinA displays higher stability against proteolytic enzymes

The stability of recombinant BinA, PEGylated BinA derivatives, and pro-BinA/pro-BinB purified from spore-crystals, was assayed in complex proteolytic environments using three proteases (trypsin, chymotrypsin and proteinase K) in independent experiments. The small shift in molecular weight for both pro-BinA and pro-BinB proteins on incubation with trypsin (Fig. 6.8A) is expected from reports in the literature, representing an activation step [27]. Compared to pro-BinA, pro-BinB was found to be susceptible to proteolytic cleavage by proteinase K even with 1 hr incubation, as evident from the decrease in BinB band intensity and appearance of multiple protein bands migrating with higher mobility (Fig. 6.8A, lane 4). The recombinant BinA and PEG-BinA conjugates were found to be stable against trypsin and broad-spectrum proteinase K proteases even upon overnight incubation at protease: protein ratio of as high as 1:10 for trypsin and 1:50 for proteinase K (Fig. 6.8C). All the proteins, however, were found to be susceptible to degradation by chymotrypsin protease upon overnight incubation (Fig. 6.8).

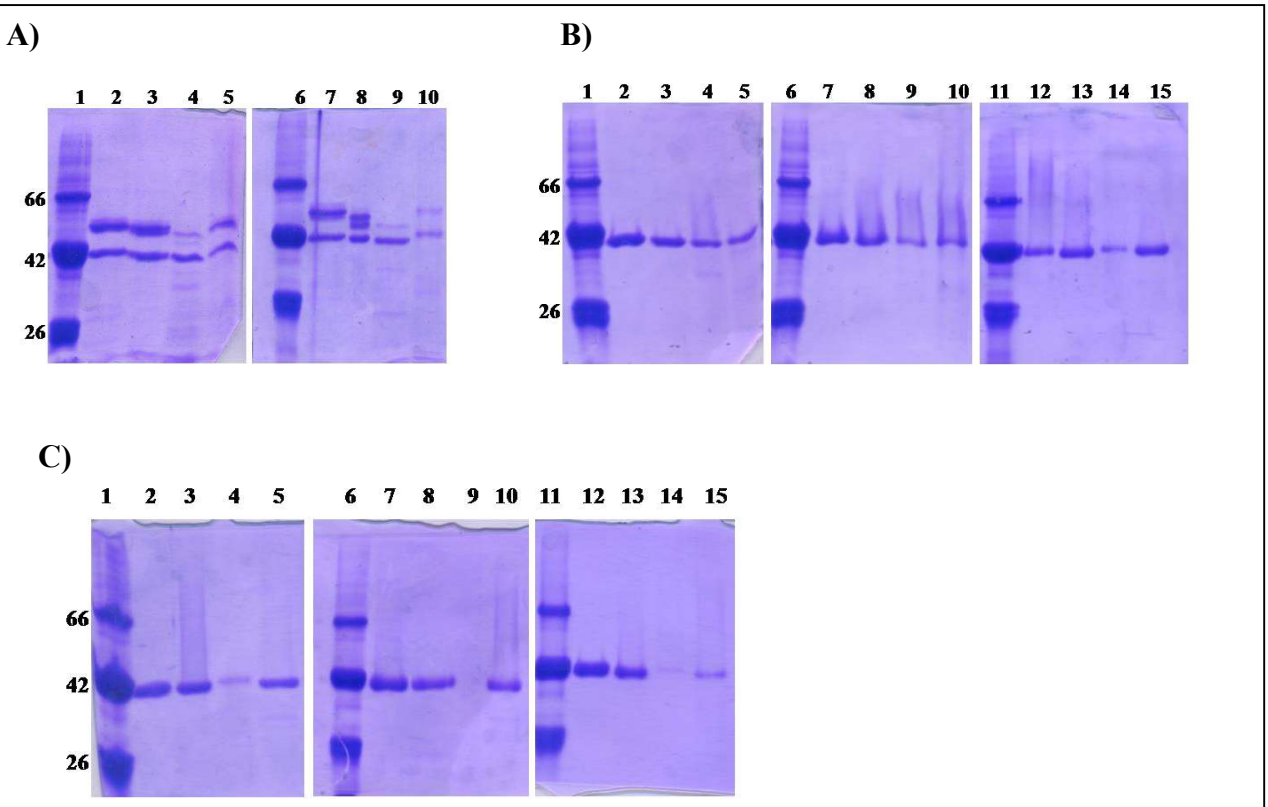


Figure 6.8. Assessment of protein stability against proteolytic enzymes. A) 12% SDS-PAGE analysis of pro-BinA and pro-BinB proteins purified from spore crystals of *L. sphaericus* after incubation with proteases for 1 hr (Lanes 3, 4 and 5) and after over-night incubation with proteases (Lanes 8, 9 and 10). Lanes 1 and 6, molecular weight markers; Lanes 2 and 7, proteins without proteases; Lanes 3 and 8, incubation with trypsin; Lane 4 and 9, incubation with proteinase K; Lanes 5 and 10, incubation with chymotrypsin. **B**) 12% SDS-PAGE analysis of proteolytic stability of recombinant BinA (Lanes 2, 3, 4 and 5), m1BinA isoform (Lanes 7, 8, 9 and 10) and m2BinA isoform (Lanes 12, 13, 14 and 15) after 1 hr incubation with proteases. Lanes 1, 6 and 11, molecular weight markers; Lanes 2, 7 and 12, proteins without proteases; Lanes 3, 8 and 13, incubation with trypsin; Lanes 4, 9 and 14, incubation with chymotrypsin; Lanes 5, 10 and 15, incubation with proteinase K. **C**) 12% SDS-PAGE analysis of proteolytic stability of recombinant BinA (Lanes 2, 3, 4 and 5), m1BinA (Lanes 7, 8, 9 and 10) and m2BinA (Lanes 12, 13, 14 and 15) after overnight incubation with proteases. Lanes 1,6 and 11, molecular weight markers; Lanes 2, 7 and 12, proteins without proteases; Lanes 3, 8 and 13, incubation with trypsin; Lanes 4, 9 and 14, incubation with chymotrypsin; Lanes 5, 10 and 15, incubation with proteinase K.

6.4.8. SPR analysis of PEG-BinA conjugate interaction with glycoprotein

The interaction between PEGylated BinA (m2BinA) and thyroglobulin (glycoprotein) was quantified independently in real-time with SPR from the increase in SPR signal as a function of analyte concentration. The best profile fitting was achieved with the integrated biphasic model of association for the association kinetics of thyroglobulin with m2BinA that describes interactions of two independent phases. The association and dissociation rate constants for thyroglobulin for the first phase (*k*a1 and *k*d1) were estimated to be 8 x10^4 M^{-1} s^{-1} and 1.4 s^{-1}, while the rate constants for the second phase (*k*a2 and *k*d2) were estimated to be 92 M^{-1} s^{-1} and 0.019 s^{-1}, respectively (Fig. 6.9A). The residual plot shows the goodness of fit (Fig. 6.9B).

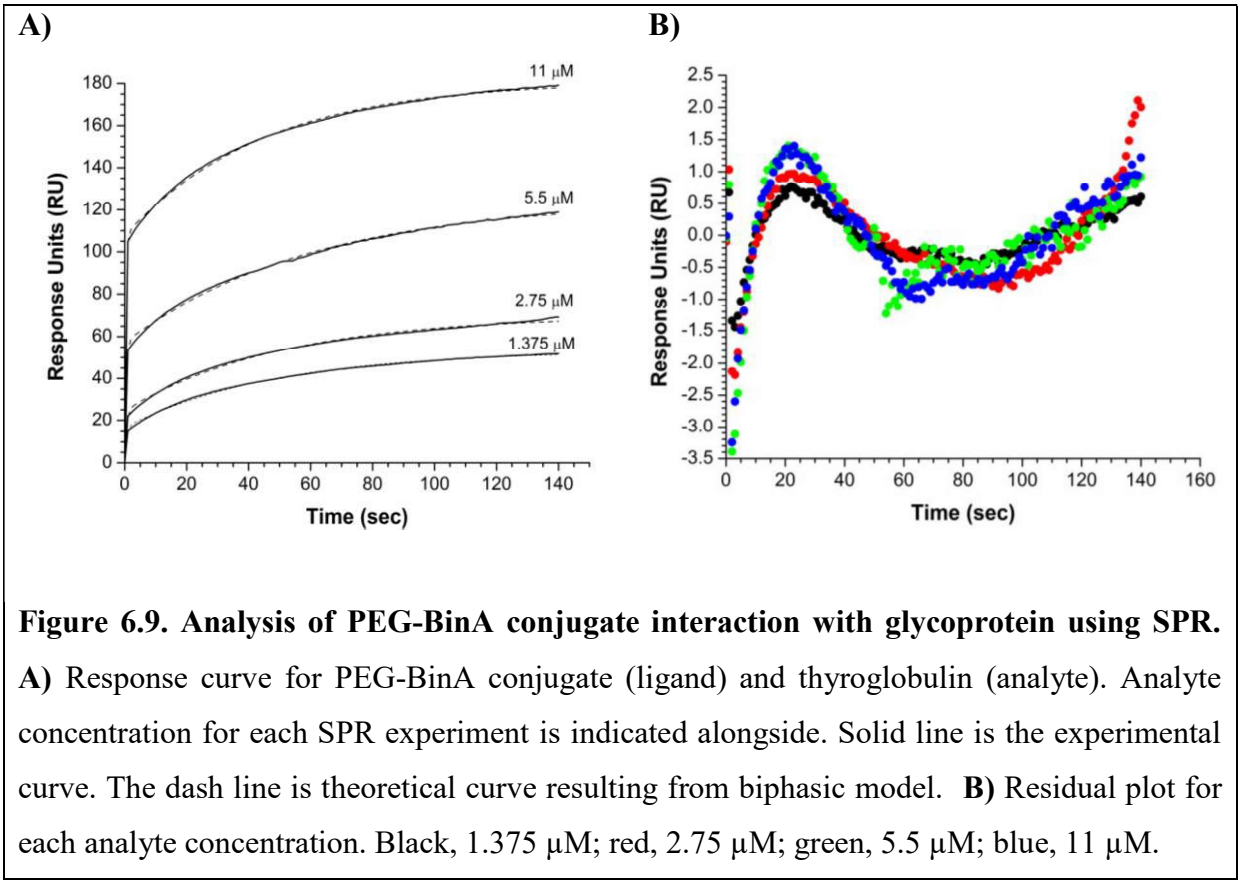


Figure 6.9. Analysis of PEG-BinA conjugate interaction with glycoprotein using SPR. A) Response curve for PEG-BinA conjugate (ligand) and thyroglobulin (analyte). Analyte concentration for each SPR experiment is indicated alongside. Solid line is the experimental curve. The dash line is theoretical curve resulting from biphasic model. **B)** Residual plot for each analyte concentration. Black, 1.375 µM; red, 2.75 µM; green, 5.5 µM; blue, 11 µM.

6.5. DISCUSSION

Mosquitoes are known to spread many human diseases like West Nile fever, dengue, malaria, zika, chikungunya, accounting for millions of deaths annually. *Culex* is known to transmit West Nile virus and Japanese encephalitis (WHO official website., http://www.who.int/en/). Binary toxin from *L. sphaericus* displays mosquito-larvicidal activity against *Culex* and *Anopheles*, and has been found to be safe for a vast range of non-target organism [19]. At LC_{50}, *L. sphaericus* mortality peaks at 48 hr post-treatment due to internalization process [34]. Studies have shown that BinA component of the binary toxin, alone, displays toxicity without BinB, *albeit* at reduced levels [32, 88, 150].

The study presented here is the first reported evaluation of the effect of PEGylation on the activity of BinA larvicidal protein. Recombinant BinA was PEGylated using 750 Da methoxy polyethylene glycol isocyanate as the PEGylation agent. Two PEGylated BinA isoforms, synthesized at pH 7.2 and 8.5 respectively, were purified to homogeneity through size exclusion and anion exchange chromatography. The two-stage column chromatography purification of PEG-BinA conjugates ensured recovery of homogeneous polymer-modified BinA isoforms and complete removal of unreacted excess of PEGylation reagent.

The elution profile from size exclusion chromatography, mobility on native-PAGE, and DLS experiments showed that the PEGylated BinA isoforms exist in the same monomeric state as the unmodified recombinant protein (Fig. 6.1, 6.2 & 6.4). The positive immunostaining with anti-polyethylene glycol and anti-PEG (methoxy group) antibodies, and the observed increased mass in MALDI-TOF experiments confirmed single site modification of recombinant BinA with the PEG moiety (Fig. 6.2B, C & 6.3). Spectroscopic studies further revealed that chemical conjugation did not induce observable conformational change in the protein structure. The CD spectra and deconvolution suggested similarity in the secondary structures of the proteins with nearly 48% residues in β-structure and 2% in α-

helices. The estimated secondary structure matches reasonably with the secondary structure (42% β-structure and 7% α-helices) in the recently resolved X-ray structure of the active region of BinA protein [40].

One of the expected results of PEGylation is the altered thermal stability of target protein. The thermal stability of PEGylated BinA protein was investigated using CD and DSF experiments (Fig. 6.6 & 6.7). An increase of nearly 6 ^{o}C in the denaturation temperature of secondary structure and an increase of nearly 3 ^{o}C in the transition temperature of tertiary structure, upon PEGylation in one of the isoform (m1BinA), clearly indicated the enhancement in the thermal stability of the modified BinA. The other isoform (m2BinA) though did not show enhancement in T_m of the secondary structure, yet its tertiary structure was more stable than the unmodified BinA protein by 3 ^{o}C. Since both the PEG-BinA conjugates appear to be mono-PEGylated, the obtained differences in the T_m values together with changes in DLS profile may suggest differences in site of PEGylation. Also, the observed thermal stability may be independent of site of PEG attachment.

The PEGylated BinA proteins were assayed for mosquito-larvicidal activity against 3rd instar *Culex* larvae. The PEGylated BinA isoforms alone, without BinB, exhibited significantly enhanced median LC_{50} value of ~3.4 ng/mL, compared to LC_{50} values of 21.1 ng/mL for the recombinant (unmodified) BinA protein (Table 6.1). Nearly same lethal concentration was observed for the two isoforms of PEGylated BinA and these compare favourably against the LC_{50} value for BinA/B used world-wide. The LC_{50} values for recombinant BinA differs marginally from the earlier [32] reported toxicity (LC_{50} 16.5 ng/mL), may be due to the variation in mosquito population. The observed increase of nearly 6-fold in BinA mosquito-larvicidal activity is the first evidence of improving protein bio-pesticidal efficacy using PEGylation. The exact mechanism for the increased larvicidal activity of polymer modified BinA is not clear. However, affinity of PEGylated BinA

towards glycoprotein does not change significantly compared to non-PEGylated BinA. The higher larvicidal activity of bioconjugate may be due to enhanced accumulation of active toxin owing to higher stability and bioavailability.

The proteolytic stability analysis for recombinant BinA and PEGylated BinA isoforms, and mixture of pro-BinA and pro-BinB purified from spore-crystals against trypsin, chymotrypsin and proteinase K revealed that BinA is highly stable against trypsin and broad-spectrum proteinase K (Fig. 6.8). BinB, however, was found to be susceptible to degradation by these proteases and can be adjudged to be unstable in complex proteolytic environments. Taken together, these studies suggest that PEGylated BinA alone, without BinB, can be an effective biological control agent. BinA protein can be purified in large scale through recombinant technology, and easily PEGylated. Field trials are required further to adjudge usefulness of this approach in control of mosquito vector diseases.

6.6. CONCLUSIONS

- The chapter discusses completely new approach of increasing the mosquito-larvicidal activity of *L. sphaericus* binary toxin by PEGylating the toxic BinA protein component.
- Recombinant BinA protein was PEGylated using mPEG-ISC -750 as PEGylating reagent.
- The homogenously modified biopolymers displayed preservation of the toxin structure with improved thermal stability.
- The PEGylated protein isoforms show a significant 6-fold enhancement in its activity against *Culex* larvae, which may not depend on the site of PEG attachment.
- The toxicity of PEGylated BinA isoforms matches favourably against the combination of BinA and BinB proteins and can be helpful in reducing dependency on less stable BinB protein to achieve maximal mosquito-larvicidal activity.

Chapter 7

Conclusion & Future Prospects

CONCLUSION

Since the discovery of bacterial strains with the potential for use in mosquito control programs [151], *L. sphaericus* (Ls) has proved to be highly effective and eco-friendly mosquito larvicide [18]. The larvicidal activity of toxic Ls strains is mainly due to BinAB (binary) proteins that exist as a parasporal crystal in the bacterium. The focus of recent studies world-over has been to develop more effective bacterial strains and to elucidate mechanism of larvicidal toxicity of BinAB. This chapter summarizes the main conclusions from the present thesis and highlights some key aspects in understanding the molecular mechanism of action of Ls BinAB toxin.

As Ls spores are used widely for mosquito control, identification of all the components of spores used in mosquito control programs is essential. The present work clearly establishes the identity of the high molecular weight complex (110 and 125 kDa) proteins from the Ls spore crystals, which usually co-purify with BinA/B toxin. These are the products of highly conserved surface layer protein (SlpC) and do not represent oligomeric binary toxin, as thought previously. The purified SlpC protein of ISPC-8 strain showed poor toxicity against the third instar *Culex* larvae. As no homologous *slp*C-like gene is identified in *E. coli* genome, recombinant BinA/B proteins produced homogeneously using *E. coli* based high level protein expression system can be employed for mosquito-control programs.

Interaction of BinA/B with mosquito gut receptor protein (Cqm1) is essential for the larvicidal activity. Prior to the study reported in the present thesis, Cqm1 was purified using SF9 cell lines, which showed weak α-glucosidase activity, and its biological activity was not determined. The present thesis gives a detailed biophysical and biochemical characterization of the recombinant receptor protein from *Culex* purified using *E. coli* based expression system. Recombinant Cqm1 protein exists as a stable dimer which is consistent with its apical localization in lipid rafts. The present thesis reports for the first time that Cqm1 is an

amylomaltase with high glycosyltransferase activity. This property and the observation that Cqm1 can hydrolyze glycogen and sucrose is suggestive of its role in carbohydrate metabolism in mosquitoes. Ca(II) provides thermal stability to the protein and is not essential for catalytic activity. BinB receptor binding component, and not BinA toxic component, reduces catalytic activity of Cqm1 protein. BinB shows high affinity towards Cqm1 (*K*D, 9.8 nM), while BinA interaction with Cqm1 are 1000-fold weaker. Notably, the estimated *K*D values match well with dissociation constants reported earlier with larvae brush border membrane fractions.

Crystal structure analysis revealed presence of lectin-like and aerolysin-like domains in BinA [40]. The BinA protein, however, does not display hemagglutination or hemolytic activity. The present thesis demonstrates for the first time that BinA binds the structurally diverse glycosylated proteins with differential affinity and interacts with simple sugars. The protein interacts with Cqm1 polypeptide non-specifically. The receptor protein for binary toxin is attached to the cell by a GPI anchor. The core of GPI consists of phosphatidylinositol, glycans comprising one glucosamine and at least three mannoses, and a terminal phosphoethanolamine. The carbohydrate composition of glycan core of GPI varies between species and also for different receptors within a species [49]. Presently information about GPI anchor of BinAB receptor is not available. Since BinA displays only mild affinity (*K*D ~ 2 μM) towards Cqm1, simple sugars, as well as for BinB, it is likely that dual interaction of BinA, interaction of lectin-like domain with glycan core of GPI anchor and aerolysin-like pore forming domain with the receptor polypeptide, results in the toxin-receptor specificity in *Culex*. Similar avidity effect has been observed for the *Aeromonas hydrophila* aerolysin toxin [152]. In contrast, BinB binds Cqm1, but not Aam1, with high affinity and its lectin-like domain may not be competent to bind glycans [40, 57]. The strong affinity of BinB towards receptor polypeptide of *Culex* may itself be responsible for highly specific homing of BinB

onto the receptor molecule. Further, the Aam1 receptor of *Aedes aegypti*, and not of *Culex*, is glycosylated [53, 57]. It can be thought that non-specific BinA binding to the glycosylated receptor may block its intracellular translocation resulting in non-toxic accumulation of BinA in the posterior mid gut region of *Aedes aegypti*. Localization of BinA/BinB in the target *Culex* as well as refractory *Aedes* species has been observed in fluorescence studies [58]. A schematic of the suggested events is proposed in Fig. 7.1.

The heteromeric complex of BinA/B may form pores during internalization by a mechanism similar to aerolysin [40, 153]. Once inside the *Culex* cell, BinA causes autophagy, vacuolation and phagocytosis [51]. However, none of these events are thought to cause cell death. It is being suggested in the thesis that cytotoxicity of BinA may be linked to its competence to bind diverse and functionally critical glycosylated proteins. We expect that BinA target proteins may contain fucose/arabinose sugars. Both these sugars inhibit BinA larvicidal activity against *Culex* larvae and were identified by high-throughput protocols in the present thesis. Disruption of protein N-glycosylation through mutations, chemical inhibitors or RNAi, has been found to be fatal for the survival of insects, as N-glycosylation has been reported to be critical for post-embryonic development and metamorphosis in holometabolous insects [154, 155].

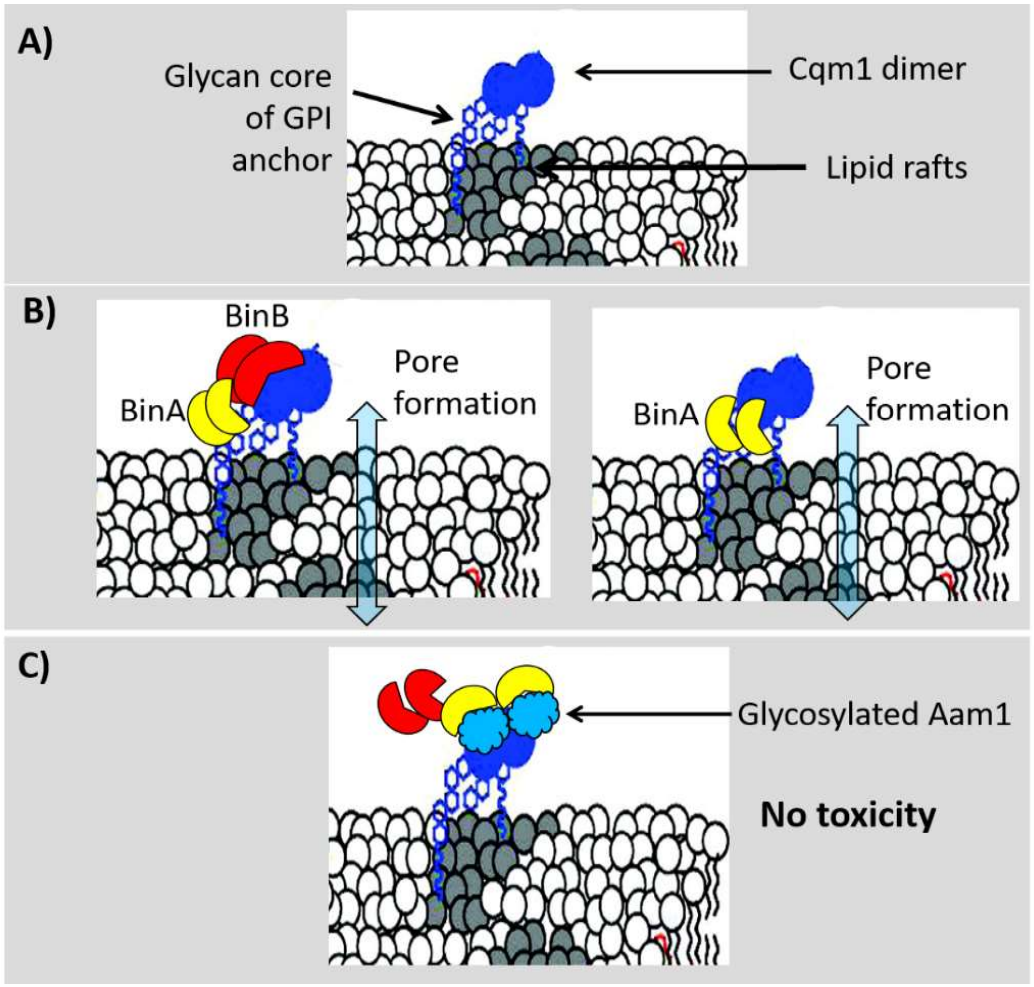


Figure 7.1. A schematic of the events suggested in the present thesis for recognition and accumulation of BinAB protein components on their target receptor in *Culex* (Cqm1) and *Aedes* (Aam1). **A)** Dimeric status of Cqm1 protein supports its association with lipid rafts. **B)** Schematic for the dual mechanism of receptor recognition by BinA in the presence and absence of BinB. The panel highlights the high affinity of BinB for Cqm1 polypeptide. **C)** Proposed reason for refractoriness of *Aedes*.

In view of the important role glycan specificity of BinA may hold, mutations to improve larvicidal activity and broaden specificity of the toxin, may compromise its activity against already known targets. In such a scenario, alternative modes for toxin improvisation, like chemical modification of BinA toxin and/or construction of Ls and Bti chimeric proteins/strains [34] appear to be more favorable approaches.

The present thesis discusses a new approach of increasing the mosquito-larvicidal activity of *L. sphaericus* binary toxin by PEGylating the toxic BinA protein component. PEGylation has proven to be a potent tool for chemical modifications of drugs in real-life applications. PEGylation of BinA results in a rewarding 6-fold increase in its larvicidal activity against *Culex* larvae. The enhanced toxicity may not depend on the site of PEG attachment, and matches favourably against the combination of BinA and BinB proteins. The homogeneously modified biopolymers also display preservation of the toxin structure with improved thermal stability. It can be helpful in reducing dependency on less stable BinB protein to achieve maximal mosquito-larvicidal activity. This is in-line with WHO recommendations on improvement and efficient implementation of mosquito control interventions.

FUTURE PROSPECTS

Many studies have been reported for understanding the cytopathological events involved in toxicity. Apoptosis involving intrinsic pathway has also been suggested as one of the probable events [52]. The study reported enhanced activity of caspase-9 and caspase-3 apoptotic proteins and linked toxicity with mitochondrial response. The present thesis and recent study by Bideshi et al., [34], however, suggests involvement of endoplasmic reticulum, which is the site for N-linked protein glycosylation, and folding and quality control in eukaryotic cells, as the target of binary toxin. Future studies to probe activity of caspase-12 like proteins, which is induced on ER-stress in humans and can also activate caspase-9, are needed to link these events and elucidate the mechanism at the molecular level.

Management of the developing resistance among mosquito population against *L. sphaericus* is necessary to maintain the utility and continued use of this bacterium in mosquito control programs. Development of recombinant toxins is needed to improve the target specificity of such efficient insecticides to combat mosquito borne diseases.

References

1. World Health Organisation. WHO/TDR/Handbook for dengue outbreak contingency planning, 2016.
2. World Health Organisation. WHO Malaria Report, 2017.
3. Muspratt J. Destruction of the larvae of *Anopheles gambiae* giles by Coelomomyces fungus. Bull World Health Organ. 1963; 29:81−86.
4. World Health Organization. Report of the Seventh Scientific Working Group on Biological Control of Vectors. Document TDR/BCV/SWG-7/84.3*.
5. Schnepf E, Crickmore N, Van Rie J, Lereclus D, Baum J, Feitelson J, Zeigler DR, Dean DH. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. 1998; 62(3):775−806.
6. Wirth MC, mosquito resistance to bacterial larvicidal toxins. The Open Toxinology Journal. 2010, 3(1):126−140.
7. Cohen S, Albeck S, Ben-Dov E, Cahan R, Firer M, Zaritsky A, Dym O. Cyt1Aa toxin: crystal structure reveals implications for its membrane-perforating function. J. Mol. Biol. 2011; 413(4):804−814.
8. Bravo A, Gill SS, Soberón M, 2005. *Bacillus thuringiensis* Mechanisms and Use, In: Comprehensive Molecular Insect Science. Elsevier BV, pp. 175–206.
9. Promdonkoy B, Ellar DJ. Investigation of the pore-forming mechanism of a cytolytic delta-endotoxin from *Bacillus thuringiensis*. Biochem. J. 2003; 374(1):255−259.
10. Butko P, Huang F, Pusztai-Carey M, Surewicz WK. Interaction of the delta-endotoxin CytA from *Bacillus thuringiensis* var. *israelensis* with lipid membranes. Biochemistry. 1997; 36(42):12862−12868.
11. Regis L, Silva-Filha MH, Nielsen-LeRoux C, Charles JF. Bacteriological larvicides of dipteran disease vectors. Trends Parasitol. 2001; 17(8):377−380.

12. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera- a visualization system for exploratory research and analysis. J. Comput. Chem. 2004; 25(13):1605−1612.
13. Lacey LA. *Bacillus thuringiensis* serovariety *israelensis* and *Bacillus sphaericus* for mosquito control. J. Am. Mosq. Control Assoc. 2007; 23(2):133−163.
14. Kellen WR, Clark TB, Lindegren JE, Ho BC, Rogoff MH, Singer S. *Bacillus sphaericus* Neide as a pathogen of mosquitoes. J. Invertebr. Pathol. 1965; 7(4):442−448.
15. Russell BL, Jelley SA, Yousten AA. Carbohydrate metabolism in the mosquito pathogen *Bacillus sphaericus* 2362. Appl. Environ. Microbiol. 1989; 55(2):294−297.
16. Ahmed I, Yokota A, Yamazoe A, Fujiwara T. Proposal of *Lysinibacillus boronitolerans* gen. nov. sp. nov., and transfer of *Bacillus fusiformis* to *Lysinibacillus fusiformis* comb. nov. and *Bacillus sphaericus* to *Lysinibacillus sphaericus* comb. nov. Int. J. Syst. Evol. Microbiol. 2007; 57(5):1117−1125.
17. Myers P, Yousten AA. Toxic activity of *Bacillus sphaericus* SSII-1 for mosquito larvae. Infect Immun. 1978; 19(3):1047−1053.
18. World Health Organization, 1985. Informal consultation on the development of *Bacillus sphaericus* as a microbial larvicide. TDR/BCV/Sphaericus/85.3. World Health Organization, Geneva.
19. Lacey LA, Mulla MS, 1990. Safety of *Bacillus thuringiensis* (H-14) and *Bacillus sphaericus* to nontarget organisms in the aquatic environment, In: M. Laird, L.A. Lacey, E.W. Davidson, eds. Safety of Microbial Insecticides. Boca Raton, FL: CRC Press, pp. 169–188.
20. Lacey LA, Siegel JP, 2000. Safety and ecotoxicology of entomopathogenic bacteria. In: J.F. Charles, A. Delecluse, C. Nielsen-LeRoux, eds. Entomopathogenic Bacteria: From

laboratory to field application. Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 253–273.

21. Brown ID, Watson TM, Carter J, Purdie DM, Kay BH. Toxicity of VectoLex (*Bacillus sphaericus*) products to selected Australian mosquito and non-target species. J. Econ. Entomol. 2004; 97(1):51−58.
22. Nishiwaki H, Nakashima K, Ishida C, Kawamura T, Matsuda K. Cloning, functional characterization, and mode of action of a novel insecticidal pore-forming toxin, sphaericolysin, produced by *Bacillus sphaericus*. Appl. Environ. Microbiol. 2007; 73(10):3404−3411.
23. Jones GW, Nielsen-Leroux C, Yang Y, Yuan Z, Dumas VF, Monnerat RG, Berry C. A new Cry toxin with a unique two-component dependency from *Bacillus sphaericus*. FASEB J. 2007; 21(14):4112−4120.
24. Jones GW, Wirth MC, Monnerat RG, Berry C. The Cry48Aa-Cry49Aa binary toxin from *Bacillus sphaericus* exhibits highly restricted target specificity. Environ. Microbiol. 2008; 10(9):2418−2424.
25. Baumann P, Unterman BM, Baumann L, Broadwell AH, Abbene SJ, Bowditch RD. Purification of the larvicidal toxin of *Bacillus sphaericus* and evidence for high-molecular-weight precursors. J. Bacteriol. 1985; 163(2):738−747.
26. Baumann P, Clark MA, Baumann L, Broadwell AH. *Bacillus sphaericus* as a mosquito pathogen: properties of the organism and its toxins. Microbiol Rev. 1991; 55(3):425−436.
27. Broadwell AH, Baumann P. Proteolysis in the gut of mosquito larvae results in further activation of the *Bacillus sphaericus* toxin. Appl Environ Microbiol. 1987; 53(6):1333−1337.

28. Silva-Filha MH, Nielsen-LeRoux C, Charles JF. Identification of the receptor for *Bacillus sphaericus* crystal toxin in the brush border membrane of the mosquito *Culex pipiens* (Diptera: Culicidae). Insect Biochem. Mol. Biol. 1999; 29(8):711−721.
29. Darboux I, Nielsen-LeRoux C, Charles JF, Pauron D. The receptor of *Bacillus sphaericus* binary toxin in *Culex pipiens* (Diptera: Culicidae) midgut: molecular cloning and expression. Insect Biochem. Mol. Biol. 2001; 31(10):981−990.
30. Broadwell AH, Baumann L, Baumann P. The 42- and 51-kilodalton mosquitocidal proteins of *Bacillus sphaericus* 2362: construction of recombinants with enhanced expression and in vivo studies of processing and toxicity. J. Bacteriol. 1990; 172(5):2217−2223.
31. Oei C, Hindley J, Berry C. An analysis of the genes encoding the 51.4- and 41.9-kDa toxins of *Bacillus sphaericus* 2297 by deletion mutagenesis: the construction of fusion proteins. FEMS Microbiol. Lett. 1990; 60(3):265−273.
32. Kale A, Hire RS, Hadapad AB, D'Souza SF, Kumar V. Interaction between mosquito-larvicidal *Lysinibacillus sphaericus* binary toxin components: analysis of complex formation. Insect Biochem. Mol. Biol. 2013; 43(11):1045−1054.
33. Sharma M, Hire RS, Hadapad AB, Gupta GD, Kumar V. PEGylation Enhances Mosquito-Larvicidal Activity of *Lysinibacillus sphaericus* Binary Toxin. Bioconjug. Chem. 2017 15; 28(2):410−418.
34. Bideshi DK, Park HW, Hice RH, Wirth MC, Federici BA. Highly effective broad spectrum chimeric larvicide that targets vector mosquitoes using a lipophilic protein. Sci. Rep. 2017; 7(1):11282.
35. Cokmus C, Davidson EW, Cooper K. Electrophysiological effects of *Bacillus sphaericus* binary toxin on cultured mosquito cells. J. Invertebr. Pathol. 1997; 69(3):197−204.

36. Schwartz JL, Potvin L, Coux F, Charles JF, Berry C, Humphreys MJ, Jones AF, Bernhart I, Dalla Serra M, Menestrina G. Permeabilization of model lipid membranes by *Bacillus sphaericus* mosquitocidal binary toxin and its individual components. J. Membr. Biol. 2001; 184(2):171−83.
37. Smith AW, Cámara-Artigas A, Brune DC, Allen JP. Implications of high-molecular-weight oligomers of the binary toxin from *Bacillus sphaericus*. J. Invertebr. Pathol. 2005; 88(1):27−33.
38. Pei J, Kim BH, Grishin NV. PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res. 2008; 36(7):2295−2300.
39. The UniProt Consortium. UniProt: the universal protein knowledgebase**.** Nucleic Acids Res. 2017; 45: D158−D169.
40. Colletier JP, Sawaya MR, Gingery M, Rodriguez JA, Cascio D, Brewster AS, Michels-Clark T, Hice RH, Coquelle N, Boutet S, Williams GJ, Messerschmidt M, DePonte DP, Sierra RG, Laksmono H, Koglin JE, Hunter MS, Park HW, Uervirojnangkoorn M, Bideshi DK, Brunger AT, Federici BA, Sauter NK, Eisenberg DS. De novo phasing with X-ray laser reveals mosquito larvicide BinAB structure. Nature. 2016; 539(7627):43−47.
41. Srisucharitpanit K, Yao M, Promdonkoy B, Chimnaronk S, Tanaka I, Boonserm P. Crystal structure of BinB: a receptor binding component of the binary toxin from *Lysinibacillus sphaericus.* Proteins. 2014; 82(10):2703−2712.
42. Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. 2004; 382(3):769−781.
43. Murzin AG, Lesk AM, Chothia C. β-Trefoil fold. Patterns of structure and sequence in the Kunitz inhibitors interleukins-1 beta and 1 alpha and fibroblast growth factors. J. Mol. Biol. 1992; 223(2):531−543.

44. Baenziger JU, Fiete D. Structural determinants of *Ricinus communis* agglutinin and toxin specificity for oligosaccharides. J. Biol. Chem. 1979; 254 (19): 9795–9799.

45. Rosado CJ, Kondos S, Bull TE, Kuiper MJ, Law RH, Buckle AM, Voskoboinik I, Bird PI, Trapani JA, Whisstock JC, Dunstone MA. The MACPF/CDC family of pore-forming toxins. Cell Microbiol. 2008; 10(9):1765−1774.

46. Diep DB, Nelson KL, Raja SM, Pleshak EN, Buckley JT. Glycosylphosphatidylinositol anchors of membrane glycoproteins are binding determinants for the channel-forming toxin aerolysin. J. Biol. Chem. 1998; 273(4):2355−2360.

47. Kelker MS, Berry C, Evans SL, Pai R, McCaskill DG, Wang NX, Russell JC, Baker MD, Yang C, Pflugrath JW, Wade M, Wess TJ, Narva KE. Structural and biophysical characterization of *Bacillus thuringiensis* insecticidal proteins Cry34Ab1 and Cry35Ab1. PLoS One. 2014; 9(11):e112555.

48. Szczesny P, Iacovache I, Muszewska A, Ginalski K, van der Goot FG, Grynberg M. Extending the aerolysin family: from bacteria to vertebrates. PLoS One. 2011; 6(6):e20349.

49. Paulick MG, Bertozzi CR. The Glycosylphosphatidylinositol Anchor: A complex membrane-anchoring structure for proteins. Biochemistry. 2008; 47(27): 6991–7000.

50. Kinoshita T. Glycosylphosphatidylinositol (GPI) Anchors: Biochemistry and Cell Biology: Introduction to a Thematic Review Series. J. Lipid Res. 2016; 57(1):4−5.

51. Opota O, Gauthier NC, Doye A, Berry C, Gounon P, Lemichez E, Pauron D. *Bacillus sphaericus* binary toxin elicits host cell autophagy as a response to intoxication. PLoS One. 2011; 6(2):e14682.

52. Tangsongcharoen C, Chomanee N, Promdonkoy B, Boonserm P. *Lysinibacillus sphaericus* binary toxin induces apoptosis in susceptible *Culex quinquefasciatus* larvae. J. Invertebr. Pathol. 2015; 128:57−63.

53. Ferreira LM, Romão TP, de-Melo-Neto OP, Silva-Filha MH. The orthologue to the Cpm1/Cqm1 receptor in *Aedes aegypti* is expressed as a midgut GPI-anchored α-glucosidase, which does not bind to the insecticidal binary toxin. Insect Biochem. Mol. Biol. 2010; 40(8):604−610.
54. Opota O, Charles JF, Warot S, Pauron D, Darboux I. Identification and characterization of the receptor for the *Bacillus sphaericus* binary toxin in the malaria vector mosquito, *Anopheles gambiae*. Comp Biochem. Physiol. B. Biochem. Mol. Biol. 2008; 149(3):419−427.
55. Silva-Filha MH, Nielsen-Leroux C, Charles JF. Binding kinetics of *Bacillus sphaericus* binary toxin to midgut brush-border membranes of *Anopheles* and *Culex* sp. mosquito larvae. Eur. J. Biochem. 1997; 247(3):754−761.
56. Nielsen-Leroux C, Charles JF. Binding of *Bacillus sphaericus* binary toxin to a specific receptor on midgut brush-border membranes from mosquito larvae. Eur. J. Biochem. 1992; 210(2):585−590.
57. Ferreira LM, Romão TP, Nascimento NA, Costa Mda C, Rezende AM, de-Melo-Neto OP, Silva-Filha MH. Non-conserved residues between Cqm1 and Aam1 mosquito α-glucosidases are critical for the capacity of Cqm1 to bind the Binary toxin from *Lysinibacillus sphaericus*. Insect Biochem. Mol. Biol. 2014; 50:34−42.
58. Lekakarn H, Promdonkoy B, Boonserm P. Interaction of *Lysinibacillus sphaericus* binary toxin with mosquito larval gut cells: Binding and internalization. J. Invertebr. Pathol. 2015; 132:125−31.
59. Darboux I, Pauchet Y, Castella C, Silva-Filha MH, Nielsen-LeRoux C, Charles JF, Pauron D. Loss of the membrane anchor of the target receptor is a mechanism of bioinsecticide resistance. Proc. Natl. Acad. Sci. U S A. 2002; 99(9):5830−5835.

60. Darboux I, Charles JF, Pauchet Y, Warot S, Pauron D. Transposon-mediated resistance to *Bacillus sphaericus* in a field-evolved population of *Culex* pipiens (Diptera: Culicidae). Cell Microbiol. 2007; 9(8):2022−2029.

61. Romão TP, de Melo Chalegre KD, Key S, Ayres CF, Fontes de Oliveira CM, de-Melo-Neto OP, Silva-Filha MH. A second independent resistance mechanism to *Bacillus sphaericus* binary toxin targets its alpha-glucosidase receptor in *Culex quinquefasciatus*. FEBS J. 2006; 273(7):1556−1568.

62. Guo QY, Cai QX, Yan JP, Hu XM, Zheng DS, Yuan ZM. Single nucleotide deletion of cqm1 gene results in the development of resistance to *Bacillus sphaericus* in *Culex quinquefasciatus*. J. Insect. Physiol. 2013; 59(9):967−973.

63. Georghiou GP, Wirth MC. Influence of Exposure to Single versus Multiple Toxins of *Bacillus thuringiensis* subsp. *israelensis* on Development of Resistance in the Mosquito *Culex quinquefasciatus* (Diptera: Culicidae). Appl. Environ. Microbiol. 1997; 63(3):1095−1101.

64. Wirth MC, Georghiou GP, Federici BA. CytA enables CryIV endotoxins of *Bacillus thuringiensis* to overcome high levels of CryIV resistance in the mosquito, *Culex quinquefasciatus*. Proc. Natl. Acad. Sci. U S A. 1997; 94(20):10536−10540.

65. Park HW, Bideshi DK, Federici BA. Recombinant strain of *Bacillus thuringiensis* producing Cyt1A, Cry11B, and the *Bacillus sphaericus* binary toxin. Appl. Environ. Microbiol. 2003; 69(2):1331−1334.

66. Park HW, Bideshi DK, Wirth MC, Johnson JJ, Walton WE, Federici BA. Recombinant larvicidal bacteria with markedly improved efficacy against *Culex* vectors of west nile virus. Am. J. Trop. Med. Hyg. 2005; 72(6):732−738.

67. Abuchowski A, van Es T, Palczuk NC, Davis FF. Alteration of immunological properties of bovine serum albumin by covalent attachment of polyethylene glycol. J. Biol. Chem. 1977; 252(11):3578−3581.

68. Abuchowski A, McCoy JR, Palczuk NC, van Es T, Davis FF. Effect of covalent attachment of polyethylene glycol on immunogenicity and circulating life of bovine liver catalase. J. Biol. Chem. 1977; 252(11):3582−3586.

69. Harris JM, Martin NE, Modi M. Pegylation: a novel process for modifying pharmacokinetics. Clin. Pharmacokinet. 2001; 40(7):539−551.

70. Roberts MJ, Bentley MD, Harris JM. Chemistry for peptide and protein PEGylation. Adv. Drug Deliv. Rev. 2002; 54(4):459−476.

71. Veronese FM, Pasut G. Improvement of Drug Therapy by Covalent PEG Conjugation: An Overview From a Research Laboratory. Israel Journal of Chemistry, 2010; 50:151−159.

72. Caliceti P, Veronese FM. Pharmacokinetic and biodistribution properties of poly(ethylene glycol)-protein conjugates. Adv. Drug Deliv. Rev. 2003; 55(10):1261−1277.

73. Alconcel SNS, Baas AS and Maynard HD. FDA-approved poly (ethylene glycol)-protein conjugate drugs. Polym. Chem. 2011, **2**:1442–1448.

74. Sambrook J, Russell RW, 2001. Molecular cloning: A laboratory manual, 3rd ed. Cold spring harbor laboratory press, cold spring harbor, New York.

75. Hadapad AB, Hire RS, Vijayalakshmi N, Dongre TK. UV protectants for the biopesticide based on *Bacillus sphaericus* Neide and their role in protecting the binary toxins from UV radiation. J. Invertebr. Pathol. 2009; 100(3):147−152.

76. Savitsky P, Bray J, Cooper CDO, et al. High-throughput production of human proteins for crystallization: The SGC experience. Journal of Structural Biology. 2010; 172(1):3−13.

77. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A, 2005. Protein Identification and Analysis Tools on the ExPASy Server, In: Walker JM, (Ed) The Proteomics Protocols Handbook. Humana Press Inc, Totowa, NJ. pp 571−607.
78. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 1976; 72:248−254.
79. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the Folin phenol reagent. J. Biol. Chem. 1951; 193(1):265−275.
80. Perez-Iratxeta C, Andrade-Navarro MA. K2D2: estimation of protein secondary structure from circular dichroism spectra. BMC Struct. Biol. 2008; 8:25.
81. Niesen FH, Berglund H, Vedadi M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat. Protoc. 2007; 2(9):2212−2221.
82. Brogioni B, Berti F. Surface plasmon resonance for the characterization of bacterial polysaccharide antigens: a review. Med. Chem. Comm. 2014; 5:1058−1066.
83. Kumar S, Eremina N, Barth A. Detection of ligand binding to proteins through observation of hydration water. J. Phys. Chem. B. 2012; 116(48):13968−13974.
84. Hadapad AB, Vijayalakshmi N, Hire RS, Dongre TK. Effect of ultraviolet radiation on spore viability and mosquitocidal activity of an indigenous ISPC-8 *Bacillus sphaericus* Neide strain. Acta Trop. 2008; 107(2):113−116.
85. Abbott WS. A method of computing the effectiveness of an insecticide. 1925. J. Am. Mosq. Control Assoc. 1987; 3(2):302−303.
86. Finney DJ, 1952. Probit Analysis, 2nd ed. Cambridge University Press, New York, 1952. 22.5 × 14 cm. xiv + 318 pp. J. Pharm. Sci., 41: 627.
87. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.r-project.org/.

88. Hire RS, Hadapad AB, Vijayalakshmi N, Dongre TK. Characterization of highly toxic indigenous strains of mosquitocidal organism *Bacillus sphaericus*. FEMS Microbiol. Lett. 2010; 305(2):155−161.

89. Bowditch RD, Baumann P, Yousten AA. Cloning and sequencing of the gene encoding a 125-kilodalton surface-layer protein from *Bacillus sphaericus* 2362 and of a related cryptic gene. J. Bacteriol. 1989; 171(8):4178−4188.

90. Sangthongpitag K, Penfold RJ, Delaney SF, Rogers PL. Cloning and expression of the *Bacillus sphaericus* 2362 mosquitocidal genes in a non-toxic unicellular cyanobacterium, Synechococcus PCC6301. Appl. Microbiol. Biotechnol. 1997; 47(4):379−384.

91. Yuan Z, Rang C, Maroun RC, Juárez-Pérez V, Frutos R, Pasteur N, Vendrely C, Charles JF, Nielsen-Leroux C. Identification and molecular structural prediction analysis of a toxicity determinant in the *Bacillus sphaericus* crystal larvicidal toxin. Eur. J. Biochem. 2001; 268(9):2751−2760.

92. Shanmugavelu M, Rajamohan F, Kathirvel M, Elangovan G, Dean DH, Jayaraman K. Functional complementation of nontoxic mutant binary toxins of *Bacillus sphaericus* 1593M generated by site-directed mutagenesis. Appl. Environ. Microbiol. 1998; 64(2):756−759.

93. Broadwell AH, Baumann P. Sporulation-associated activation of *Bacillus sphaericus* larvicide. Appl. Environ. Microbiol. 1986; 52(4):758−764.

94. Hu X, Li J, Hansen BM, Yuan Z. Phylogenetic analysis and heterologous expression of surface layer protein SlpC of *Bacillus sphaericus* C3-41. Biosci. Biotechnol. Biochem. 2008; 72(5):1257−1263.

95. Lozano LC, Ayala JA, Dussán J. *Lysinibacillus sphaericus* S-layer protein toxicity against *Culex quinquefasciatus*. Biotechnol. Lett. 2011; 33(10):2037−2041.

96. Lozano LC, Dussán J. Synergistic Activity Between S-Layer Protein and Spore-Crystal Preparations from *Lysinibacillus sphaericus* Against *Culex quinquefasciatus* Larvae. Curr. Microbiol. 2017; 74(3):371−376.

97. Nascimento NA, Ferreira LM, Romão TP, Correia DM, Vasconcelos CR, Rezende AM, Costa SG, Genta FA, de-Melo-Neto OP, Silva-Filha MH. N-glycosylation influences the catalytic activity of mosquito α-glucosidases associated with susceptibility or refractoriness to *Lysinibacillus sphaericus*. Insect Biochem. Mol. Biol. 2017; 81:62−71.

98. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014; 42,490−495.

99. Stam MR, Danchin EG, Rancurel C, Coutinho PM, Henrissat B. Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins. Protein Eng. Des. Sel. 2006; 19(12):555−562.

100. Chiba S, 1988. α-Glucosidases, In: Handbook of amylases and related enzymes. The Amylase Research Society of Japan, Pergamon Press, Oxford, United Kingdom. pp 104−105.

101. Kato N, Suyama S, Shirokane M, Kato M, Kobayashi T, Tsukagoshi N. Novel alpha-glucosidase from *Aspergillus nidulans* with strong transglycosylation activity. Appl. Environ. Microbiol. 2002; 68(3):1250−1256.

102. Chalegre KD, Tavares DA, Romão TP, de Menezes HS, Nascimento NA, de Oliveira CM, de-Melo-Neto OP, Silva-Filha MH. Co-selection and replacement of resistance alleles to *Lysinibacillus sphaericus* in a *Culex quinquefasciatus* colony. FEBS J. 2015; 282(18):3592−3602.

103. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. 2006; 1(6):2856−2860.

104. Gattiker A, Bienvenut WV, Bairoch A, Gasteiger E. FindPept, a tool to identify unmatched masses in peptide mass fingerprinting protein identification. Proteomics. 2002; 2(10):1435−1444.

105. Eswar N, Eramian D, Webb B, Shen MY, Sali A. Protein structure modelling with MODELLER. Methods Mol. Biol. 2008; 426:145−159.

106. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Gallo Cassarino T, Bertoni M, Bordoli L, Schwede T. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. 2014; 42:252−258.

107. Ravaud S, Robert X, Watzlawick H, Haser R, Mattes R, Aghajari N. Trehalulose synthase native and carbohydrate complexed structures provide insights into sucrose isomerization. J. Biol. Chem. 2007; 282(38):28126−28136.

108. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, Hess B, Lindahl E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics. 2013; 29(7):845−854.

109. Jones DT, Cozzetto D. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. Bioinformatics. 2015; 31(6):857−863.

110. Kobayashi M, Hondoh H, Mori H, Saburi W, Okuyama M, Kimura A. Calcium ion-dependent increase in thermostability of dextran glucosidase from *Streptococcus* mutants. Biosci. Biotechnol. Biochem. 2011; 75(8):1557−1563.

111. Paladino S, Sarnataro D, Pillich R, Tivodar S, Nitsch L, Zurzolo C. Protein oligomerization modulates raft partitioning and apical sorting of GPI-anchored proteins. J. Cell. Biol. 2004; 167(4):699−709.

112. Pauchet Y, Luton F, Castella C, Charles JF, Romey G, Pauron D. Effects of a mosquitocidal toxin on a mammalian epithelial cell line expressing its target receptor. Cell Microbiol. 2005; 7(9):1335−1344.

113. Charles JF, Silva-Filha MH, Nielsen-LeRoux C, Humphreys MJ, Berry C. Binding of the 51- and 42-kDa individual components from the *Bacillus sphaericus* crystal toxin to mosquito larval midgut membranes from *Culex* and *Anopheles* sp. (Diptera: Culicidae). FEMS Microbiol. Lett. 1997; 156(1):153−159.

114. Deng X, Petitjean M, Teste MA, Kooli W, Tranier S, François JM, Parrou JL. Similarities and differences in the biochemical and enzymological properties of the four isomaltases from *Saccharomyces cerevisiae*. FEBS Open Bio. 2014; 4:200−212.

115. Linser PJ, Smith KE, Seron TJ, Neira Oviedo M. Carbonic anhydrases and anion transport in mosquito midgut pH regulation. J. Exp. Biol. 2009; 212(11):1662−1671.

116. Dadd, RH. Alkalinity within the midgut of mosquito larvae with alkaline-active digestive enzymes. J. Insect. Physiol. 1975; 21:1847–53.

117. Shen X, Saburi W, Gai Z, Kato K, Ojima-Kato T, Yu J, Komoda K, Kido Y, Matsui H, Mori H, Yao M. Structural analysis of the α-glucosidase HaG provides new insights into substrate specificity and catalytic mechanism. Acta Crystallogr. D Biol. Crystallogr. 2015; 71(6):1382−1391.

118. Price DR, Karley AJ, Ashford DA, Isaacs HV, Pownall ME, Wilkinson HS, Gatehouse JA, Douglas AE. Molecular characterisation of a candidate gut sucrase in the pea aphid, *Acyrthosiphon pisum*. Insect Biochem. Mol. Biol. 2007; 37(4):307−317.

119. Wongchawalit J, Yamamoto T, Nakai H, Kim YM, Sato N, Nishimoto M, Okuyama M, Mori H, Saji O, Chanchao C, Wongsiri S, Surarit R, Svasti J, Chiba S, Kimura A. Purification and characterization of alpha-glucosidase I from Japanese honeybee (*Apiscerana japonica*) and molecular cloning of its cDNA. Biosci. Biotechnol. Biochem. 2006; 70(12):2889−2898.

120. Ashford DA, Smith WA, Douglas AE. Living on a high sugar diet: the fate of sucrose ingested by a phloem-feeding insect, the pea aphid *Acyrthosiphon pisum*. J. Insect Physiol. 2000; 46(3):335−341.

121. Karley AJ, Ashford DA, Minto LM, Pritchard J, Douglas AE. The significance of gut sucrase activity for osmoregulation in the pea aphid, *Acyrthosiphon pisum*. J. Insect Physiol. 2005; 51(12):1313−1319.

122. Berry C. The bacterium, *Lysinibacillus sphaericus*, as an insect pathogen. J. Invertebr. Pathol. 2012; 109(1):1−10.

123. Davidson EW. Binding of the *Bacillus sphaericus* (Eubacteriales: Bacillaceae) toxin to midgut cells of mosquito (Diptera: Culicidae) larvae: relationship to host range. J Med Entomol. 1988; 25(3):151−157.

124. Hart GW. Thematic minireview series on glycobiology and extracellular matrices: glycan functions pervade biology at all levels. J. Biol. Chem. 2013; 288(10):6903.

125. Varki A. Biological roles of glycans. Glycobiology. 2017; 27(1):3−49.

126. Holdgate GA, Anderson M, Edfeldt F, Geschwindner S. Affinity-based, biophysical methods to detect and analyze ligand binding to recombinant proteins: matching high information content with high throughput. J. Struct. Biol. 2010; 172(1):142−157.

127. Ciulli A. Biophysical screening for the discovery of small-molecule ligands. Methods Mol. Biol. 2013; 1008:357−388.

128. Shangary S, Singh J, Kamboj SS, Kamboj KK, Sandhu RS. Purification and properties of four monocot lectins from the family Araceae. Phytochemistry. 1995; 40(2):449−455.

129. Wang R, Sun S, Bekos EJ, Bright FV. Dynamics surrounding Cys-34 in native, chemically denatured, and silica-adsorbed bovine serum albumin. Anal. Chem. 1995; 67(1):149−159.

130. Schellman JA. Temperature, stability, and the hydrophobic interaction. Biophys. J. 1997; 73(6):2960−2964.

131. Matulis D, Kranz JK, Salemme FR, Todd MJ. Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. Biochemistry. 2005; 44(13):5258−5266.

132. Vedadi M, Niesen FH, Allali-Hassani A, Fedorov OY, Finerty PJ Jr, WasneyGA,Yeung R, Arrowsmith C, Ball LJ, Berglund H, Hui R, Marsden BD, Nordlund P, Sundstrom M, Weigelt J, Edwards AM. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. Proc. Natl. Acad. Sci. USA. 2006; 103(43):15835−15840.

133. Gupta D, Dam TK, Oscarson S, Brewer CF. Thermodynamics of lectin-carbohydrate interactions. Binding of the core trimannoside of asparagine-linked carbohydrates and deoxy analogs to concanavalin A. J. Biol. Chem. 1997; 272(10):6388−6392.

134. Tschampel SM, Woods RJ. Quantifying the role of water in protein-carbohydrate interactions. J. Phys. Chem A. 2003; 107(43):9175−9181.

135. Baenziger JU, Fiete D. Structure of the complex oligosaccharides of fetuin. J. Biol. Chem. 1979; 254(3):789−795.

136. Yet MG, Chin CC, Wold F. The covalent structure of individual N-linked glycopeptides from ovomucoid and asialofetuin. J. Biol. Chem. 1988; 263(1):111−117.

137. Tsuji T, Yamamoto K, Irimura T, Osawa T. Structure of carbohydrate unit A of porcine thyroglobulin. Biochem. J. 1981; 195(3):691−699.

138. Yamamoto K, Tsuji T, Irimura T, Osawa T. The structure of carbohydrate unit B of porcine thyroglobulin. Biochem. J. 1981; 195(3):701−713.

139. Federici BA, Park HW, Bideshi DK, Wirth MC, Johnson JJ, Sakano Y, Tang M. Developing recombinant bacteria for control of mosquito larvae. J. Am. Mosq. Control Assoc. 2007; 23(2):164−175.

140. Shen H, Brandt A, Bissinger BW, Gunnoe B, Michael Roe R. Novel insecticide polymer chemistry to reduce the enzymatic digestion of a protein pesticide, trypsin modulating oostatic factor (TMOF). Pesticide Biochemistry and Physiology, 2009; 93:144–152.

141. Boccù E, Velo GP, Veronese FM. Pharmacokinetic properties of polyethyleneglycol derivatized superoxide dismutase. Pharmacol Res Commun. 1982; 14(2):113−120.

142. Harris JM, Chess RB. Effect of pegylation on pharmaceuticals. Nat. Rev. Drug Discov. 2003; 2(3):214−221.

143. An Q, Lei Y, Jia N, Zhang X, Bai Y, Yi J, Chen R, Xia A, Yang J, Wei S, Cheng X, Fan A, Mu S, Xu Z. Effect of site-directed PEGylation of trichosanthin on its biological activity, immunogenicity, and pharmacokinetics. Biomol. Eng. 2007; 24(6):643−649.

144. Jeffers LA, Shen H, Khalil S, Bissinger BW, Brandt A, Gunnoe TB, Roe RM. Enhanced activity of an insecticidal protein, trypsin modulating oostatic factor (TMOF), through conjugation with aliphatic polyethylene glycol. Pest Manag. Sci. 2012; 68(1):49−59.

145. Kurfürst MM. Detection and molecular weight determination of polyethyleneglycol-modified hirudin by staining after sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Anal. Biochem. 1992; 200(2):244−248.

146. Skoog B. A method for qualitative and/or quantitative detection of polyethylene glycols in biological fluids. VOX Sang. 1979; 37:345–349.

147. Greenfield NJ. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nat. Protoc. 2006; 1(6):2527−2535.

148. Grimsley GR, Scholtz JM, Pace CN. A summary of the measured pK values of the ionizable groups in folded proteins. Protein Sci. 2009; 18(1):247−251.

149. Freskgård PO, Mårtensson LG, Jonasson P, Jonsson BH, Carlsson U. Assignment of the contribution of the tryptophan residues to the circular dichroism spectrum of human carbonic anhydrase II. Biochemistry. 1994; 33(47):14281−14288.

150. Nicolas L, Nielsen-Leroux C, Charles JF, Delécluse A. Respective role of the 42- and 51-kDa components of the *Bacillus sphaericus* toxin overexpressed in *Bacillus thuringiensis*. FEMS Microbiol Lett. 1993; 106(3):275−280.

151. Singer S. Insecticidal activity of recent bacterial isolates and their toxins against mosquito larvae. Nature. 1973; 244(5411):110−111.

152. Hong Y, Ohishi K, Inoue N, Kang JY, Shime H, Horiguchi Y, van der Goot FG, Sugimoto N, Kinoshita T. Requirement of N-glycan on GPI-anchored proteins for efficient binding of aerolysin but not *Clostridium septicum* alpha-toxin. EMBO J. 2002; 21(19):5047−5056.

153. Degiacomi MT, Iacovache I, Pernot L, Chami M, Kudryashev M, Stahlberg H, vanderGoot FG, Dal Peraro M. Molecular assembly of the aerolysin pore reveals a swirling membrane-insertion mechanism. Nat. Chem. Biol. 2013; 9(10):623−629.

154. Walski T, Van Damme EJ, Smargiasso N, Christiaens O, De Pauw E, Smagghe G. Protein N-glycosylation and N-glycan trimming are required for postembryonic development of the pest beetle *Tribolium castaneum*. Sci Rep. 2016; 6:35151.

155. Walski T, De Schutter K, Van Damme EJ, Smagghe G. Diversity and functions of protein glycosylation in insects. Insect Biochem. Mol. Biol. 2017; 83:21−34.

Appendix

Appendix A

Crystallization and preliminary structural analysis of Cqm1, receptor protein for binary toxin

The receptor for binary toxin (BinA/B) has been identified as 66 kDa, α-glucosidase, which remains bound to the epithelial membrane in larval midgut through the GPI anchor. The active core of the protein (residues 23-560) has been successfully crystallized using D-arabinose and $CaCl_2$ as additives. These additives were selected by screening for sugar substrates, using thermofluor shift assay, that may impart conformational stability to the protein. The crystals of Cqm1 were obtained using JCSG+ screen. The crystals diffracted to a resolution of 2.8 Å using laboratory X-ray source and belong to space group $P2_12_12$, with unit cell parameters, a=191.3, b=205.3, c= 59.0 Å.

1. Materials and Methods

1.1. Expression and purification of Cqm1 protein

Detailed cloning, expression and purification protocols of the active core of Cqm1 protein (residues 23-560), without signal peptide and GPI anchor residues, has been described in chapter 4. The expressed protein was purified to homogeneity using IMAC and AEC chromatography methods.

1.2. Thermofluor shift assay to screen sugar substrates

Identification of a probable additive for Cqm1 crystallization, to enhance conformational and thermal stability of the protein, was attempted through thermofluor shift assay. 29 different sugars were incubated with Cqm1 (2 μM) in 1:100 molar ratio in separate reactions for 1 hr in 25 μL reaction volume in buffer A. Following incubation freshly diluted SYPRO orange dye (1:000, v/v) was added and the measurements were done in duplicates, in a 96 well plate, using CFX96^{TM} Real time system (Bio-Rad) in FRET mode. The samples were heated from

20 to 90 ℃ at a ramp rate of 1 ℃/min. Thermal unfolding of Cqm1 alone and Hen egg white (HEW) lysozyme were also probed in the same plate as an internal control. The experimental data was processed using software available with CFX96^{TM} Real time system and, melting temperature was determined from the first derivative of each of the thermofluor-shift melting curve.

1.3. Crystallization and X-Ray diffraction data collection

The Cqm1 protein in buffer A (20 mM HEPES pH 7.5, 20 mM NaCl) was incubated with D-arabinose (protein: sugar; 1:10 molar ratio) for 1 hr at 4 ℃. Excess sugar was removed using ultra-centrifugal device (10 kDa cutoff) with repeated buffer exchange and the protein was finally concentrated to 15 mg/mL. $CaCl_2$ was added into the concentrated protein sample to a final concentration, 0.1 mM. Crystallization was performed by sitting and hanging-drop vapor-diffusion methods in 24- and 96-well plates at 295 K respectively. Initial screening was performed by sitting-drop vapour-diffusion method using JCSG+ and PACT crystallization screening suites from Qiagen. Drops were prepared by mixing 2 µL of Cqm1 protein (with additives) with an equivalent volume of reservoir solution and were equilibrated against 100 µL of reservoir solution in sitting drop method. Crystals for diffraction experiments were grown in the selected condition using hanging-drop vapour- diffusion method in which drops were prepared by mixing 2.5 µL of Cqm1 protein with an equivalent volume of reservoir solution and were equilibrated against 400 µL of reservoir solution. Protein crystals were briefly soaked in cryoprotecting solutions consisting of 10% and 20% (v/v) glycerol included in the mother liquor, respectively, and flash frozen in liquid nitrogen steam.

The diffraction intensity data were collected using fine focus X-ray source (Microstar, Bruker) under cryo-conditions. A total of 120 frames were collected at an exposure time of 2 min for each image. The diffraction images were recorded on an image plate detector (MAR

RESEARCH) using Cu Kα radiation with 1.0° oscillation per image. The data were processed using XDS/autoPROC software [1].

2. Results and discussion

Thermal stability of Cqm1 protein was observed to be buffer dependent. For instance, melting temperature (*T*m) of 48 ℃ was observed in phosphate buffer pH 8, while a higher *T*m of 52 ℃ was observed in HEPES buffer pH 7.5. Hence, the buffer of the protein was exchanged to buffer A prior to crystallization. Ca (II) was found to be essential for thermal stability of the protein (Fig. 4.4C, D). Thermo-fluor shift assay against different sugar substrates revealed that none of the used sugars affected the *T*m of the protein. However, a distinct temperature dependent melting profile for the tertiary structure of Cqm1 protein was observed in the presence of D-arabinose (Fig. A1). The Cqm1 protein was equilibrated with D-arabinose prior to crystallization. This strategy certainly helped crystallization, as Cqm1 without D-arabinose additive, did not crystallize.

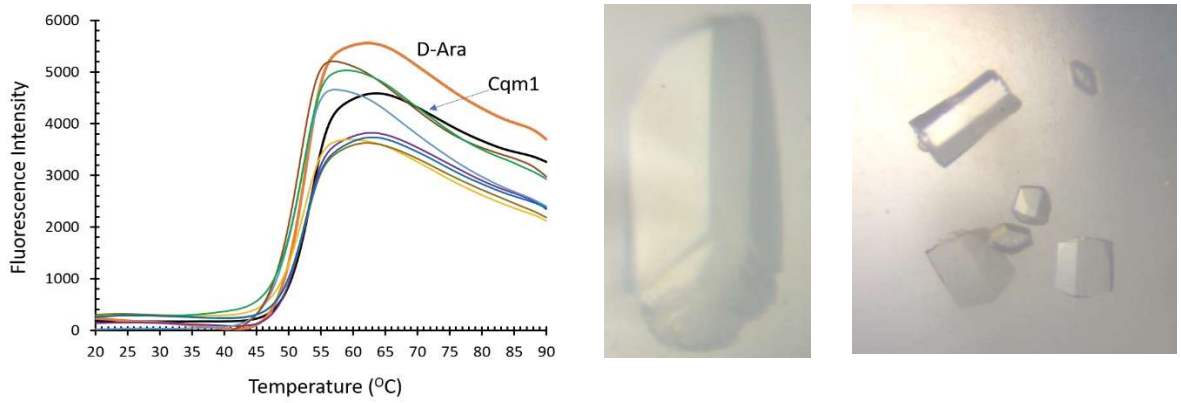
Best quality crystals appeared in one of the conditions of the JCSG+ screen consisting of 5 mM $CoCl_2$, 5 mM $CdCl_2$, 5 mM $MgCl_2$, 5 mM $NiCl_2$, 0.1 mM HEPES pH 7.5 and 12% PEG 3350 at 295 K. The crystals grew to their maximal size of ~500 µm within 1 week (Fig. A2). The crystals were reproduced using hanging-drop method for diffraction data collection (Fig. A3).

X-ray diffraction data of the Cqm1 crystal were collected using a laboratory X-ray source and an Image plate detector. Crystals diffracted to 2.8 Å resolution (Fig. A4). The data collection statistics is given in table A1. The crystals belong to the orthorhombic space group $P2_12_12$ with unit cell parameters a=191.3, b=205.3, c= 59.0 Å, α=β=γ=90°. The Mathews coefficient (*V*m) was estimated to be 2.24 $Å^3$ Da^{-1} with four molecules in the asymmetric unit, corresponding to a solvent content of 45.2 % [2]. Structure determination is in progress.

Table A1: Summary of crystallographic data	
Resolution	2.8 Å
Space group	$P2_12_12$
Unit cell parameters	a=191.3, b=205.3, c= 59.0 Å; α= β= γ= 90 º
Data Resolution (Å)	39 – 2.8 (2.95 – 2.80) *
Total no. of observations	242945 (36526)
Number of unique observations	56101 (8382)
Multiplicity	4.3 (4.4)
Completeness (%)	96.1 (99.7)
R_{merge}	0.108 (0.348)
$I/\sigma(I)$	9.7 (3.4)

*) Values in the bracket are for the highest resolution shell

1) **2)** **3)**



4)

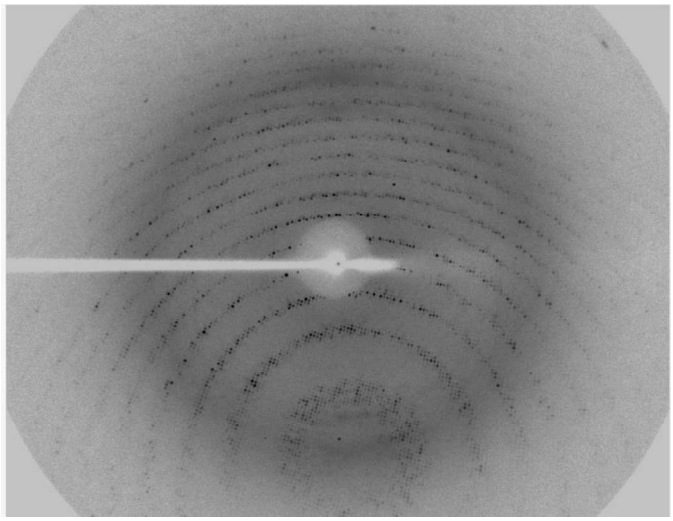


Figure A. Crystallization and X-ray diffraction analysis of Cqm1. 1) Thermo-fluor shift assay to determine the melting profile of Cqm1 protein in the presence of different sugar substrates. Melting profile of few selected sugars is shown. **2)** Crystals of Cqm1 grown during initial crystallization trial using sitting-drop method. **3)** Crystals of Cqm1 grown using hanging drop method used for X-ray diffraction experiments. **4)** X-ray diffraction image of Cqm1 crystal.

References

1. Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, et al. Data processing and analysis with the autoPROC toolbox. Acta Crystallogr D Biol Crystallogr. 2011; 67: 293–302.
2. Matthews BW. Solvent content of protein crystals. J Mol Biol. 1968; 33: 491−497.

Appendix-B

The CDS of *cqm*1 gene, (binary toxin receptor protein, Cqm1) from *Culex quinquefasciatus* isolate Trombay, GenBank entry: KY929304

```
1    ATG CGA CCG CTG GGA GCT TTG AGC TTA GTC GCT CTG TTG GCG ACG   45
1    Met Arg Pro Leu Gly Ala Leu Ser Leu Val Ala Leu Leu Ala Thr   15

46   ACC GTC AAC GGG CTA GCG ATT CGC GAA CCG GAC TCG AAG GAC TGG   90
16   Thr Val Asn Gly Leu Ala Ile Arg Glu Pro Asp Ser Lys Asp Trp   30

91   TAT CAG CAC GCG ACG TTC TAC CAG ATC TAC CCG AGG TCG TTT TTG  135
31   Tyr Gln His Ala Thr Phe Tyr Gln Ile Tyr Pro Arg Ser Phe Leu   45

136  GAC AGC AAC GGC GAT GGG ATA GGT GAT TTG GCG GGG ATC ACC TCC  180
46   Asp Ser Asn Gly Asp Gly Ile Gly Asp Leu Ala Gly Ile Thr Ser   60

181  AAG ATG AAG TAC TTG GCG GAT ATT GGG ATT GAC GCG ACG TGG TTG  225
61   Lys Met Lys Tyr Leu Ala Asp Ile Gly Ile Asp Ala Thr Trp Leu   75

226  AGC CCG CCG TTC AAG TCA CCT CTG AAA GAC TTT GGG TAT GAT GTG  270
76   Ser Pro Pro Phe Lys Ser Pro Leu Lys Asp Phe Gly Tyr Asp Val   90

271  TCG GAT TTC TAC GCT ATC CAG CCG GAG TAC GGG AAT TTG ACG GAT  315
91   Ser Asp Phe Tyr Ala Ile Gln Pro Glu Tyr Gly Asn Leu Thr Asp  105

316  TTT GAC AAG TTG GTG GAG GAA TCG CAC AAG AAT GGG ATC AAG CTG  360
106  Phe Asp Lys Leu Val Glu Glu Ser His Lys Asn Gly Ile Lys Leu  120

361  ATG CTG GAC TTT ATT CCG AAC CAC TCG AGT GAT CAG CAC GAG TGG  405
121  Met Leu Asp Phe Ile Pro Asn His Ser Ser Asp Gln His Glu Trp  135

406  TTT GTG AAG TCT GTG GTG CGG GAT CCG GAG TAT AGT GAC TTT TAT  450
136  Phe Val Lys Ser Val Val Arg Asp Pro Glu Tyr Ser Asp Phe Tyr  150

451  GTG TGG AGA CCT CCG GCG ACA GGG GGT GGT CCG CCG AAT AAT TGG  495
151  Val Trp Arg Pro Pro Ala Thr Gly Gly Gly Pro Pro Asn Asn Trp  165

496  ATC TCG GTA TTT GGA GGG TCG GCG TGG ACA TAT AAT CAA GCG AGA  540
166  Ile Ser Val Phe Gly Gly Ser Ala Trp Thr Tyr Asn Gln Ala Arg  180

541  GGG GAG TAC TAT CTG CAC CAG TTT ACG CCT CAG CAG CCG GAT TTG  585
181  Gly Glu Tyr Tyr Leu His Gln Phe Thr Pro Gln Gln Pro Asp Leu  195

586  AAC TAC CGG AAT CCG AAG GTG CTG GCG GAG ATG ACC AAG ATG TTG  630
196  Asn Tyr Arg Asn Pro Lys Val Leu Ala Glu Met Thr Lys Met Leu  210

631  TTC TTC TGG TTG GAC CGT GGA GTG GAT GGG TTC CGG TTG GAC GCT  675
211  Phe Phe Trp Leu Asp Arg Gly Val Asp Gly Phe Arg Leu Asp Ala  225

676  ATC AAC CAC ATG TTC GAG GAT GAG CAG TTC CGG GAT GAG CCG TTG  720
226  Ile Asn His Met Phe Glu Asp Glu Gln Phe Arg Asp Glu Pro Leu  240

721  TCT GGG TGG GGA CAG CCG GGA GAG TAC GAT TCT TTG GAT CAC ATT  765
241  Ser Gly Trp Gly Gln Pro Gly Glu Tyr Asp Ser Leu Asp His Ile  255

766  TAT ACC AAG GAC ATC CCG GAT GTG TAT GAT GTG GTT TAC AAC TGG  810
256  Tyr Thr Lys Asp Ile Pro Asp Val Tyr Asp Val Val Tyr Asn Trp  270
```

```
811  CGG GAT CAG ATG GAT AAG TAT TCC GCG GAG AAG GGC CGC ACT ATC  855
271  Arg Asp Gln Met Asp Lys Tyr Ser Ala Glu Lys Gly Arg Thr Ile  285

856  ATT CTG ATG ACG GAA GCG TAT TCG AGC ATC GAG GGA ACG ATG CTG  900
286  Ile Leu Met Thr Glu Ala Tyr Ser Ser Ile Glu Gly Thr Met Leu  300

901  TAC TAC GAG AGT GCG GAC CGG AAG CGC CAG GGA GCT CAC ATG CCG  945
301  Tyr Tyr Glu Ser Ala Asp Arg Lys Arg Gln Gly Ala His Met Pro  315

946  TTC AAC TTC CAG CTG ATC TAT GAC TTT AAG AAG GAG CAG AAC GCC  990
316  Phe Asn Phe Gln Leu Ile Tyr Asp Phe Lys Lys Glu Gln Asn Ala  330

991  GTT GGG CTG AAG AGT TCG ATC GAC TGG TGG ATG AAC AAC ATG CCG 1035
331  Val Gly Leu Lys Ser Ser Ile Asp Trp Trp Met Asn Asn Met Pro  345

1036 GCT CGA CAC ACT CCA AGC TGG GTA GCT GGA TCG CAC GAT CAT TCA 1080
346  Ala Arg His Thr Pro Ser Trp Val Ala Gly Ser His Asp His Ser  360

1081 CGG GTA GCG TCA CGA GTT GGA CTA GAT CGG GTT GAT CAG GTC ATG 1125
361  Arg Val Ala Ser Arg Val Gly Leu Asp Arg Val Asp Gln Val Met  375

1126 ACG TTG CTG CAT ACC CTG CCG GGG ACT AGT ATT ACG TAT TAC GTA 1170
376  Thr Leu Leu His Thr Leu Pro Gly Thr Ser Ile Thr Tyr Tyr Val  390

1171 AGT ACC AGT GAG TTC CTC CAA GAT TCT GTC TAA CAT TTG TTT TCC 1215
391  Ser Thr Ser Glu Phe Leu Gln Asp Ser Val End His Leu Phe Ser  405

1216 AGG GTG AAG AGG TAG CCA TGC AGG ACT TTA AGG AAG CTC AGC AGT 1260
406  Arg Val Lys Arg End Pro Cys Arg Thr Leu Arg Lys Leu Ser Ser  420

1261 TTG ACA ACC GGG ATC CGA ACC GAA CGC CTA TGC AGT GGG ACT CGT 1305
421  Leu Thr Thr Gly Ile Arg Thr Glu Arg Leu Cys Ser Gly Thr Arg  435

1306 CGA CCA GTG CCG GGT TCA GTA CCA ACA CCA ACA CCT GGC TCC GAG 1350
436  Arg Pro Val Pro Gly Ser Val Pro Thr Pro Thr Pro Gly Ser Glu  450

1351 TTC ATC CGG ATT ACG CTC GGT ACA ACG TGG ACG TGA TGC AAA AGA 1395
451  Phe Ile Arg Ile Thr Leu Gly Thr Thr Trp Thr End Cys Lys Arg  465

1396 ATC CAC AAA GTA CCT TCC ACC ACT TCC AGC ATC TGA CCA AGC TGC 1440
466  Ile His Lys Val Pro Ser Thr Thr Ser Ser Ile End Pro Ser Cys  480

1441 GAC GGC ACC GGA CGA TGC AGA GTG GTG AGT ACG TGC ACA AGA CGG 1485
481  Asp Gly Thr Gly Arg Cys Arg Val Val Ser Thr Cys Thr Arg Arg  495

1486 TCG GAA CCA AGG TGT ACG CTT TGC TGA GGG AAC TCC GTG GTG AGG 1530
496  Ser Glu Pro Arg Cys Thr Leu Cys End Gly Asn Ser Val Val Arg  510


1531 ATT CGT TCC TGA CGG TGC TGA ACA TGG CCG GAG CAG AGG ATA CCG 1575
511  Ile Arg Ser End Arg Cys End Thr Trp Pro Glu Gln Arg Ile Pro  525

1576 TGG ATC TGG GAG ATT TCG TGA ATC TTC CGC AGA AGA TGC GAG TTG 1620
526  Trp Ile Trp Glu Ile Ser End Ile Phe Arg Arg Arg Cys Glu Leu  540

1621 AGG TGG CGC AAC CGA ACT CCA AGT CGA AGG CGG GGT AAG AAC TCT 1665
541  Arg Trp Arg Asn Arg Thr Pro Ser Arg Arg Arg Gly Lys Asn Ser  555

1666 AAC TTC TCT TAT AGA AGT ATT TCT AAC ATG TTT TCC CCT AAC AGC 1710
556  Asn Phe Ser Tyr Arg Ser Ile Ser Asn Met Phe Ser Pro Asn Ser  570
```

```
1711 AAT GAG GTC GAC ATC AGC AAA TTG ACG CTG GGA CCG TAT GAT TCG 1755
571  Asn Glu Val Asp Ile Ser Lys Leu Thr Leu Gly Pro Tyr Asp Ser  585

1756 GTT GTG CTG AGA GCA ACG GTA TCG TCG GCA GCT GCT ATC AAC CTT 1800
586  Val Val Leu Arg Ala Thr Val Ser Ser Ala Ala Ala Ile Asn Leu  600

1801 TCG ATT GGA TTG CTG CTA GCG ATT ATG GCC AGG TAT ATT TTC GTG 1845
601  Ser Ile Gly Leu Leu Leu Ala Ile Met Ala Arg Tyr Ile Phe Val  615

1846 TAG  1848
616  End   616
```

Appendix-C

The CDS of *cqm*1 gene, (binary toxin receptor protein, Cqm1) from *Culex quinquefasciatus* isolate Mankhurd, GenBank entry: KY929305

```
1    ATG CGA CCG CTG GGA GCT TTG AGC TTA GTC GCT CTG TTG GCG ACG   45
1    Met Arg Pro Leu Gly Ala Leu Ser Leu Val Ala Leu Leu Ala Thr   15

46   ACC GTC AAC GGG CTA GCG ATT CGC GAA CCG GTC TCG AAG GAC TGG   90
16   Thr Val Asn Gly Leu Ala Ile Arg Glu Pro Val Ser Lys Asp Trp   30

91   TAT CAG CAC GCG ACG TTC TAC CAG ATC TAC CCG AGG TCG TTT TTG  135
31   Tyr Gln His Ala Thr Phe Tyr Gln Ile Tyr Pro Arg Ser Phe Leu   45

136  GAC AGC AAC GGC GAT GGG ATA GGT GAT TTG GCG GGG ATC ACC TCC  180
46   Asp Ser Asn Gly Asp Gly Ile Gly Asp Leu Ala Gly Ile Thr Ser   60

181  AAG ATG AAG TAC TTG GCG GAT ATT GGG ATT GAC GCG ACG TGG TTG  225
61   Lys Met Lys Tyr Leu Ala Asp Ile Gly Ile Asp Ala Thr Trp Leu   75

226  AGC CCG CCG TTC AAG TCA CCT CTG AAA GAC TTT GGG TAT GAT GTG  270
76   Ser Pro Pro Phe Lys Ser Pro Leu Lys Asp Phe Gly Tyr Asp Val   90

271  TCG GAT TTC TAC GCT ATC CAG CCG GAG TAC GGG AAT TTG ACG GAT  315
91   Ser Asp Phe Tyr Ala Ile Gln Pro Glu Tyr Gly Asn Leu Thr Asp  105

316  TTT GAC AAG TTG GTG GAG GAA TCG CAC AAG AAT GGG ATC AAG CTG  360
106  Phe Asp Lys Leu Val Glu Glu Ser His Lys Asn Gly Ile Lys Leu  120

361  ATG CTG GAC TTT ATT CCG AAC CAC TCG AGT GAT CAG CAC GAG TGG  405
121  Met Leu Asp Phe Ile Pro Asn His Ser Ser Asp Gln His Glu Trp  135

406  TTT GTG AAG TCT GTG GTG CGG GAT CCG GAG TAT AGT GAC TTT TAT  450
136  Phe Val Lys Ser Val Val Arg Asp Pro Glu Tyr Ser Asp Phe Tyr  150

451  GTG TGG AGA CCT CCG GCG ACA GGG GGT GGT CCG CCG AAT AAT TGG  495
151  Val Trp Arg Pro Pro Ala Thr Gly Gly Gly Pro Pro Asn Asn Trp  165

496  ATC TCG GTA TTT GGA GGG TCG GCG TGG ACA TAT AAT CAA GCG AGA  540
166  Ile Ser Val Phe Gly Gly Ser Ala Trp Thr Tyr Asn Gln Ala Arg  180

541  GGG GAG TAC TAT CTG CAC CAG TTT ACG CCT CAG CAG CCG GAT TTG  585
181  Gly Glu Tyr Tyr Leu His Gln Phe Thr Pro Gln Gln Pro Asp Leu  195

586  AAC TAC CGG AAT CCG AAG GTG CTG GCG GAG ATG ACC AAG ATG TTG  630
196  Asn Tyr Arg Asn Pro Lys Val Leu Ala Glu Met Thr Lys Met Leu  210

631  TTC TTC TGG TTG GAC CGT GGA GTG GAT GGG TTC CGG TTG GAC GCT  675
211  Phe Phe Trp Leu Asp Arg Gly Val Asp Gly Phe Arg Leu Asp Ala  225

676  ATC AAC CAC ATG TTC GAG GAT GAG CAG TTC CGG GAT GAG CCG TTG  720
226  Ile Asn His Met Phe Glu Asp Glu Gln Phe Arg Asp Glu Pro Leu  240

721  TCT GGG TGG GGA CAG CCG GGA GAG TAC GAT TCT TTG GAT CAC ATT  765
241  Ser Gly Trp Gly Gln Pro Gly Glu Tyr Asp Ser Leu Asp His Ile  255


766  TAT ACC AAG GAC ATC CCG GAT GTG TAT GAT GTG GTT TAC AAC TGG  810
256  Tyr Thr Lys Asp Ile Pro Asp Val Tyr Asp Val Val Tyr Asn Trp  270
```

```
811  CGG GAT CAG ATG GAT AAG TAT TCC GCG GAG AAG GGC CGC ACT ATC  855
271  Arg Asp Gln Met Asp Lys Tyr Ser Ala Glu Lys Gly Arg Thr Ile  285

856  ATT CTG ATG ACG GAA GCG TAT TCG AGC ATC GAG GGA ACG ATG CTG  900
286  Ile Leu Met Thr Glu Ala Tyr Ser Ser Ile Glu Gly Thr Met Leu  300

901  TAC TAC GAG AGT GCG GAC CGG AAG CGC CAG GGA GCT CAC ATG CCG  945
301  Tyr Tyr Glu Ser Ala Asp Arg Lys Arg Gln Gly Ala His Met Pro  315

946  TTC AAC TTC CAG CTG ATC TAT GAC TTT AAG AAG GAG CAG AAC GCC  990
316  Phe Asn Phe Gln Leu Ile Tyr Asp Phe Lys Lys Glu Gln Asn Ala  330

991  GTT GGG CTG AAG AGT TCG ATC GAC TGG TGG ATG AAC AAT ATG CCG 1035
331  Val Gly Leu Lys Ser Ser Ile Asp Trp Trp Met Asn Asn Met Pro  345

1036 GCT CGA CAC ACT CCA AGC TGG GTA GCT GGA TCG CAC GAT CAT TCA 1080
346  Ala Arg His Thr Pro Ser Trp Val Ala Gly Ser His Asp His Ser  360

1081 CGG GTA GCG TCA CGA GTT GGA CTA GAT CGG GTT GAT CAG GTC ATG 1125
361  Arg Val Ala Ser Arg Val Gly Leu Asp Arg Val Asp Gln Val Met  375

1126 ACG TTG CTG CAT ACC CTG CCG GGG ACT AGT ATT ACG TAT TAC GTA 1170
376  Thr Leu Leu His Thr Leu Pro Gly Thr Ser Ile Thr Tyr Tyr Val  390

1171 AGT ACC AGT GAG TTC CTC CAA GAT TCT GTC TAA CAT TTG TTT TCC 1215
391  Ser Thr Ser Glu Phe Leu Gln Asp Ser Val End His Leu Phe Ser  405

1216 AGG GTG AAG AGG TAG CCA TGC AGG ACT TTA AGG AAG CTC AGC AGT 1260
406  Arg Val Lys Arg End Pro Cys Arg Thr Leu Arg Lys Leu Ser Ser  420

1261 TTG ACA ACC GGG ATC CGA ACC GAA CGC CTA TGC AGT GGG ACT CGT 1305
421  Leu Thr Thr Gly Ile Arg Thr Glu Arg Leu Cys Ser Gly Thr Arg  435

1306 CGA CCA GTG CCG GGT TCA GTA CCA ACA CCA ACA CCT GGC TCC GAG 1350
436  Arg Pro Val Pro Gly Ser Val Pro Thr Pro Thr Pro Gly Ser Glu  450

1351 TTC ATC CGG ATT ACG CTC GGT ACA ACG TGG ACG TGA TGC AAA AGA 1395
451  Phe Ile Arg Ile Thr Leu Gly Thr Thr Trp Thr End Cys Lys Arg  465

1396 ATC CAC AAA GTA CCT TCC ACC ACT TCC AGC ATC TGA CCA AGC TGC 1440
466  Ile His Lys Val Pro Ser Thr Thr Ser Ser Ile End Pro Ser Cys  480

1441 GAC GGC ACC GGA CGA TGC AGA GTG GTG AGT ACG TGC ACA AGA CGG 1485
481  Asp Gly Thr Gly Arg Cys Arg Val Val Ser Thr Cys Thr Arg Arg  495

1486 TCG GAA CCA AGG TGT ACG CTA TGC TGA GGG AAC TCC GTG GTG AGG 1530
496  Ser Glu Pro Arg Cys Thr Leu Cys End Gly Asn Ser Val Val Arg  510

1531 ATT CGT TCC TGA CGG TGC TGA ACA TGG CCG GAG CAG AGG ATA CCG 1575
511  Ile Arg Ser End Arg Cys End Thr Trp Pro Glu Gln Arg Ile Pro  525

1576 TGG ATC TGG GTG ATT TCG TGA ATC TTC CGC AGA AGA TGC GAG TTG 1620
526  Trp Ile Trp Val Ile Ser End Ile Phe Arg Arg Arg Cys Glu Leu  540

1621 AGG TGG CGC AAC CGA ACT CCA AGT CGA AGG CGG GGT AAG AAC TCT 1665
541  Arg Trp Arg Asn Arg Thr Pro Ser Arg Arg Arg Gly Lys Asn Ser  555


1666 AAC TTC TCT TAT AGA AGT ATT TCT AAC ATG TTT TCC CCT AAC AGC 1710
556  Asn Phe Ser Tyr Arg Ser Ile Ser Asn Met Phe Ser Pro Asn Ser  570
```

```
1711 AAT GAG GTC GAC ATC AGC AAA TTG ACG CTG GGA CCG TAT GAT TCG 1755
571  Asn Glu Val Asp Ile Ser Lys Leu Thr Leu Gly Pro Tyr Asp Ser  585

1756 GTT GTG CTG AGA GCA ACG GTA TCT ACG GCA GCT GCT ATC AAC CTT 1800
586  Val Val Leu Arg Ala Thr Val Ser Thr Ala Ala Ala Ile Asn Leu  600

1801 TCG ATT GGA TTG CTG CTA GCG ATT ATG GCC AGG TAT ATT TTC GTG 1845
601  Ser Ile Gly Leu Leu Leu Ala Ile Met Ala Arg Tyr Ile Phe Val  615

1846 TAG  1848
616  End   616
```

Appendix-D

The CDS of synthetic *cqm*1 gene, GenBank entry: MG211157

```
1    ATG CGC CCG CTG GGT GCT CTG TCT CTG GTG GCT CTG CTG GCT ACG   45
1    Met Arg Pro Leu Gly Ala Leu Ser Leu Val Ala Leu Leu Ala Thr   15

46   ACC GTG AAT GGC CTG GCT ATC CGT GAA CCG GAC TCA AAA GAT TGG   90
16   Thr Val Asn Gly Leu Ala Ile Arg Glu Pro Asp Ser Lys Asp Trp   30

91   TAT CAG CAT GCA ACC TTC TAT CAA ATT TAC CCG CGT AGC TTT CTG  135
31   Tyr Gln His Ala Thr Phe Tyr Gln Ile Tyr Pro Arg Ser Phe Leu   45

136  GAT TCT AAC GGT GAC GGC ATT GGT GAT CTG GCT GGC ATC ACG TCT  180
46   Asp Ser Asn Gly Asp Gly Ile Gly Asp Leu Ala Gly Ile Thr Ser   60

181  AAA ATG AAA TAT CTG GCG GAC ATT GGT ATC GAT GCC ACC TGG CTG  225
61   Lys Met Lys Tyr Leu Ala Asp Ile Gly Ile Asp Ala Thr Trp Leu   75

226  TCT CCG CCG TTT AAA AGT CCG CTG AAA GAT TTC GGC TAT GAT GTT  270
76   Ser Pro Pro Phe Lys Ser Pro Leu Lys Asp Phe Gly Tyr Asp Val   90

271  TCT GAC TTT TAC GCA ATC CAG CCG GAA TAT GGT AAT CTG ACC GAT  315
91   Ser Asp Phe Tyr Ala Ile Gln Pro Glu Tyr Gly Asn Leu Thr Asp  105

316  TTC GAC AAA CTG GTC GAA GAA AGT CAT AAA AAC GGC ATT AAA CTG  360
106  Phe Asp Lys Leu Val Glu Glu Ser His Lys Asn Gly Ile Lys Leu  120

361  ATG CTG GAC TTC ATC CCG AAT CAT AGC TCT GAT CAG CAC GAA TGG  405
121  Met Leu Asp Phe Ile Pro Asn His Ser Ser Asp Gln His Glu Trp  135

406  TTT GTT AAA TCG GTG GTT CGT GAC CCG GAA TAT AGC GAT TTT TAC  450
136  Phe Val Lys Ser Val Val Arg Asp Pro Glu Tyr Ser Asp Phe Tyr  150

451  GTC TGG CGC CCG CCG GCA ACC GGC GGT GGC CCG CCG AAC AAT TGG  495
151  Val Trp Arg Pro Pro Ala Thr Gly Gly Gly Pro Pro Asn Asn Trp  165

496  ATT TCC GTT TTC GGT GGC TCA GCA TGG ACG TAC AAC CAG GCT CGT  540
166  Ile Ser Val Phe Gly Gly Ser Ala Trp Thr Tyr Asn Gln Ala Arg  180

541  GGT GAA TAT TAC CTG CAC CAA TTT ACC CCG CAG CAA CCG GAT CTG  585
181  Gly Glu Tyr Tyr Leu His Gln Phe Thr Pro Gln Gln Pro Asp Leu  195

586  AAC TAT CGC AAT CCG AAA GTG CTG GCG GAA ATG ACC AAA ATG CTG  630
196  Asn Tyr Arg Asn Pro Lys Val Leu Ala Glu Met Thr Lys Met Leu  210

631  TTT TTC TGG CTG GAC CGT GGC GTT GAT GGT TTT CGC CTG GAT GCC  675
211  Phe Phe Trp Leu Asp Arg Gly Val Asp Gly Phe Arg Leu Asp Ala  225

676  ATC AAC CAC ATG TTT GAA GAC GAA CAA TTC CGC GAT GAA CCG CTG  720
226  Ile Asn His Met Phe Glu Asp Glu Gln Phe Arg Asp Glu Pro Leu  240

721  TCC GGT TGG GGT CAG CCG GGT GAA TAC GAT TCA CTG GAC CAC ATT  765
241  Ser Gly Trp Gly Gln Pro Gly Glu Tyr Asp Ser Leu Asp His Ile  255

766  TAT ACG AAA GAT ATT CCG GAC GTG TAT GAT GTC GTG TAC AAT TGG  810
256  Tyr Thr Lys Asp Ile Pro Asp Val Tyr Asp Val Val Tyr Asn Trp  270

811  CGT GAC CAG ATG GAT AAA TAC TCC GCG GAA AAA GGT CGC ACG ATT  855
271  Arg Asp Gln Met Asp Lys Tyr Ser Ala Glu Lys Gly Arg Thr Ile  285
```

```
856  ATC CTG ATG ACC GAA GCC TAT AGT TCC ATT GAA GGC ACC ATG CTG  900
286  Ile Leu Met Thr Glu Ala Tyr Ser Ser Ile Glu Gly Thr Met Leu  300

901  TAT TAC GAA TCA GCG GAT CGT AAA CGC CAG GGT GCC CAC ATG CCG  945
301  Tyr Tyr Glu Ser Ala Asp Arg Lys Arg Gln Gly Ala His Met Pro  315

946  TTC AAC TTC CAA CTG ATC TAC GAT TTC AAA AAA GAA CAG AAT GCG  990
316  Phe Asn Phe Gln Leu Ile Tyr Asp Phe Lys Lys Glu Gln Asn Ala  330

991  GTT GGC CTG AAA TCA TCG ATC GAT TGG TGG ATG AAC AAT ATG CCG 1035
331  Val Gly Leu Lys Ser Ser Ile Asp Trp Trp Met Asn Asn Met Pro  345

1036 GCC CGT CAT ACC CCG AGC TGG GTG GCC GGC AGT CAT GAT CAC TCC 1080
346  Ala Arg His Thr Pro Ser Trp Val Ala Gly Ser His Asp His Ser  360

1081 CGT GTG GCC TCA CGC GTT GGT CTG GAC CGC GTT GAT CAG GTC ATG 1125
361  Arg Val Ala Ser Arg Val Gly Leu Asp Arg Val Asp Gln Val Met  375

1126 ACC CTG CTG CAC ACC CTG CCG GGC ACG AGC ATC ACC TAT TAC GGT 1170
376  Thr Leu Leu His Thr Leu Pro Gly Thr Ser Ile Thr Tyr Tyr Gly  390

1171 GAA GAA GTT GCA ATG CAG GAT TTT AAA GAA GCT CAG CAA TTC GAC 1215
391  Glu Glu Val Ala Met Gln Asp Phe Lys Glu Ala Gln Gln Phe Asp  405

1216 AAC CGT GAT CCG AAT CGC ACG CCG ATG CAG TGG GAT AGC TCT ACC 1260
406  Asn Arg Asp Pro Asn Arg Thr Pro Met Gln Trp Asp Ser Ser Thr  420

1261 TCG GCA GGC TTT AGC ACC AAC ACG AAT ACC TGG CTG CGT GTG CAT 1305
421  Ser Ala Gly Phe Ser Thr Asn Thr Asn Thr Trp Leu Arg Val His  435

1306 CCG GAC TAT GCT CGC TAC AAC GTG GAT GTT ATG CAG AAA AAT CCG 1350
436  Pro Asp Tyr Ala Arg Tyr Asn Val Asp Val Met Gln Lys Asn Pro  450

1351 CAA TCG ACC TTT CAT CAC TTC CAA CAT CTG ACG AAA CTG CGT CGC 1395
451  Gln Ser Thr Phe His His Phe Gln His Leu Thr Lys Leu Arg Arg  465

1396 CAC CGT ACC ATG CAG AGC GGC GAA TAC GTC CAC AAA ACG GTG GGC 1440
466  His Arg Thr Met Gln Ser Gly Glu Tyr Val His Lys Thr Val Gly  480

1441 ACC AAA GTT TAT GCG CTG CTG CGT GAA CTG CGC GGT GAA GAT AGT 1485
481  Thr Lys Val Tyr Ala Leu Leu Arg Glu Leu Arg Gly Glu Asp Ser  495

1486 TTC CTG ACG GTG CTG AAC ATG GCA GGC GCT GAA GAC ACC GTC GAT 1530
496  Phe Leu Thr Val Leu Asn Met Ala Gly Ala Glu Asp Thr Val Asp  510

1531 CTG GGT GAC TTT GTG AAT CTG CCG CAG AAA ATG CGT GTC GAA GTG 1575
511  Leu Gly Asp Phe Val Asn Leu Pro Gln Lys Met Arg Val Glu Val  525

1576 GCA CAA CCG AAC TCA AAA TCG AAA GCT GGC AAT GAA GTC GAC ATT 1620
526  Ala Gln Pro Asn Ser Lys Ser Lys Ala Gly Asn Glu Val Asp Ile  540

1621 TCT AAA CTG ACG CTG GGT CCG TAT GAT AGT GTT GTC CTG CGC GCG 1665
541  Ser Lys Leu Thr Leu Gly Pro Tyr Asp Ser Val Val Leu Arg Ala  555

1666 ACC GTG AGT TCC GCG GCG GCT ATC AAT CTG TCA ATC GGT CTG CTG 1710
556  Thr Val Ser Ser Ala Ala Ala Ile Asn Leu Ser Ile Gly Leu Leu  570

1711 CTG GCA ATT ATG GCA CGC TAC ATC TTC GTC TAA  1743
571  Leu Ala Ile Met Ala Arg Tyr Ile Phe Val End   581
```

Appendix E

Journal Publications

Insect Biochemistry and Molecular Biology 93 (2018) 37–46

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Receptor protein of *Lysinibacillus sphaericus* mosquito-larvicidal toxin displays amylomaltase activity

Mahima Sharma [a, b], Gagan D. Gupta [a, b], Vinay Kumar [a, b, *]

[a] Radiation Biology & Health Sciences Division, Bhabha Atomic Research Centre, Mumbai 400085, India
[b] Homi Bhabha National Institute, Training School Complex, Anushaktinagar, Mumbai 400094, India

ARTICLE INFO

Article history:
Received 21 August 2017
Received in revised form
20 October 2017
Accepted 3 December 2017
Available online 8 December 2017

Keywords:
Culex quinquefasciatus
BinAB toxin binding receptor
Cqm1
α-glucosidase
Amylomaltase

ABSTRACT

The activated binary toxin (BinAB) from *Lysinibacillus sphaericus* binds to surface receptor protein (Cqm1) on the midgut cell membrane and kills *Culex quinquefasciatus* larvae on internalization. Cqm1 is attached to cells via a glycosyl-phosphatidylinositol (GPI) anchor. It has been classified as a member of glycoside hydrolase family 13 of the CAZy database. Here, we report characterization of the ordered domain (residues 23–560) of Cqm1. Gene expressing Cqm1 of BinAB susceptible mosquito was chemically synthesized and the protein was purified using *E. coli* expression system. Values for the Michaelis-Menten kinetics parameters towards 4-nitrophenyl α-D-glucopyranoside (α-pNPG) substrate were estimated to be 0.44 mM (*K*m) and 1.9 s^{-1} (*k*cat). Thin layer chromatography experiments established Cqm1 as α-glucosidase competent to cleave α-1,4-glycosidic bonds of maltose and maltotriose with high glycosyltransferase activity to form glucose-oligomers. The observed hydrolysis and synthesis of glucose-oligomers is consistent with open and accessible active-site in the structural model. The protein also hydrolyses glycogen and sucrose. These activities suggest that Cqm1 may be involved in carbohydrate metabolism in mosquitoes. Further, toxic BinA component does not inhibit α-glucosidase activity of Cqm1, while BinB reduced the activity by nearly 50%. The surface plasmon resonance study reveals strong binding of BinB with Cqm1 (*K*d, 9.8 nM). BinA interaction with Cqm1 however, is 1000-fold weaker. Notably the estimated *K*d values match well with dissociation constants reported earlier with larvae brush border membrane fractions. The Cqm1 protein forms a stable dimer that is consistent with its apical localization in lipid rafts. Its melting temperature (T_m) as observed by thermofluor-shift assay is 51.5 °C and Ca^{2+} provides structural stability to the protein.

1. Introduction

High efficiency, selective spectrum of action, and eco-friendliness, has made *Lysinibacillus sphaericus* binary (BinAB) toxin a successful biological larvicide commercially available (World Health Organization, TDR/BCV/Sphaericus/85.3, 1985; Lacey, 2007). Following solubilisation and activation in the mosquito gut, BinAB binds to a specific receptor on the midgut cell membrane of the susceptible mosquito larvae, leading to larval death (Charles et al., 1997; Davidson, 1988; Oei et al., 1992). The BinB component is responsible for binding to receptor, while BinA subunit exerts larvicidal activity (Charles et al., 1997). BinA and BinB (BinA/B) proteins act synergistically and display maximum larvicidal activity in equimolar ratio. The pre-formed covalent complex of BinA/B proteins and PEGylated BinA alone, also show high larvicidal activity (Kale et al., 2013; Sharma et al., 2017).

The BinAB toxin binding receptor in *Culex quinquefasciatus* has been identified as Cqm1 (*C. quinquefasciatus* maltase 1) protein, which is anchored to the apical membrane of the midgut epithelium cells via glycosyl-phosphatidylinositol (GPI) tail (Darboux et al., 2001; Silva-Filha et al., 1999). Cpm1 (*C. pipiens* maltase 1), the ortholog from *C. pipiens*, was firstly described and found to display α-glucosidase activity (Darboux et al., 2001; Silva-Filha et al., 1999). Homologous Agm3 and Aam1 proteins have also been characterized in *Anopheles gambiae* and *Aedes aegypti* species, respectively. Recently, enzymatic property of the receptor protein from *Culex* and *Aedes* has been investigated using substrate 4-nitrophenyl α-D-glucopyranoside (α-pNPG), and *Aedes* Aam1

* Corresponding author. Radiation Biology & Health Sciences Division, Bhabha Atomic Research Centre, Mumbai 400085, India.
E-mail address: vinay@barc.gov.in (V. Kumar).

https://doi.org/10.1016/j.ibmb.2017.12.002
0965-1748/

protein was found to display remarkably higher catalytic efficiency as compared to Cqm1 (Nascimento et al., 2017). Cqm1 and its orthologs have been classified into GH13 subfamily 17 (GH13_17) of the CAZy database, which includes α-amylases, α-glucosidases, cyclodextringlycosyltransferases (CGTase), branching enzymes and cyclomaltodextrinases (http://www.cazy.org/; Lombard et al., 2014; Stam et al., 2006).

α-glucosidases are essential for carbohydrate metabolism in insects and show diverse substrate specificities (against maltose, maltose-oligosaccharides, sucrose etc.). These preferentially hydrolyse 1,4-glycosidic bonds of α-D-glucosides at nonreducing ends. Three types of α-glucosidases have been defined broadly based on their substrate specificity (Chiba, 1988). Type I α-glucosidase hydrolyses heteroside substrate (sucrose and arylglucoside) more rapidly than holoside substrate (malto-oligosaccharides). Whereas, type II and type III enzymes prefer holoside substrate. Type III is also capable of hydrolysing polysaccharide substrate. Some α-glucosidases are competent to catalyse transglycosylation to produce α-1,4 linked maltose oligosaccharides (Kato et al., 2002).

BinAB does not display toxicity against *Aedes aegypti* mosquitoes. The recent crystal structure of BinAB (Colletier et al., 2016) is expected to facilitate broadening of BinAB specificity against refractory mosquito species. Also, resistance to BinAB in susceptible mosquito population is mostly reported due to mutation in the receptor (Darboux et al., 2002, 2007; Guo et al., 2013; Romão et al., 2006; Chalegre et al., 2015). Characterizing biological activity of Cqm1 is important in determining its physiological role in mosquitoes, as this could provide novel leads for mosquito control programmes. Also, understanding the relationship between the physiological role of Cqm1 and its function as receptor of *L. sphaericus* BinAB toxin might give a new insight in elucidating the mechanism of action of binary toxin (Nascimento et al., 2017).

In the present study, we provide detailed biochemical, biophysical and enzymatic characterization of homogeneously purified Cqm1 receptor protein from *Culex quinquefasciatus*. Cqm1 forms stable dimer and Ca^{2+} was found to enhance thermal stability of the protein. We observed that Cqm1 hydrolyses substrates with α-1,2- and α-1,4-glycosidic bonds and, interestingly, displays glycosyltransferase activity with glucose-oligomers. We also observed that BinB, and not BinA, can reduce catalytic activity of Cqm1 protein. BinB also binds Cqm1 with high affinity, whereas BinA interaction with the receptor is ~1000 fold weaker.

2. Materials and methods

2.1. Materials

Restriction enzymes and *Taq* polymerase were purchased from New England Biolabs. pNIC28-Bsa4 vector was from Opher Gileadi (Addgene plasmid # 26103). Luria Bertani (LB) broth and LB agar were procured from HiMedia Laboratories. *E. coli* XL-10 gold and BL21Star (DE3) bacterial strains were from Stratagene and Novagen, respectively. Superdex™ 200 10/300 GL column, gel filtration protein molecular weight markers and Ni–IDA matrix were from GE Healthcare. High Q cartridge and Thin Layer Chromatography (TLC) plates (Silica gel 60 F254) were from Bio-Rad Laboratories and Merck, respectively. Isopropyl β-D-1-thiogalactopyranoside (IPTG), ethylene glycol-bis(β-amino ethyl ether)-N,N,N′,N′-tetra-acetic acid (EGTA), phenylmethylsulfonyl fluoride (PMSF), Trypsin (proteomics grade, from porcine pancreas), Thrombin (from bovine plasma), SYPRO orange dye and α-pNPG were obtained from Sigma. Sugars and other chemicals were procured from SRL, India.

2.2. Mosquito colony

A BinAB susceptible *Culex quinquefasciatus* colony was established from local strain found in Trombay and Mankhurd regions, Mumbai, India, and were reared without exposure to *L. sphaericus*. The susceptibility was tested against equimolar concentrations of BinA/B proteins (Sharma et al., 2017). *Culex* nucleus culture was maintained at 28 ± 2 °C and 85% relative humidity in a BARC laboratory.

2.3. Plasmid constructs

The *cqm*1 gene of BinAB susceptible *Culex quinquefasciatus* mosquito colony was PCR amplified and sequenced (GenBank accession number: KY929304). The local strain showed polymorphism at six nucleotide sites (GenBank accession number: KY929305). The gene has 1840 bp containing a 1743 bp coding sequence (CDS), which encodes for 580 amino acid long Cqm1 protein. A *cqm*1 gene with CDS of 1743 bp was synthesized chemically with codons optimized for expression in *E. coli* (GenBank accession number: MG211157) through the facilities available with Biotech desk Pvt. Ltd (Bangalore, India) based on *cqm*1 sequence submitted to GenBank (accession number: KY929304).

For expressing soluble Cqm1 protein, truncated synthetic *cqm*1 gene (1614 bp) lacking 66 bases from the 5′-end and 60 bases from the 3′-end was PCR amplified using forward primer 5′-TAC TTC CAA TCC ATG CGT GAA CCG GAC TCA AAA G-3′ and reverse primer 5′-TAT CCA CCT TTA CTG CGC GGA ACT CAC GGT CGC-3'. The PCR amplified truncated *cqm*1 gene was cloned into pET28a based pNIC28-Bsa4 expression vector with an N-terminus poly-His tag (pNIC28-*cqm*1 construct) using ligation independent cloning method and was transformed into the cloning host, *E. coli* XL-10 gold competent cells. The transformed cells were plated onto LB agar plates containing kanamycin (50 μg/ml) and incubated overnight at 37 °C. Positively transformed colonies were selected through colony PCR using Taq DNA polymerase. The recombinant plasmid was isolated using plasmid purification kit (Qiagen) and was transformed into BL21 Star (DE3) competent cells for expression of recombinant Cqm1 protein (538 residues) lacking disordered signal and GPI-anchoring amino acid residues, and fused with an N-terminus poly-His tag. The tag consists of 6 His residues and 16 other residues including TEV protease site.

2.4. Protein purification

The *E. coli* BL21 Star (DE3) culture, harbouring pNIC28-*cqm1* construct, was grown at 37 °C overnight. The pre-culture was diluted (1:100) into fresh LB broth supplemented with kanamycin (50 μg/ml) and was grown at 37 °C until the optical density at 600 nm reached approximately 0.7. The temperature for culturing was then reduced to 18 °C, and the inducer IPTG was added to a final concentration of 0.5 mM. After overnight growth, cells were harvested by centrifugation at 10,000 xg for 5 min. The cell pellet was suspended in the lysis buffer [25 mM Tris–HCl pH 8.0, 100 mM NaCl, 5% sucrose, 1 mM dithiothreitol, 1 mM PMSF, 1 mg/ml lysozyme and 1 tablet of Complete protease inhibitor cocktail (Roche)] and subsequently sonicated in pulse mode (3 sec ON, 3 sec OFF) by using Vibra Cell sonicator. Crude whole cell lysate was centrifuged at 21,000 xg at 4 °C for 30 min. Recombinant Cqm1 protein was isolated from the cell free extract by poly-His tag immobilized metal ion affinity (IMAC) chromatography using Ni–IDA matrix which was pre-equilibrated with buffer A [25 mM Tris-HCl (pH 8), 100 mM NaCl] containing 25 mM imidazole. The bound proteins were eluted using imidazole concentration gradient ranging from 25–1000 mM in the same buffer. Imidazole was removed by overnight dialysis at

4 °C against buffer B [25 mM Tris-HCl (pH 8), 25 mM NaCl]. For further purification, the protein was loaded onto anion exchange chromatography (AEC) column pre-equilibrated with buffer B. The bound Cqm1 protein was eluted using 25–1000 mM NaCl gradient in the same buffer. The buffer of the purified Cqm1 protein was exchanged to buffer B by overnight dialysis at 4 °C. The protein concentration was estimated using modified Folin-Lowry's method (Lowry et al., 1951) using bovine serum albumin (BSA, Sigma) as a standard. Total of 12 mg Cqm1 was purified from 1 L *E. coli* culture. Purified protein was analysed by 15% SDS–PAGE.

Recombinant BinA and BinB proteins were purified as described earlier (Kale et al., 2013; Sharma et al., 2017). Briefly, recombinant BinA (active form; 15–366 amino acids) and BinB (residues 1–480) were expressed in BL21 (DE3) and pLysS host cells, respectively, as fusion protein with N-terminal poly-His tag and a thrombin cleavage site between poly-His tag and target protein. The recombinant proteins were purified from soluble cell lysate using IMAC and AEC matrices. Purified BinA was treated with thrombin for removal of poly-His tag. Purified BinB protein was treated with trypsin (trypsin: BinB: 1:1000, w/w) at 20 °C overnight. The reaction was stopped with 10 mM PMSF. Thrombin treated BinA and trypsin activated BinB (residues 17–448) were further purified by AEC. Buffer of BinA and BinB proteins was exchanged to buffer B by overnight dialysis.

2.5. Oligomeric status of Cqm1 protein

The recombinant Cqm1 protein, treated with TEV protease, was loaded onto size-exclusion chromatography (SEC) Superdex™ 200 10/300 GL column, pre-calibrated with gel filtration protein molecular weight markers (Ribonuclease A, 13.7 kDa; Carbonic anhydrase, 29 kDa; Ovalbumin, 44 kDa; Conalbumin, 75 kDa; Aldolase, 158 kDa; Ferritin, 440 kDa; Thyroglobulin, 669 kDa). The eluted peaks from gel filtration experiments were adjudged on 10% native-PAGE and 15% SDS–PAGE. The SEC purified protein was used for biochemical and biophysical analysis. The integrity and identity of the recombinant protein was confirmed by mass spectrometry and peptide mass fingerprinting.

The purified Cqm1 protein (21 μM) in buffer A was analysed by dynamic light scattering (DLS) using Malvern Zetasizer Nano ZS instrument, to determine its hydrodynamic size. Three sets of measurements with 12 acquisitions in each set were collected at 25 °C. The data analysis was done with the help of Malvern Zetasizer software suite (http://www.malvern.com). To assess stability of Cqm1 oligomer, the protein was incubated with 0.05% triton and with 500 mM sodium chloride in independent reactions at room temperature for 1 hr and the protein profile was observed through DLS analysis.

2.6. Folding status and thermal stability of Cqm1 protein

Circular dichroism (CD) spectrum for Cqm1 protein (0.75 mg/ml, 12 μM) in buffer C (25 mM potassium phosphate, pH 8.0) was measured on a JASCO spectrometer (J-815) in the far-UV (200–260 nm) at 20 °C in 1 mm pathlength quartz cuvette. Each spectrum was obtained by averaging three individual scans. The spectrum of the buffer blank was subtracted from the protein CD spectrum. Deconvolution and analysis of the CD spectra were carried out using K2D2 software (http://www.ogic.ca/projects/k2d2; Perez-Iratxeta and Andrade-Navarro, 2007). Thermal denaturation profile of the protein was recorded using Circular dichroism ellipticity at 210 nm through variable-temperature scan in the temperature range 20–75 °C with a scan rate of 1 °C/min in two buffer conditions; buffer B and buffer C.

Intrinsic fluorescence of Cqm1 protein (20 μM) with or without 8 M urea was monitored on a JASCO spectrofluorometer (FP-8500) using a 1 cm cuvette, with excitation at 280 nm and emission spectrum recorded over 300–400 nm at 25 °C. Each spectrum was obtained by averaging three individual scans. Intrinsic fluorescence measurements were also performed with Cqm1 protein (20 μM) in buffer D [20 mM HEPES buffer (pH 7.5) 100 mM NaCl] in the presence of 5 mM EGTA. Each fluorescence spectrum was obtained by averaging three individual scans.

Thermofluor-shift melting curve measurements were carried out to study thermal stability of Cqm1 protein as described by Niesen et al. (2007). The measurements were performed in a 96 well plate in 25 μL reaction, in triplicates, containing recombinant Cqm1 protein (2 μM) in buffer C and freshly diluted SYPRO orange dye (1:1000, v/v). The plate was sealed and data was recorded on CFX96™ Real time system (Bio-Rad) in FRET mode. The protein samples were heated from 20 to 95 °C at ramp rate of 1.0 °C/min. Thermal unfolding of Hen Egg White (HEW) lysozyme was also probed in the same 96 well plate as an internal control. The observed T_m of 71.5 °C for HEW lysozyme matches with the values reported in literature. For probing thermal stability due to binding of Ca^{2+}, the measurements were also performed with Cqm1 (2 μM) in buffer D and freshly diluted SYPRO orange (1:1000, v/v), in the presence of EGTA (5 mM) or EGTA (5 mM) and $CaCl_2$ (0.5, 1.0 and 10.0 mM). Experimental data were processed using software available with CFX96™ Real time system and melting temperatures (T_m) were determined from the first derivative of each of the thermofluor-shift melting curve.

2.7. Enzymatic assay

α-glucosidase activity of Cqm1 protein was assayed by Michaelis–Menten kinetics using α-pNPG as substrate. The enzyme activity was determined in 800 μL reaction mix with 65 nM Cqm1 in buffer E (50 mM sodium phosphate buffer, pH 6.8) and 100 μM glutathione (reduced) with α-pNPG concentration varying from 0.05 to 4 mM. The kinetic reaction was measured at 25 °C over a period of 20 min for each substrate concentration with data recording at an interval of 2 min using 1 cm path-length quartz cuvette. The amount of p-nitrophenol produced during the course of the reaction was measured at 405 nm using the Time Course Measurement Program available with JASCO UV–Vis spectrophotometer. The measured absorbance at 405 nm (OD_{405}) was plotted against time, and initial velocity (Vo) for each substrate concentration was estimated from the gradient. The obtained values of Vo were plotted against substrate concentration. The kinetic constants (*V*max, *K*m and *k*cat) were calculated by non-linear regression with Michaelis-Menten equation using Origin software version 6.

Thin layer chromatography was performed to identify the specificity of Cqm1 protein against α-1,1-glycosidic, α-1,2-glycosidic, α-1,4-glycosidic and α-1,6-glycosidic bonds. Trehalose with α-1,1-glycosidic bond, sucrose with α-1,2-glycosidic bond, maltose and maltotriose with α-1,4-glycosidic bond, and isomaltose with α-1,6-glycosidic bond, were used as substrates. The polysaccharide glycogen was also used as substrate in an independent experiment. Cqm1 (5 μM) was incubated with individual sugars (1: 250, w/w) in 50 μL reaction volume at 37 °C for about 2 and 24 hr in independent reactions and also for about 5 h with glycogen and sucrose. Higher amounts of Cqm1, compared to kinetic analysis, were used to detect products in TLC experiments. The products from the reaction of Cqm1 with glucose, trehalose, sucrose, maltose, maltotriose, isomaltose and glycogen were resolved on silica gel based TLC plates using chloroform/acetic acid/water (3: 3.5: 0.5, v/v) as the solvent. The products were detected by incubation in an acetone solution of $AgNO_3$ and 40% NaOH/ethanol (1: 19 v /v).

2.8. Effect of temperature, pH, Ca^{2+} and BinA/B proteins on enzymatic activity

The effect of temperature on α-glucosidase activity of Cqm1 was investigated over a temperature range of 25–60 °C using 65 nM Cqm1 in buffer E. The effect of pH on α-glucosidase activity of Cqm1 (65 M) was analysed at 37 °C over a pH range of 4.6–11 using sodium acetate buffer (pH 4.6 and 5.6), potassium phosphate buffer (pH 5, 8 and 8.5), sodium citrate buffer (pH 6), sodium phosphate buffer (pH 6.8), HEPES buffer (pH 7 and 7.5) and CAPS buffer (pH 10 and 11). The ionic strength of different buffer systems used in this experiment may vary and the present study cannot rule out effect of such variation on activity. The effect of Ca^{2+} on α-glucosidase activity of Cqm1 protein was examined using Cqm1 (65 nM) in buffer D. Cqm1 protein was incubated with EGTA (5 mM) for about 1 hr and the excess of EGTA was removed by dialysis. The dialysed protein was incubated with 10 mM Ca^{2+} at 37 °C for about 1 hr prior to addition of substrate. For probing the effect of BinA/BinB proteins on enzymatic activity against substrate α-pNPG, Cqm1 (65 nM) was incubated with BinA or BinB proteins (300 nM in buffer E) for about 1 hr at 37 °C, prior to addition of the substrate. In the control experiments, BinA and BinB did not display observable hydrolytic activity against α-pNPG.

These activity assays were performed in 200 μl reaction volume using 100 μM glutathione (reduced) and substrate α-pNPG (1 mM). Incubation time for the enzymatic reaction was 20 min and the reaction was stopped by adding 800 μl of 100 mM sodium carbonate buffer. The absorbance of accumulated p-nitrophenol product was measured at 405 nm using JASCO UV–Vis spectrophotometer. Absorbance for each reaction was converted into activity units (μmol/min/mg). The optimum temperature and pH, and effect of Ca^{2+} and BinA/BinB, were determined by plotting the activity units.

2.9. Surface plasmon resonance analysis of BinA and BinB interactions

The instrument used for Surface plasmon resonance analysis was Biacore T200 Surface plasmon resonance (SPR) system with CM5 Biacore sensor chip. For each experiment at 25 °C nearly 1240 response units of Cqm1 (ligand) were immobilized onto the CM5 Chip via amine coupling followed by extensive washing with buffer F [50 mM potassium phosphate (pH 8) and 50 mM NaCl]. Different concentrations (0.039, 0.078, 0.156 and 0.3125 μM) of aqueous BinA and BinB (analytes) in buffer F were injected onto the Cqm1-bound sensor chip at 30 μL/min flow-rate in independent experiments. Final experiments for BinA were performed using higher concentrations of the protein (1.25, 2.5, 5, 10 and 20 μM), as lower concentrations did not results in significant response units (RU). The analyte was allowed to interact with the immobilized Cqm1 for 60 sec before washing off with buffer F. The data were processed and association/dissociation rate constants were analyzed using Biacore kinetic evaluation software (V3.1) provided with the instrument.

2.10. Structure modelling, molecular dynamics and bioinformatics analysis

The online servers of Modeller (http://toolkit.tuebingen.mpg.de/modeller; Eswar et al., 2008) and SWISS-MODEL (http://swissmodel.expasy.org; Biasini et al., 2014) were used to generate homology models for Cqm1 protein. Several templates with sequence identity between 30 and 36% with Cqm1 sequence were selected by both the servers. The accepted structural model corresponding to residues 23–556 of Cqm1 and a Ca^{2+} were constructed using the template structure of trehalulose synthase (MutB) from *Pseudo mesoacidophile* (PDB code, 1ZJA; Ravaud et al., 2007). Sequence identity of 34% (similarity 38%) between Cqm1 and MutB was returned by the servers. Molecular dynamics (MD) simulations of the constructed Cqm1 model were performed using OPLS-AA/L force field and GROMACS package (Pronk et al., 2013). The overall electric charge of the system was compensated by adding Na^{+} ions. The structural models derived from SWISS-MODEL with or without Ca^{2+}were subjected to initial energy minimization, 100 ps temperature equilibration at 298 K using Berendsen thermostat and 100 ps pressure equilibrium at 1 atm using Berendsen barostat. The production run trajectory was simulated for 40 ns in 10,000,000 steps of 4 fs each and structures were written every 10 ps.

3. Results and discussion

3.1. Purification of recombinant proteins

Attempts to purify recombinant Cqm1 protein of 580 amino acids using codon optimized synthetic gene and different expression systems, including GST and trigger factor translational fusion with Cqm1 sequence, were not successful. Three major problems were encountered: 1) protein expression was not observed, 2) total of expressed protein was in inclusion body, and 3) refolding of the protein from detergent solubilized inclusion body did not yield folded protein. The Cqm1 protein sequence was subjected to analysis for intrinsically disordered regions (Jones and Cozzetto, 2015; http://bioinf.cs.ucl.ac.uk/psipred/). The N-terminus 22 residues and C-terminus 20 residues were identified as disordered (Supplementary Fig. S1). The truncated *cqm*1 gene (1614 bp) lacking 66 bases from the 5′-end and 60 bases from the 3′-end was sub-cloned into pET28a based pNIC28-Bsa4 expression vector for purification of the protein using *E. coli* expression system. The N-terminal poly-His tagged recombinant Cqm1 protein, lacking N-terminus 22 signal peptide residues and 20 C-terminus GPI anchoring residues, was purified to homogeneity by three stage column chromatography methods; IMAC, AEC and SEC (Supplementary Fig. S2). The mass of 62378 Da determined by MALDI-TOF matches with estimated mass of 62329 Da deduced from the amino acid sequence encoded by the cloned gene. The peptide mass fingerprinting analysis confirms identity of the purified protein to be Cqm1 (Supplementary Fig. S3). Recombinant BinA (active form; 15–366 amino acids) and BinB (residues 17–448) were also purified to homogeneity using IMAC and AEC chromatography methods (Supplementary Fig. S2).

3.2. Folding and oligomeric status of recombinant Cqm1 protein

The secondary structure conformation of recombinant Cqm1 protein was characterized by CD spectroscopy (Fig. 1A). The % secondary structure elements determined with the help of K2D2 web-tool reveals mixed α/β structure (α, ~8%; β, ~40%). Thermal denaturation of secondary structure of Cqm1 over a temperature range 20–75 °C was carried out in two different buffer conditions; 10 mM Tris-HCl (pH 8), 10 mM NaCl, and 25 mM potassium phosphate buffer (pH 8). The protein does not show secondary structure denaturation in the probed temperature range in either of the buffer conditions.

Intrinsic protein fluorescence shows an emission maximum (λ_{EM}) at 335 nm, which is typical of partially buried tryptophan residues (Fig. 1B). Fluorescence spectral shift to red (λ_{EM} ~352 nm) and quenching of fluorescence intensity on incubation with 8M urea, owing to denaturation of the protein in the presence of chaotropic agent, suggest that the purified recombinant protein has folded tertiary structure. Also, protein fluorescence reduced marginally in the presence of EGTA that specifically chelates bound Ca^{2+}. Several members of GH13 have been reported to bind Ca^{2+}.

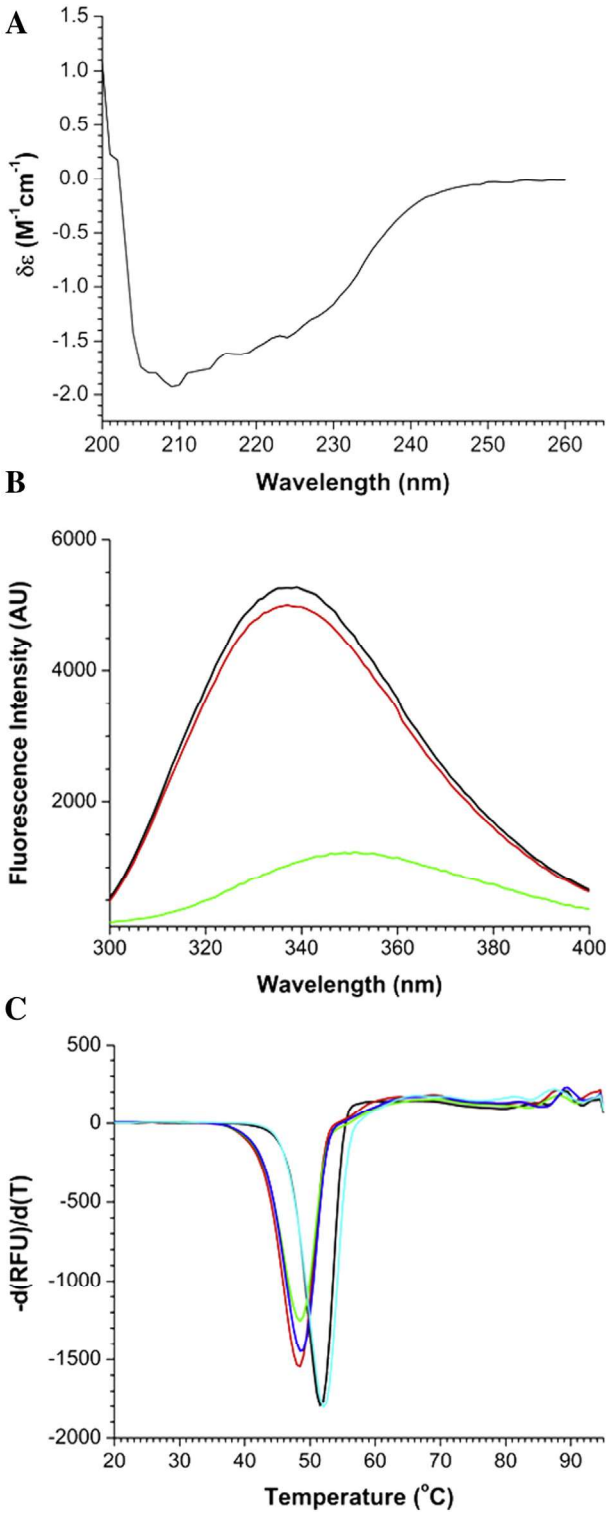


Fig. 1. Recombinant Cqm1 adopts native structure. **A**) CD spectrum of Cqm1 protein (0.75 mg/ml protein in 25 mM potassium phosphate, pH 8.0). Deconvolution of the spectrum reveals mixed α/β structure (α, ~8%; β, ~40%). **B**) An emission maximum at 335 nm in intrinsic fluorescence of Cqm1 (black) is shifted to 352 nm with loss of intensity on denaturation with 8M urea (green). Fluorescence intensity also decreases marginally in the presence of EGTA (5 mM; red). **C**) Derivative plot showing thermofluor-shift melting curve of Cqm1 (black), Cqm1 in the presence of 5 mM EGTA (red), and Cqm1 in the presence of EGTA (5 mM) and Ca^{2+} (0.5 mM, green; 1 mM, blue; 10 mM, cyan). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

The apparent melting temperature of the tertiary structure (T_m) of Cqm1 is estimated to be 51.5 °C by thermofluor-shift analysis in the presence of HEPES buffer (pH, 7.5) (Fig. 1C and Supplementary Fig. S4). Decrease in T_m (to 48 °C) is observed in the presence of EGTA. The addition of 10 mM Ca^{2+} rescued T_m to 52 °C suggesting that the divalent cation stabilizes the protein. A similar trend has also been observed for dextran glucosidase from *Streptococcus* mutant (Kobayashi et al., 2011). Influence of the buffer was also observed on T_m of Cqm1. The protein displays lower stability (T_m ~48 °C) in the presence of phosphate buffer (pH, 8.0).

The recombinant Cqm1 protein elutes as dimer on SEC column with an estimated molecular weight of 120 kDa suggesting homodimer status of the protein (Fig. 2A). The protein migrates as a single band on 10% native-PAGE with migration close to that of 140 kDa marker protein (Fig. 2B). A hydrodynamic radius of about 4.36 ± 1.73 nm corresponding to a molecular weight of about 106 kDa is estimated in dynamic light scattering analysis. Marginal increase in the hydrodynamic radius of the protein is observed in the presence of 0.05% triton and 500 mM sodium chloride (Fig. 2C). However, there was no significant change in the polydispersity in the presence of triton and salt (Supplementary Table S1). Data taken together suggests that Cqm1 exists in solution as a stable homodimer, and both hydrophobic and ionic residues may contribute towards dimeric interface. The stable oligomer of Cqm1 may be required for its apical localization in lipid rafts, as observed for other GPI-anchored proteins (Paladino et al., 2004). It has been suggested earlier that oligomer of Cpm1 (Cqm1 ortholog from *Culex pipiens*) promotes the opening of ionic pores by interacting with BinAB (Pauchet et al., 2005).

3.3. Structure modeling and molecular dynamics simulations

No structure of GH13_17 subfamily members is available in the Protein Data Bank. The homology models returned by MODELLER and SWISS-PDB were based on MutB template and were used for structure-activity analysis. The two models are similar with an RMSD of 0.49 Å for 436 pair of atoms on superposition in Chimera suite (Pettersen et al., 2004). The Cqm1 model (MODELLER scores: zDOPE score, −0.83; GA341 score, 1.0) is reliable, as active-site and Ca^{2+} binding residues of MutB align with the invariant residues of CAZy GH13_17 subfamily, and as stereochemical analysis of phi-psi dihedral angles indicates nearly 87.3% of residues in the most favored regions with none in the disallowed regions of the Ramachandran plot.

The structural models of Cqm1 reveal three independent domains (A–C), conserved Ca^{2+} binding residues (Asp-46, Asn-48, Asp-50 and Asp-54) and different conformations of loop regions (residues 241–245, 273–283, 306–311, 322–327, 369–373, 399–406, 509–517, 531–533). The domain A is a $(\beta/\alpha)_8$-barrel (residues 23–128 and 197–471), domain B is a loop between the β3 strand and α3 helix of domain A (residues 129–196), and domain C is an extension at the C-terminus characterized by a Greek key conformation constituted by residues 472–556 (Fig. 3A). The active site scaffold is constituted by Asp-85, Tyr-88, His-128, Arg-222, Asp-224, Ala-225, Glu-290, Phe-318, His-357, Asp-358 and Arg-407 of domain A (catalytic domain) (Supplementary Fig. S5). These residues are invariant in the characterized proteins of CAZy GH13_17 subfamily. Catalytic triad (Asp-224, Glu-290 and Asp-358) can be expected to perform catalysis.

The Ca^{2+} binding residues reside in the loop-1 connecting β1 and α1 elements of the $(\beta/\alpha)_8$-barrel in the structural model and are strictly conserved in GH13_17 subfamily. We observe that EGTA inhibits α-glucosidase activity of Cqm1 protein completely (Fig. 5B) and Ca^{2+} also enhanced thermal stability of the tertiary structure, as adjudged from thermofluor-shift analysis (Fig. 1C). During the 40

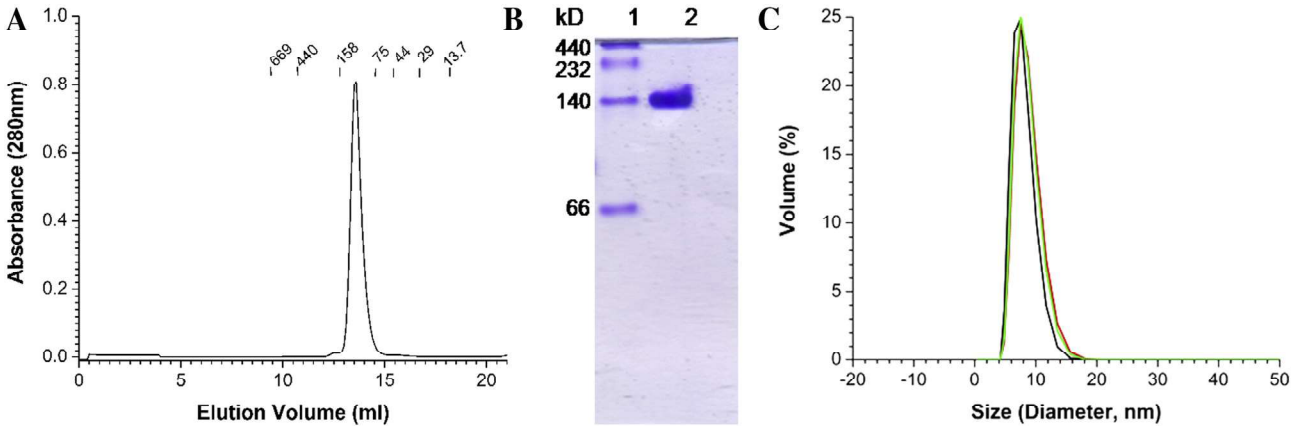


Fig. 2. Oligomeric status of Cqm1. **A**) Size exclusion chromatography elution profile on Superdex 200 column in buffer A (25 mM Tris-HCl pH 8, 100 mM NaCl). Elution volume for the standard protein molecular weight markers (Ribonuclease A, 13.7 kDa; Carbonic anhydrase, 29 kDa; Ovalbumin, 44 kDa; Conalbumin, 75 kDa; Aldolase, 158 kDa; Ferritin, 440 kDa; Thyroglobulin, 669 kDa) are also shown. **B**) 10% native-PAGE analysis; Lane 1, protein markers for native-PAGE; Lane 2, Cqm1 protein eluted from Superdex 200 column. The positions of molecular weight markers are identified. **C**) Dynamic light scattering analysis of Cqm1 (black), in the presence of 500 mM sodium chloride (red), and in the presence of 0.05% triton (green). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

ns molecular dynamics (MD) simulation, loop defining Ca^{2+} binding site is found to maintain its conformation when divalent cation is not bound to it. The peptide fragments showing diversity in the two homology models also show higher flexibility in MD simulations. Notably, loop-4 between β4 and α4 elements in the catalytic domain of Cqm1 and consisting of residues 236–253 shows high RMSF of about 6 Å (Fig. 3B). However, despite higher flexibility it does not shield the active-site. Compared to *Halomonas* sp. H11 α-glucosidase, Cqm1 sequence harbours 6 residue deletions in the loop-4 (Supplementary Fig. S5; Shen et al., 2015).

3.4. Interaction of BinA/B proteins with Cqm1

Interaction between Cqm1 and BinA/B proteins were evaluated in real-time with Surface plasmon resonance. The Cqm1 (ligand) was immobilized on a CM5 Biacore chip and BinA/B proteins were used in mobile phase. The interaction of the two toxin components with receptor Cqm1 was confirmed by the increase in SPR signal as a function of concentration. The results showed increased mass density at the sensor surface due to BinB binding (Fig. 4A). The binding curves could be explained with model that corresponds to 1:1 interaction of Cqm1 and BinB with excellent fit between experimental data and the fitted curve (Chi^2, 0.0284). The association and dissociation rate constants (*k*a and *k*d) are $7.12 \times 10^4\ M^{-1}\ s^{-1}$ and $0.0007\ s^{-1}$ that suggests strong binding with equilibrium dissociation constant (*K*d) value of 9.8 nM. This matches well with the earlier reported *K*D of 7–20 nM for BinB binding to larvae midgut preparations (Nielsen-LeRoux and Charles, 1992; Charles et al., 1997). In comparison, BinA does not show interaction with Cqm1 at lower concentrations. The interaction of Cqm1 and BinA are detected at higher concentration of BinA (Fig. 4B) with estimated *k*a and *k*d values of $1.49 \times 10^4\ M^{-1}\ s^{-1}$ and $0.032\ s^{-1}$ that results in *K*d value of 2.2 μM (1:1 interaction of Cqm1 and BinA with Chi^2 value of 0.237). This also agrees well with the IC_{50} of about 2 μM observed for BinA with brush border membrane fractions (Charles et al., 1997). The SPR analysis thus suggests specific and strong interaction of BinB with Cqm1, while interaction of BinA with Cqm1 may be non-specific and transient. Significantly retarded mobility of

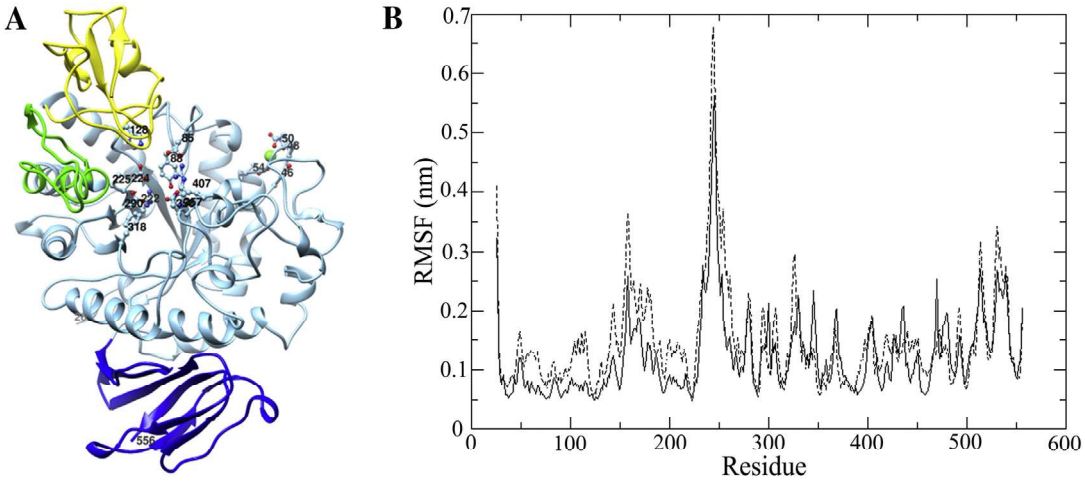


Fig. 3. Structural model and MD simulations. **A**) Ribbon model of Cqm1 protein (domain A, cyan; domain B, yellow and domain C, blue) showing also active-site and Ca^{2+} binding residues in ball-and-stick representation. Ca^{2+} is shown as green sphere. The loop-4 between β4 and α4 elements is shown in green color. **B**) The root mean square fluctuation (RMSF) of 23–558 residues about the average position during the 40 ns of MD simulations is shown. RMSF values for Cqm1 without Ca^{2+} is shown in solid line and with bound $Ca^{2+}$is given in dash line. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

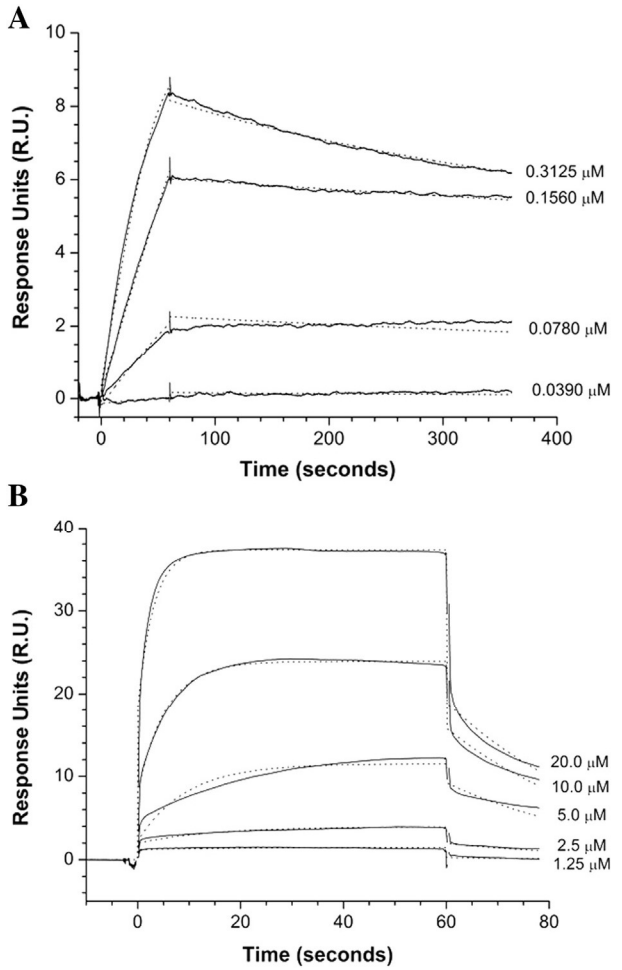


Fig. 4. Interaction of Cqm1 with BinAB toxin components. Nearly 1240 response units of Cqm1 were immobilized on to CM5 Biacore chip that was subsequently washed thoroughly with buffer F, and different concentrations of the analyte (BinA or BinB) were injected over Cqm1 surface in independent experiments. The response was monitored for 60 secs before washing with buffer F. **A**) BinB as analyte. **B**) BinA as analyte. The black trace (solid line) at each concentration indicates experimental determination. The interactions could best be modeled by a 1:1 interaction model (dotted traces) in both the studies.

Cqm1 in the presence of BinB on 10% native-PAGE and independent migration of BinA and Cqm1 supports these results (Supplementary Fig. S6).

3.5. Enzymatic activity of Cqm1 protein

The Michaelis-Menten kinetics parameters towards α-pNPG were estimated from nonlinear regression fit to the velocity-substrate profile (Fig. 5A). The maximum velocity (*V*max), affinity constant (*K*m) and catalytic constant (*k*cat) for α-pNPG hydrolysis are estimated to be 1.81 μmol $min^{-1}mg^{-1}$, 0.44 mM and 1.9 s^{-1}, respectively. The low *k*cat value observed for Cqm1 is in agreement with the values generally observed for maltases (Deng et al., 2014). The resulting specific activity and catalytic efficiency (*k*cat/*K*m) of Cqm1 are calculated to be 3.2 Units/mg, and 4318 $M^{-1}s^{-1}$, respectively. The observed catalytic efficiency of Cqm1 compares favorably with the homologous protein of *Aedes aegypti*, but is much higher than the value of about 40 $M^{-1}s^{-1}$ reported for the *Culex* protein expressed along with un-structured signal peptide residues (Nascimento et al., 2017).

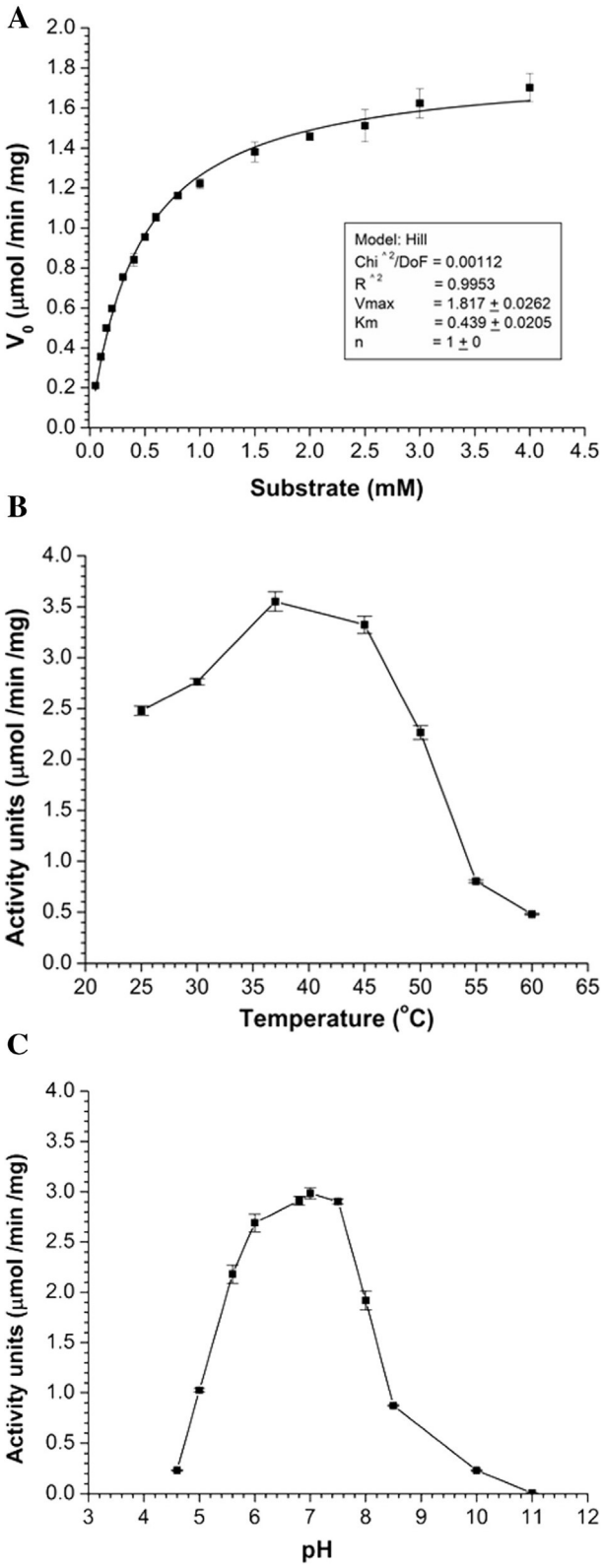


Fig. 5. Kinetic analysis of α-glucosidase activity of Cqm1 protein. A) The Michaelis-Menten plot using Cqm1 (65 nM) at pH 6.8 and temperature 25 °C is shown. The kinetic parameters were obtained by nonlinear regression fit to the Michaelis-Menten equation using Origin software. **B**) Plot depicting variation of enzymatic activity at pH 6.8 with varying temperature (25–60 °C) using Cqm1 (65 nM) and substrate (1 mM). **C**) Plot displaying variation of enzymatic activity with varying pH (4.6–11) at 37 °C using Cqm1 (65 nM) and substrate (1 mM). The vertical bars represent standard errors observed in three independent experiments. The substrate α-pNPG was used for all the experiments.

The optimum temperature for hydrolytic activity against α-pNPG is estimated to be about 37 °C. The activity increases steadily from 70% at 25 °C to the maximum at 37 °C, but declines at higher temperatures above 50 °C (Fig. 5B). The Cqm1 protein displays significant hydrolytic activity at pH values above 5 with the maximum activity at pH value of about 7.2. Better than 70% activity is observed within the pH range of 6.0–8.0. However, enzyme activity rapidly decreases at pH value above 8.0 and Cqm1 is completely inactive at alkaline pH value of 11.0 (Fig. 5C). The observed pH optimum relates well to the physiological conditions found in the posterior half of larval midgut (pH of 7–8 in *Culex*), where fluorescently labelled BinAB were observed to co-localize (Dadd, 1975; Lekakarn et al., 2015).

The enzymatic activity of Cqm1 against α-pNPG is not affected in the presence of BinA. BinA displays larvicidal activity (Hire et al., 2009; Sharma et al., 2017) and displays transient interaction with Cqm1. However, BinB, which binds with high affinity to Cqm1, reduces α-glucosidase activity by nearly 50% (Fig. 6B). Since BinA is known to be the toxic factor, it rules out the possibility discussed earlier (Silva-Filha et al., 1999) that larvicidal activity may be due to inhibition of receptor biochemical activity.

Recombinant Cqm1 enzyme was incubated with different disaccharides and glycogen at 37 °C for different time intervals, and reaction products were analysed by Thin Layer Chromatography (TLC) to probe substrate specificity. Cqm1 fails to act on trehalose and isomaltose (Fig. 7A). Instead, Cqm1 shows efficient transglycosylation activity with maltose and maltotriose, as is evident from the accumulation of maltose-oligomers (Fig. 7B). Cqm1 also displays weak hydrolytic activity against sucrose, and higher molecular mass oligomers of sucrose are not observed even after 24 hr incubation. However, higher accumulation of fructose is observed when glucose was added to Cqm1/sucrose reaction mixture (Supplementary Fig. S6A). Low molecular weight saccharides, such as glucose, are observed in reaction with glycogen after about 5 hr (Supplementary Fig. S6B). The data taken together suggests that Cqm1 is competent to cleave α-1,4- and α-1,2-glycosidic bonds and synthesizes new α-1,4-linkages with shorter maltodextrins to form maltose-oligomers. Conservation of Ala-225 and Phe-318 residues is consistent with the hydrolytic activity against α-1,4-linkages. The α-1,6-glucosidases have Val at position equivalent to Ala-225 of Cqm1. The higher activity against holoside substrates (maltose and maltotriose), and poor activity against sucrose and α-aryl glucoside (α-pNPG), suggest that Cqm1 α-glucosidase should be characterized as type II or type III-glucosidase. Type II and III enzymes are usually catalogued in GH31 family of CAZy database, however.

The formation of glucosyl-enzyme intermediate resulting from the hydrolysis of glucose-oligomers precedes transglycosylation step, which seems to be preferred over hydrolysis of glucosyl-enzyme intermediate. Higher accumulation of maltose-oligomers is observed in reaction mixture with shorter incubation time of 2 hr (Fig. 7A), compared to longer incubation period of 24 hr (Fig. 7B). Also, maltose-oligomers, and not glucose, are observed in 20 min incubation reaction of Cqm1 with maltose (Supplementary Fig. S6C). This may be due to depletion of maltose-oligomer substrates of Cqm1 over longer time spans. Hydrolysis of glycogen and transglycosylation of maltose-oligomers are consistent with open and accessible active site for binding of large substrates. The short loop-4 between β4 and α4 elements in the catalytic domain permits binding of maltose-oligomers and glycogen in the active site. In comparison, longer loop-4 in *Halomonas* sp. H11 α-glucosidase sterically inhibits binding of maltose-oligomers in the active site during transglycosylation step (Supplementary Fig. S4; Shen et al., 2015).

Cqm1 displays amylomaltase activity, which catalyses transglycosylation by successive transfer of glucose moiety of maltose onto the malto-oligosaccharides, and sucrase activity, which hydrolyses sucrose to glucose and fructose. Similar activities have also been reported for α-glucosidase from *Acyrthosiphon pisum* and *Apiscerana japonica* which also belong to GH13_17 subfamily (Price et al., 2007; Wongchawalit et al., 2006). The appearance of low-molecular weight saccharides in Cqm1/glycogen reaction mixture also suggests that the protein may be involved in hydrolysis of high-molecular-mass maltose-oligomers. Thus, Cqm1 or its orthologs in *Aedes* and *Anopheles* species can be expected to participate in carbohydrate metabolism in mosquitoes to meet the high-energy demand for reproduction and for harbouring disease-causing vectors. Also, α-glucosidases seem to be essential for the partitioning of carbohydrates from the diet into carbon nutrition and osmo-regulation in insects (Ashford et al., 2000; Karley et al., 2005). Further studies can reveal if Cqm1 or its orthologs in *Aedes* and *Anopheles* can be targeted for mosquito control programs, as *Culex* mosquitoes incompetent to present Cqm1 on cell surface are known to be viable.

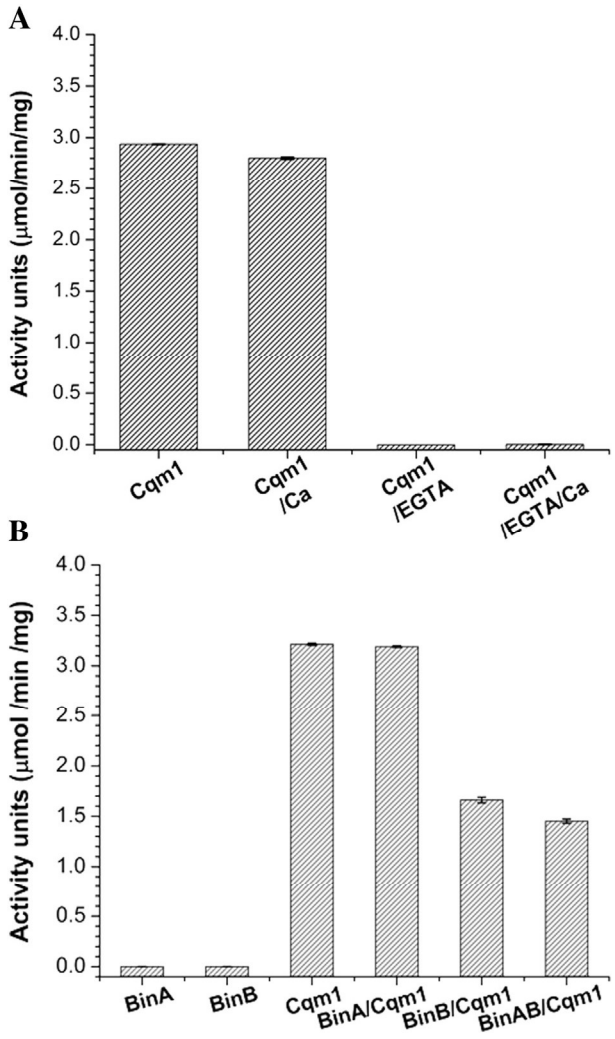


Fig. 6. Inhibition of Cqm1 α-glucosidase activity. A) Enhancement of Cqm1 (65 nM) enzymatic activity was not observed with the addition of Ca^{2+} (Cqm1/Ca). However, EGTA (5 mM) totally inhibited α-glucosidase activity (Cqm1/EGTA). Addition of Ca^{2+} (10 mM) to the dialysed fraction of Cqm1/EGTA did not rescue the activity (Cqm1/EGTA/Ca). **B**) α-glucosidase activity of Cqm1 was probed in the presence of BinA (BinA/Cqm1), BinB (BinB/Cqm1), and equimolar mixture of BinA/BinB (BinAB/Cqm1) proteins. BinA and BinB individually were also used in control experiments without Cqm1. α-glucosidase activity was examined at 37 °C and pH 7.5 using α-pNPG (1 mM) as a substrate in both the experiments. The error bars show standard errors from three independent experiments.

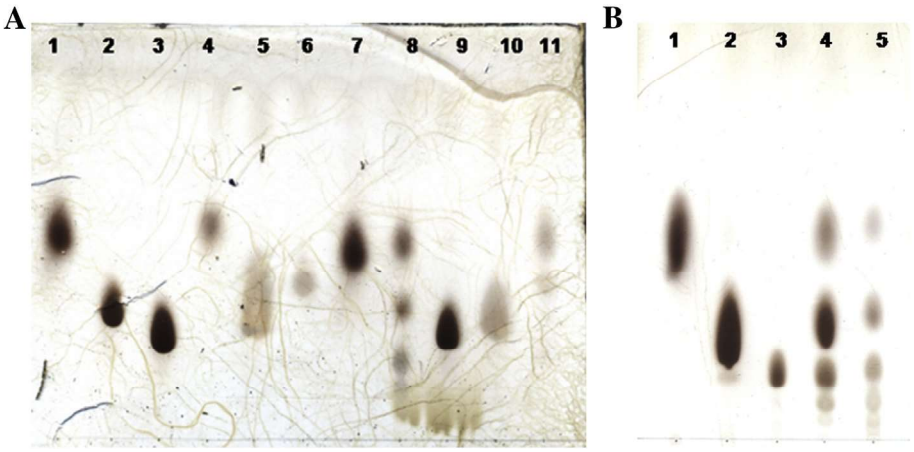


Fig. 7. Thin layer chromatography analysis. A) Lane 1, glucose; Lane 2; maltose; Lane 3, isomaltose; Lane 4, fructose; Lane 5, trehalose; Lane 6, sucrose. Glucose, maltose, isomaltose, trehalose and sucrose were also incubated individually with Cqm1 (5 μM) for overnight at 37 °C and reaction mixtures were resolved on TLC in lanes 7, 8, 9, 10 and 11, respectively. **B**) Lane 1, glucose; Lane 2; maltose; Lane 3, maltotriose. Maltose and maltotriose were individually incubated with Cqm1 (5 μM) at 37 °C for 2 hr and mixtures were resolved on TLC in lanes 4 and 5, respectively.

4. Conclusion

We have characterized *Culex* Cqm1 protein which is thought to be essential for mosquito-larvicidal activity of BinAB toxin. The recombinant Cqm1 protein exists as a stable dimer which supports its apical localization in lipid rafts. The protein catalyses transglycosylation by successive transfer of glucose moiety of maltose onto the maltose-oligosaccharides and is also competent to hydrolyse sucrose and glycogen. Cqm1 can be expected to participate in carbohydrate metabolism in mosquitoes, which makes it a valuable target for world-wide mosquito control programmes.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgment

We thank Drs. Ramesh Hire and Ashok Hadapad, Nuclear Agriculture & Biotechnology Division, for maintaining nucleus culture of *Culex quinquefasciatus* mosquitoes, Dr. Sanjeeva Srivastava, IIT Bombay, for the help with Mass spectrometry experiments, and Dr. Shubhendu Seal, GE Healthcare Life Sciences, for the help with SPR experiments. Mahima Sharma acknowledges award of DAE PhD fellowship.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ibmb.2017.12.002.

References

Ashford, D.A., Smith, W.A., Douglas, A.E., 2000. Living on a high sugar diet: the fate of sucrose ingested by a phloem-feeding insect, the pea aphid *Acyrthosiphon pisum*. J. Insect. Physiol. 46, 335–341.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L., Schwede, T., 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. 42, 252–258.

Chalegre, K.D., Tavares, D.A., Romão, T.P., de Menezes, H.S., Nascimento, N.A., de Oliveira, C.M., de-Melo-Neto, O.P., Silva-Filha, M.H., 2015. Co-selection and replacement of resistance alleles to *Lysinibacillus sphaericus* in a *Culex quinquefasciatus* colony. FEBS J. 282, 3592–3602.

Charles, J.F., Silva-Filha, M.H., Nielsen-LeRoux, C., Humphreys, M.J., Berry, C., 1997. Binding of the 51- and 42-kDa individual components from the *Bacillus sphaericus* crystal toxin to mosquito larval midgut membranes from *Culex* and *Anopheles* sp. (Diptera: Culicidae). FEMS Microbiol. Lett. 156, 153–159.

Chiba, S., 1988. α-Glucosidases. In: Handbook of Amylases and Related Enzymes. The Amylase Research Society of Japan. Pergamon Press, Oxford, United Kingdom, pp. 104–105.

Colletier, J.P., Sawaya, M.R., Gingery, M., Rodriguez, J.A., Cascio, D., Brewster, A.S., Michels-Clark, T., Hice, R.H., Coquelle, N., Boutet, S., Williams, G.J., Messerschmidt, M., DePonte, D.P., Sierra, R.G., Laksmono, H., Koglin, J.E., Hunter, M.S., Park, H.W., Uervirojnangkoorn, M., Bideshi, D.K., Brunger, A.T., Federici, B.A., Sauter, N.K., Eisenberg, D.S., 2016. De novo phasing with X-ray laser reveals mosquito larvicide BinAB structure. Nature 539, 43–47.

Dadd, R.H., 1975. Alkalinity within the midgut of mosquito larvae with alkaline-active digestive enzymes. J. Insect. Physiol. 21, 1847–1853.

Darboux, I., Nielsen-LeRoux, C., Charles, J.F., Pauron, D., 2001. The receptor of *Bacillus sphaericus* binary toxin in *Culex pipiens* (Diptera: Culicidae) midgut: molecular cloning and expression. Insect biochem. Mol. Biol. 31, 981–990.

Darboux, I., Pauchet, Y., Castella, C., Silva-Filha, M.H., Nielsen-LeRoux, C., Charles, J.F., Pauron, D., 2002. Loss of the membrane anchor of the target receptor is a mechanism of bioinsecticide resistance. Proc. Natl. Acad. Sci. U. S. A. 99, 5830–5835.

Darboux, I., Charles, J.F., Pauchet, Y., Warot, S., Pauron, D., 2007. Transposon mediated resistance to *Bacillus sphaericus* in a field-evolved population of *Culex pipiens* (Diptera: Culicidae). Cell. Microbiol. 9, 2022–2029.

Davidson, E.W., 1988. Binding of the *Bacillus sphaericus* toxin to midgut cells of mosquito (Diptera: Culicidae) larvae: relationship to host range. J. Med. Entomol. 25, 151–157.

Deng, X., Petitjean, M., Teste, M.A., Kooli, W., Tranier, S., François, J.M., Parrou, J.L., 2014. Similarities and differences in the biochemical and enzymological properties of the four isomaltases from *Saccharomyces cerevisiae*. FEBS Open Bio 4, 200–212.

Eswar, N., Eramian, D., Webb, B., Shen, M.Y., Sali, A., 2008. Protein structure modeling with MODELLER. Methods Mol. Biol. 426, 145–159.

Guo, Q.Y., Cai, Q.X., Yan, J.P., Hu, X.M., Zheng, D.S., Yuan, Z.M., 2013. Single nucleotide deletion of cqm1 gene results in the development of resistance to *Bacillus sphaericus* in *Culex quinquefasciatus*. J. Insect. Physiol. 59, 967–973.

Hire, R.S., Hadapad, A.B., Dongre, T.K., Kumar, V., 2009. Purification and characterization of mosquitocidal *Bacillus sphaericus* BinA protein. J. Invertebr. Pathol. 101, 106–111.

Jones, D.T., Cozzetto, D., 2015. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. Bioinformatics 31, 857–863.

Kale, A., Hire, R.S., Hadapad, A.B., D'souza, S.F., Kumar, V., 2013. Interaction between mosquito-larvicidal *Lysinibacillus sphaericus* Binary toxin components: analysis of complex formation. Insect. biochem. Mol. Biol. 43, 1045–1054.

Karley, A.J., Ashford, D.A., Minto, L.M., Pritchard, J., Douglas, A.E., 2005. The significance of gut sucrase activity for osmoregulation in the pea aphid, *Acyrthosiphon pisum*. J. Insect Physiol. 51, 1313–1319.

Kato, N., Suyama, S., Shirokane, M., Kato, M., Kobayashi, T., Tsukagoshi, N., 2002. Novel alpha-glucosidase from *Aspergillus nidulans* with strong transglycosylation activity. Appl. Environ. Microbiol. 68, 1250–1256.

Kobayashi, M., Hondoh, H., Mori, H., Saburi, W., Okuyama, M., Kimura, A., 2011. Calcium ion-dependent increase in thermostability of dextran glucosidase from *Streptococcus* mutants. Biosci. Biotechnol. Biochem. 75, 1557–1563.

Lacey, L.A., 2007. *Bacillus thuringiensis* serovariety *israelensis* and *Bacillus sphaericus* for mosquito control. J. Am. Mosq. Control. Assoc. 23, 133–163.

Lekakarn, H., Promdonkoy, B., Boonserm, P., 2015. Interaction of *Lysinibacillus sphaericus* binary toxin with mosquito larval gut cells: binding and internalization. J. Invertebr. Pathol. 132, 125–131.

Lombard, V., GolacondaRamulu, H., Drula, E., Coutinho, P.M., Henrissat, B., 2014. The Carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, 490–495.

Lowry, O.H., Rosebrough, N.J., Farr, A.L., Randall, R.J., 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. 193, 265–275.

Nascimento, N.A., Ferreira, L.M., Romão, T.P., Correia, D.M., Vasconcelos, C.R., Rezende, A.M., Costa, S.G., Genta, F.A., de-Melo-Neto, O.P., Silva-Filha, M.H., 2017. N-glycosylation influences the catalytic activity of mosquito α-glucosidases associated with susceptibility or refractoriness to *Lysinibacillus sphaericus*. Insect biochem. Mol. Biol. 81, 62–71.

Nielsen-LeRoux, C., Charles, J.F., 1992. Binding of *Bacillus sphaericus* binary toxin to a specific receptor on midgut brush-border membranes from mosquito larvae. Eur. J. Biochem. 210, 585–590.

Niesen, F.H., Berglund, H., Vedadi, M., 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat. Protoc. 2, 2212–2221.

Oei, C., Hindley, J., Berry, C., 1992. Binding of purified *Bacillus sphaericus* binary toxin and its deletion derivatives to *Culex quinquefasciatus* gut: elucidation of functional binding domains. J. Gen. Microbiol. 138, 1515–1526.

Paladino, S., Sarnataro, D., Pillich, R., Tivodar, S., Nitsch, L., Zurzolo, C., 2004. Protein oligomerization modulates raft partitioning and apical sorting of GPI-anchored proteins. J. Cell Biol. 167, 699–709.

Pauchet, Y., Luton, F., Castella, C., Charles, J.F., Romey, G., Pauron, D., 2005. Effects of a mosquitocidal toxin on a mammalian epithelial cell line expressing its target receptor. Cell Microbiol. 7, 1335–1344.

Perez-Iratxeta, C., Andrade-Navarro, M.A., 2007. K2D2: estimate of protein secondary structure from circular dichroism spectra. BMC Struct. Biol. 8, 25.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

Price, D.R., Karley, A.J., Ashford, D.A., Isaacs, H.V., Pownall, M.E., Wilkinson, H.S., Gatehouse, J.A., Douglas, A.E., 2007. Molecular characterization of a candidate gut sucrose in the pea aphid, *Acyrthosiphon pisum*. Insect biochem. Mol. Biol. 37, 307–317.

Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M.R., Smith, J.C., Kasson, P.M., van der Spoel, D., Hess, B., Lindahl, E., 2013. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 29, 845–854.

Ravaud, S., Robert, X., Watzlawick, H., Haser, R., Mattes, R., Aghajari, N., 2007. Trehalulose synthase native and carbohydrate complexed structures provide insights into sucrose isomerization. J. Biol. Chem. 282, 28126–28136.

Romão, T.P., de Melo Chalegre, K.D., Key, S., Ayres, C.F., Fontes de Oliveira, C.M., de-Melo-Neto, O.P., Silva-Filha, M.H., 2006. A second independent resistance mechanism to *Bacillus sphaericus* binary toxin targets its alpha-glucosidase receptor in *Culex quinquefasciatus*. FEBS J. 273, 1556–1568.

Sharma, M., Hire, R.S., Hadapad, A.B., Gupta, G.D., Kumar, V., 2017. PEGylation enhances mosquito-larvicidal activity of *Lysinibacillus sphaericus* binary toxin. Bioconjug. Chem. 28, 410–418.

Shen, X., Saburi, W., Gai, Z., Kato, K., Ojima-Kato, T., Yu, J., Komoda, K., Kido, Y., Matsui, H., Mori, H., Yao, M., 2015. Structural analysis of α-glucosidase HaG provides new insights into substrate specificity and catalytic mechanism. Acta Crystallogr. D. Biol. Crystallogr. 71, 1382–1391.

Silva-Filha, M.H., Nielsen-LeRoux, C., Charles, J.F., 1999. Identification of the receptor for *Bacillus sphaericus* crystal toxin in the brush border membrane of the mosquito *Culex pipiens* (Diptera: Culicidae). Insect biochem. Mol. Biol. 29, 711–721.

Stam, M.R., Danchin, E.G., Rancurel, C., Coutinho, P.M., Henrissat, B., 2006. Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins. Protein Eng. Des. Sel. 19, 555–562.

Wongchawalit, J., Yamamoto, T., Nakai, H., Kim, Y.M., Sato, N., Nishimoto, M., Okuyama, M., Mori, H., Saji, O., Chanchao, C., Wongsiri, S., Surarit, R., Svasti, J., Chiba, S., Kimura, A., 2006. Purification and characterization of alpha-glucosidase I from Japanese honeybee (*Apiscerana japonica*) and molecular cloning of its cDNA. Biosci. Biotechnol. Biochem. 70, 2889–2898.

World Health Organization, 1985. Informal Consultation on the Development of *Bacillus Sphaericus* as a Microbial Larvicide. TDR/BCV/Sphaericus/85.3. World Health Organization, Geneva.

Bioconjugate Chemistry

Article

pubs.acs.org/bc

PEGylation Enhances Mosquito-Larvicidal Activity of *Lysinibacillus sphaericus* Binary Toxin

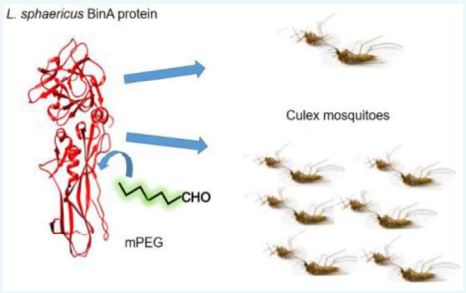
Mahima Sharma,[†,§] Ramesh S. Hire,[‡] Ashok B. Hadapad,[‡] Gagan D. Gupta,[†] and Vinay Kumar*[,†,§]

[†]Protein Crystallography Section, Radiation Biology & Health Sciences Division and [‡]Nuclear Agriculture & Biotechnology Division, Bhabha Atomic Research Centre, Mumbai 400085, India

[§]Homi Bhabha National Institute, Training School Complex, Anushaktinagar, Mumbai 400094, India

Supporting Information

ABSTRACT: Toxic strains of *Lysinibacillus sphaericus* have been used in the field for larval control of mosquito vector diseases. The high toxicity of *L. sphaericus* is attributed to the binary (BinAB) toxin produced as parasporal crystalline inclusions during the early stages of sporulation. BinA and BinB, the primary components of these spore-crystals, exert high toxicity when administered together. However, instability, short half-lives, and rapid proteolytic digestion can limit their use as an effective insecticide. BinA alone displays larvicidal toxicity, in the absence of BinB, albeit with much reduced activity. Here for the first time, we demonstrate the beneficial effect of PEGylation (covalent attachment of polyethylene glycol) on mosquito-larvicidal activity of BinA. Polymer conjugation was achieved using 750 Da polyethylene glycol (PEG) at two different pH values (pH 7.2 and 8.5). Two different isoforms of the biopolymers, purified to homogeneity, were highly water-soluble and resistant to trypsin and proteinase K. The mono-PEGylated BinA isoforms also displayed preservation of the toxin structure with improved thermal stability by about 3−5 °C, as evident from thermal denaturation studies by circular dichroism and differential scanning fluorimetry. Notably, PEGylation enhanced BinA toxicity by nearly 6-fold. The PEGylated BinA isoforms alone displayed high larvicidal activity (LC_{50} value of ∼3.4 ng/mL) against the third instar *Culex* larvae, which compares favorably against LC_{50} reported for the combination of BinA and BinB proteins. Since BinA can be synthesized easily through recombinant technology and easily PEGylated, the conjugated biopolymers offer a promising opportunity for mosquito control programs.



INTRODUCTION

The highly toxic strains (like 2362, C3−41) of *Lysinibacillus sphaericus*, a Gram-positive soil bacterium, are used as a larvicide to control populations of mosquitoes, like *Culex* and *Anopheles*, which transmit some of the deadly human diseases such as West Nile fever, Japanese encephalitis, and malaria.[1,2] *Aedes* have been reported refractory to the toxic effects of *L. sphaericus* binary toxin.[3] The high mosquito larvicidal activity of *L. sphaericus* is attributed to the binary (BinAB) toxin produced in the form of parasporal crystalline inclusions during the early stages of sporulation.[1] Binary toxin has the advantage of being highly specific to mosquitoes as targets, eliminating the risks of any effects on nontarget organisms, including other mosquito predators, chironomids and vertebrates, and thus can be considered environmentally friendly.[4−6] Binary toxin is composed of two component proteins: toxic BinA (Mr, 41.9 kDa) and receptor binding BinB (Mr, 51.4 kDa). From crystals of protoxins, upon ingestion, alkalinity in the larval midgut (pH 8−11) releases BinAB, the components of which are activated by the larval gut proteases.[7,8] Both BinA and BinB peptides are required to exert the maximum toxicity.[9] Studies have also shown that BinA alone, without BinB, shows weak toxicity against *Culex* larvae.[10−12]

Improvement in the efficacy of such protein-based biopesticides have been achieved by recombinant DNA technology by increasing their synthesis and/or enabling the production of a combination of toxins from different bacteria in a single strain. For instance, combination of BinAB and Bti toxins in the same bacterial species showed improved mosquito-larvicidal activity.[13] However, instability, short half-lives, and rapid digestion by the proteases of peptide based biopolymers calls for their frequent administration at the breeding sites and limits their use as insecticide.[14] Modification of the useful biopesticides enhancing their stability and toxicity could provide viable means to use these more effectively in mosquito control programs.

PEGylation of therapeutic proteins and peptides, by covalent linking of polyethylene glycol chains to the target molecules, is proving to be a highly potent tool in their real-life applications.[15,16] The most relevant advantages of PEGylation are known to be prolonged bioavailability, increase in stability, and reduced immunogenicity.[17−21] There are currently over 10 different PEGylated products approved by the U.S. FDA,

Received: October 1, 2016
Revised: December 29, 2016
Published: January 25, 2017

DOI: 10.1021/acs.bioconjchem.6b00565
Bioconjugate Chem. 2017, 28, 410−418

including PEG-adenosine deaminase and PEG-asparaginase.[20] Also, the use of aliphatic PEG polymer in enhancing insecticidal activity of proteins was suggested by Jeffers et al., who observed an enhancement of nearly 10-fold in toxicity of a modified decapeptide against *Aedes aegypti*.[22]

In the present study, we report conjugation of methoxy polyethylene glycol isocyanate to the toxic BinA component of mosquito-larvicidal BinAB toxin. We characterized the resulting conjugate using biophysical approaches, like MALDI-TOF, dynamic light scattering (DLS), differential scanning fluorimetry (DSF), and circular dichroism (CD), for assessing extent of modification and thermal stability of the biopolymer. The modification was found to improve larvicidal efficacy by nearly 6-fold and enhanced thermal stability of secondary and tertiary structure of the protein. The PEG conjugated BinA protein was also found to be resistant to broad-spectrum proteases. This is the first study probing the effect of PEGylation on the activity of a safe and universally used *L. sphaericus* BinAB toxin. The modification of recombinantly produced BinA can be easily achieved and can be helpful in reducing dependency on less stable BinB protein to achieve maximal mosquito-larvicidal activity. This is in line with WHO recommendations on improvement and efficient implementation of mosquito control interventions.

RESULTS

Purification of BinAB Proteins. The recombinant BinA (active form; 15−366 amino acids) was expressed in BL21 (DE3) host cells as fusion protein with N-terminal poly-His tag and a thrombin cleavage site between poly-His tag and BinA using *p*ET28a-*bin*A construct.[12] The recombinant protein was purified from soluble cell lysate using IMAC affinity chromatography and High Q anion exchange chromatography (AEC). The BinA protein was found to exist as monomer, as was evident from its elution profile on size exclusion chromatography (SEC) column. The mixture of pro-BinA (41.9 kDa) and pro-BinB (51.4 kDa) proteins (BinAB) was also purified to homogeneity from spore-crystals of highly toxic local strain of *L. sphaericus* ISPC-8 using ion-exchange and size-exclusion chromatography.[23]

Synthesis and Purification of PEGylated BinA. The purified recombinant BinA was used for the PEGylation reaction with methoxy polyethylene glycol isocyanate of MW 750 Da (mPEG-ISC-750) at two different pH values, pH 7.2 (m1BinA) and pH 8.5 (m2BinA). PEG-isocyanate derivative is useful for amine group conjugation, yielding stable urethane linkage. The reaction depends upon the nucleophilicity of the amine group. At pH 7.2, the N-terminus amine group can be expected to react with isocyanate due to its low pK_a (7.7 ± 0.5) generally found in proteins.[24] At pH 8.5, however, the amine group of accessible lysine residues (pK_a 10.4 ± 1.1) can also form a covalent linkage. After overnight reaction, the PEG conjugated proteins were purified through SEC using Superdex-200 column. Monomeric PEGylated BinA constituted the largest fraction (Figure 1) for each of the PEGylation reaction. Highly pure PEG conjugated BinA proteins were obtained using anion exchange chromatography and were used for further analysis.

Electrophoretic Analysis. The purified recombinant BinA and PEGylated BinA (m1BinA and m2BinA) proteins were electrophoresed under native as well as denaturating conditions. In the Coomassie stained 15% SDS-PAGE, the PEG conjugates showed marginal shift in the molecular weight toward the higher side (Figure 2A). The presence of PEG moiety in the PEGylated BinA isoforms was, however, confirmed with the Western blot analysis using anti-polyethylene glycol and anti-PEG (methoxy group) antibodies

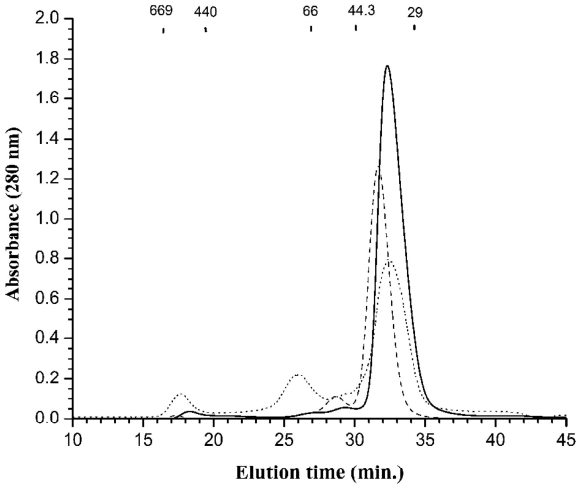


Figure 1. Purification of PEGylated BinA by SEC. Elution profiles of m1BinA (dotted line) and m2BinA (dash line) from Superdex-200 column. The elution profile of recombinant BinA (solid line) is also overlaid. In each elution profile the major peak corresponds to monomer of BinA and contains most of the BinA protein, as adjudged from SDS-PAGE analysis. The elution times of standard molecular markers are also shown. The minor peaks contain small amounts of oligomeric BinA, as adjudged by 12% SDS-PAGE.

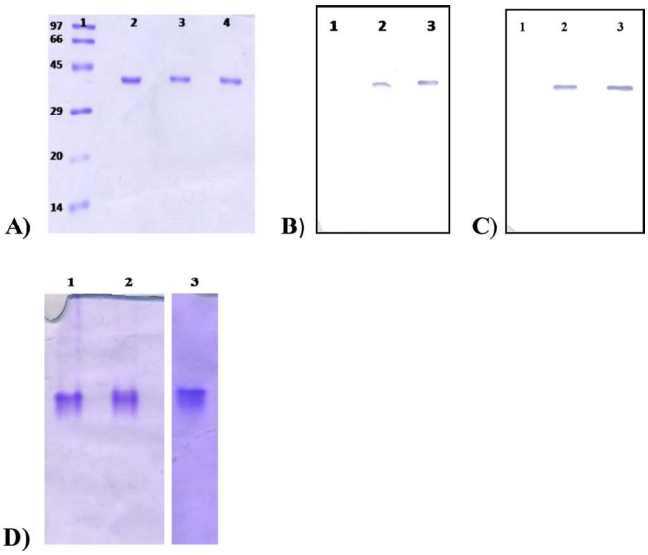


Figure 2. Electrophoretic analysis of PEGylated BinA after anion exchange chromatography. (A) 15% SDS-PAGE. Lane 1, Protein molecular weight markers; Lane 2, purified recombinant BinA; Lane 3, purified m1BinA isoform; Lane 4, purified m2BinA isoform. (B,C) Western blot analysis of m1BinA and m2BinA proteins resolved on 15% SDS-PAGE using anti-polyethylene glycol antibodies and anti-PEG (methoxy group) antibodies, respectively. The immune detection was visualized by secondary antibody and BCIP/NBT color development reagents. Unmodified BinA (lane 1) was not immunostained, while PEGylated m1BinA (Lane 2) and m2BinA (Lane 3) were clearly visible. (D) 12% native PAGE of recombinant and PEG conjugated BinA proteins. Lane 1, BinA; Lane 2, m1BinA; Lane 3, m2BinA.

DOI: 10.1021/acs.bioconjchem.6b00565
Bioconjugate Chem. 2017, 28, 410−418

(Figure 2B,C). The three proteins showed similar mobility on the native PAGE (Figure 2D). The single bands of m1BinA and m2BinA in the native PAGE suggest homogeneous-polymer modification of BinA at both the pH values.

The SDS-PAGE gel was also stained with barium iodide by the I_2−$BaCl_2$ method.[25,26] The PEGylated BinA, however, did not stain differentially using the barium-iodide method, as staining of unmodified BinA was also observed each time (SI Figure S1). We assume that the small size of the PEG moiety used in the study could not be easily detected by the traditional PEG staining method.

MALDI-TOF and DLS Analysis. The PEGylated BinA isoforms and recombinant BinA were further analyzed by MALDI-TOF to determine the state of PEGylation. The analysis showed single large spectral peak at m/z of 41 044 Da and 41 053 for the PEG conjugated BinA proteins at pH 7.2 and 8.5, respectively, whereas unmodified recombinant BinA showed a spectral peak at m/z of 40 152 Da (Figure 3). The

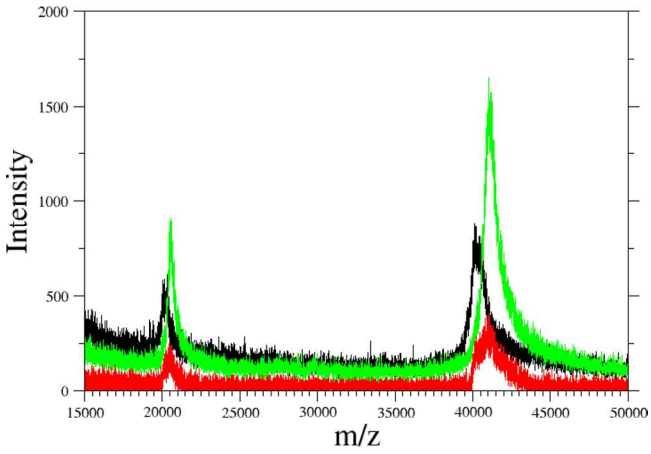


Figure 3. MALDI-TOF analysis of native and PEG conjugated BinA proteins. The MALDI-TOF analysis of recombinant BinA (black) and PEGylated BinA proteins, m1BinA (red) and m2BinA (green), showed large spectral peaks at m/z values of 40 152 Da for BinA, 41 044 Da for m1BinA, and 41 053 Da for m2BinA. Smaller peaks at about 20 500 Da correspond to doubly charged species of these proteins.

observed mass of 40 152 Da for BinA matches well with the expected mass of 40 142 Da estimated for the amino acid sequence of the protein that is composed of BinA residues 15−366 and four residues (GSHM) at the N-terminus from the poly-His tag after thrombin treatment of the expressed protein. Further, the observed mass differences between PEGylated and unmodified BinA suggest single site PEGylation of BinA at both pH values. The observation of a single peak for both the PEG-BinA conjugates also suggested that the purification protocol used was effective in obtaining homogeneous PEG polymer of modified BinA.

The DLS experiments also confirmed PEG conjugation resulting in a marginal increase in the hydrodynamic size of the modified proteins, compared to recombinant BinA. The hydrodynamic diameters for the unmodified, m1BinA, and m2BinA proteins were estimated to be 5.8 ± 0.37, 5.98 ± 0.98, and 6.1 ± 1.3 nm, respectively (SI Figure S2). The proteins were monodispersive with the estimated polydispersity of 5%, 12.5%, and 24%, respectively, suggesting that the oligomeric status of BinA did not change on modification.

Toxicity Analysis. Toxicity of recombinant BinA, PEGylated BinA proteins, and BinAB mixture purified from *L. sphaericus* spore-crystals was assayed against the third instar larvae of *Culex quinquefasciatus* from the same generation and was performed on the same day. The recombinant BinA showed toxicity against the larvae with an LC_{50} value of 21.1 ng/mL. However, interestingly, PEGylated BinA isoforms, without BinB, displayed superior insecticidal characteristics with LC_{50} values of about 3.4 ng/mL (Table 1). In comparison,

Table 1. LC_{50} Values of Recombinant BinA (BinA), PEGylated BinA Isoforms (m1BinA and m2BinA), and BinAB Mixture Purified from *L. sphaericus* Spore-Crystals against the Third Instar *Culex quinquefasciatus* Larvae

		95% confidence values			
protein	LC_{50} value (ng/mL)	lower	higher	χ^2-value	*P*-value
BinA	21.1	16.1	27.5	1.65	0.65
m1BinA	3.1	2.0	4.6	12.6	0.005
m2BinA	3.7	2.2	6.1	4.8	0.08
BinAB	6.5	4.5	9.5	32.2	<0.005

an LC_{50} value of 6.5 ng/mL was observed for the BinAB mixture. However, toxicity was not observed in the control experiments using PEGylated BSA against *Culex* and for m2BinA against *Aedes* larvae.

PEGylation Enhances Thermal Stability of BinA Protein. To probe if PEG conjugation with BinA induced conformational changes in the protein structure, secondary structure was estimated for the recombinant and PEGylated proteins using CD spectroscopy (Figure 4A). The deconvolution of CD data points, computed using the K2D2 software,[27] revealed BinA and its PEGylated isoforms to be mainly composed of β-structure with nearly 48% residues in β-strand conformation and nearly 2% residues in α-helices. The similarity of CD spectra of the three proteins suggested that the secondary structure of BinA was maintained on PEGylation. Also, positive ellipticity for the three proteins was observed in the wavelength range 245−225 nm, which probably is due to tryptophan residues as observed for HCAII enzyme.[28]

The thermal denaturation of the recombinant and PEGylated BinA proteins was studied using CD by monitoring ellipticity at 218 nm during temperature-wavelength scans. The loss of secondary structures was adjudged from the complete CD spectrum at different temperatures. The temperature range showing loss of secondary structure was used to estimate apparent transition temperature from the first derivative of CD values at 218 nm. The unmodified BinA protein showed a T_m of 55 °C, whereas the m1BinA isoform exhibited a T_m of 61 °C and m2BinA isoform showed a T_m of 55 °C (Figure 4B). The denaturation profiles were found to be buffer sensitive, which is not unexpected. For instance, in the phosphate buffer recombinant BinA and m2BinA proteins showed T_m of ∼65 °C, and m1BinA did not show loss of secondary structure until the temperature of about 75 °C in the temperature scan at 218 nm.

The influence of PEG conjugation on the stability of the tertiary structure of BinA protein was studied using DSF. The thermal denaturation was monitored in the presence of SYPRO orange dye, which is highly fluorescent in the nonpolar environment. Fluorescence was recorded as a function of temperature (Figure 5A) and the first derivative was used to

DOI: 10.1021/acs.bioconjchem.6b00565
Bioconjugate Chem. 2017, 28, 410−418

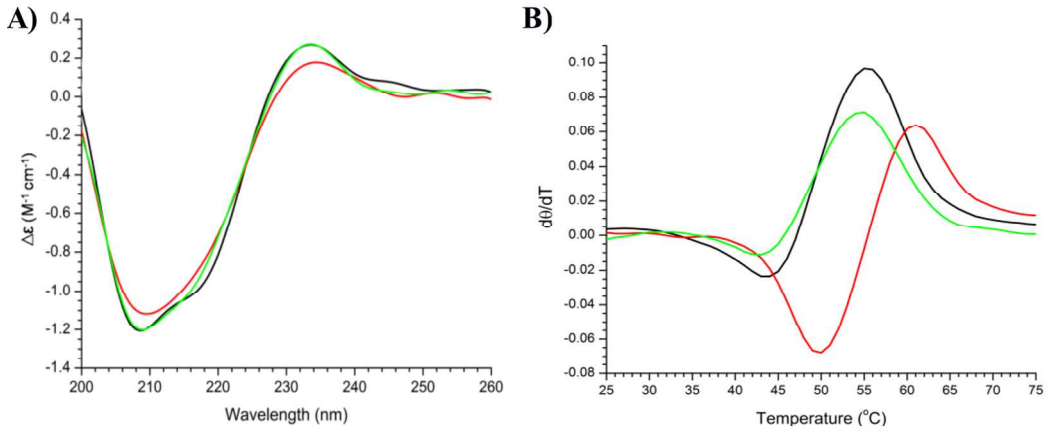


Figure 4. Analysis by CD spectroscopy. (A) CD spectra of BinA (black), m1BinA (red), and m2BinA (green) proteins at 25 °C. The contribution due to buffer was subtracted and data were smoothed with FFT algorithm with 15 points using Origin software. (B) Thermal denaturation of BinA (black) and PEGylated BinA isoforms (m1BinA, red; m2BinA, green) was measured from the first derivative of the observed ellipticity at 218 nm in temperature-wavelength scans.

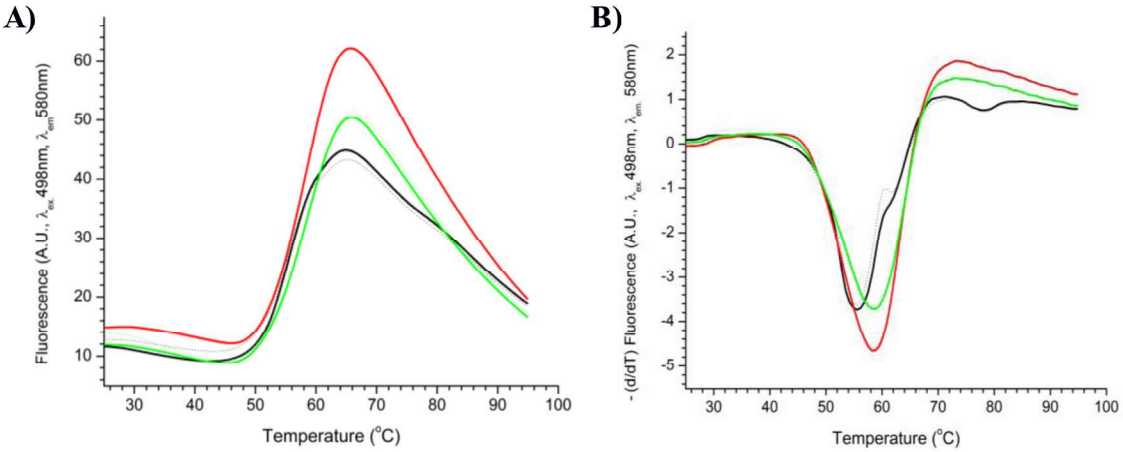


Figure 5. Thermal denaturation of the BinA proteins by Differential Scanning Fluorimetry. (A) Variation in fluorescence intensity of SYPRO orange as a function of temperature was monitored in duplicate in the presence of recombinant BinA (black lines) and PEG conjugated BinA isoforms (m1BinA, red; m2BinA, green). The λ_{ex} and λ_{em} were 498 and 580 nm, respectively. (B) First derivative of observed fluorescence was plotted as a function of temperature to determine the apparent transition temperature, T_m, which were estimated to be 55.5, 58.2, and 58.8 °C for recombinant BinA, m1BinA isoform, and m2BinA isoform, respectively.

estimate the apparent melting temperature of the tertiary structure (Figure 5B). The melting temperatures determined for recombinant BinA and PEG-BinA conjugates at pH 7.2 and 8.5 were 55.5, 58.2, and 58.8 °C, respectively. These values can be treated with high confidence as the observed T_m of 69.5 °C for HEW lysozyme (SI Figure S3), used as an internal control in the DSF experiment, matched exactly with the literature value.

The significant increases of about 6 and 3 °C in the midpoint of denaturation of BinA secondary and tertiary structures after PEGylation, respectively, clearly indicated enhanced thermal stability of BinA protein on PEG conjugation.

BinA Displays Higher Stability against Proteolytic Enzymes. The stability of recombinant BinA, PEGylated BinA derivatives, and pro-BinA/pro-BinB purified from spore-crystals was assayed in complex proteolytic environments using three proteases (trypsin, chymotrypsin, and proteinase K) in independent experiments. The small shift in molecular weight for both pro-BinA and pro-BinB proteins on incubation with trypsin (Figure 6A) is expected from reports in the literature, representing an activation step.[7] Compared to pro-BinA, pro-BinB was found to be susceptible to proteolytic cleavage by proteinase K even with 1 h incubation, as evident from the decrease in BinB band intensity and appearance of multiple protein bands migrating with higher mobility (Figure 6A, lane 4). The recombinant BinA and PEG-BinA conjugates were found to be stable against trypsin and broad-spectrum proteinase K proteases even upon overnight incubation at protease:protein ratio as high as 1:10 for trypsin and 1:50 for proteinase K (Figure 6C). All the proteins, however, were found to be susceptible to degradation by chymotrypsin protease upon overnight incubation (Figure 6).

DISCUSSION

Mosquitoes are known to spread many human diseases like West Nile fever, dengue, malaria, zika, and chikungunya, accounting for millions of death annually. *Culex* is known to transmit West Nile virus and Japanese encephalitis.[29] Binary

DOI: 10.1021/acs.bioconjchem.6b00565
Bioconjugate Chem. 2017, 28, 410−418

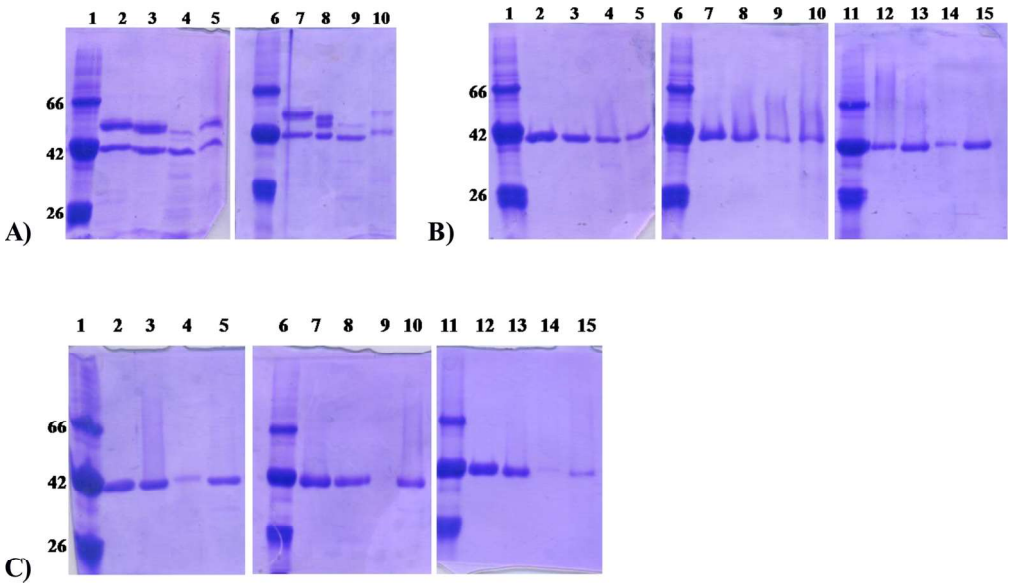


Figure 6. Assessment of protein stability against proteolytic enzymes. (A) 12% SDS-PAGE analysis of pro-BinA and pro-BinB proteins purified from spore crystals of *L. sphaericus* after incubation with proteases for 1 h (Lanes 3, 4, and 5) and after overnight incubation with proteases (Lanes 8, 9, and 10). Lanes 1 and 6, molecular weight markers; Lanes 2 and 7, proteins without proteases; Lanes 3 and 8, incubation with trypsin; Lane 4 and 9, incubation with proteinase K; Lanes 5 and 10, incubation with chymotrypsin. (B) 12% SDS-PAGE analysis of proteolytic stability of recombinant BinA (Lanes 2, 3, 4, and 5), m1BinA isoform (Lanes 7, 8, 9, and 10), and m2BinA isoform (Lanes 12, 13, 14, and 15) after 1 h incubation with proteases. Lanes 1, 6, and 11, molecular weight markers; Lanes 2, 7, and 12, proteins without proteases; Lanes 3, 8, and 13, incubation with trypsin; Lanes 4, 9, and 14, incubation with chymotrypsin; Lanes 5, 10, and 15, incubation with proteinase K. (C) 12% SDS-PAGE analysis of proteolytic stability of recombinant BinA (Lanes 2, 3, 4, and 5), m1BinA (Lanes 7, 8, 9, and 10), and m2BinA (Lanes 12, 13, 14, and 15) after overnight incubation with proteases. Lanes 1, 6, and 11, molecular weight markers; Lanes 2, 7, and 12, proteins without proteases; Lanes 3, 8, and 13, incubation with trypsin; Lanes 4, 9, and 14, incubation with chymotrypsin; Lanes 5, 10, and 15, incubation with proteinase K.

toxin from *L. sphaericus* displays mosquito-larvicidal activity against *Culex* and *Anopheles*, and has been found to be safe for a vast range of nontarget organism.[4] Studies have shown that BinA component of the binary toxin, alone, displays toxicity without BinB, albeit at reduced levels.[10−12]

The study presented here is the first reported evaluation of the effect of PEGylation on the activity of BinA larvicidal protein. Recombinant BinA was PEGylated using 750 Da methoxy polyethylene glycol isocyanate as the PEGylation agent. Two PEGylated BinA isoforms, synthesized at pH 7.2 and 8.5, respectively, were purified to homogeneity through size exclusion and anion exchange chromatography. The two-stage column chromatography purification of PEG-BinA conjugates ensured recovery of homogeneous polymer-modified BinA isoforms and complete removal of unreacted excess of PEGylation reagent.

The elution profile from size exclusion chromatography, mobility on native PAGE, and DLS experiments showed that the PEGylated BinA isoforms exist in the same monomeric state as the unmodified recombinant protein (Figures 1,2, and SI Figure S2). The positive immunostaining with anti-polyethylene glycol and anti-PEG (methoxy group) antibodies, and the observed increased mass in MALDI-TOF experiments confirmed single site modification of recombinant BinA with the PEG moiety (Figures 2 and 3). Spectroscopic studies further revealed that chemical conjugation did not induce observable conformational change in the protein structure. The CD spectra and deconvolution suggested similarity in the secondary structures of the proteins with nearly 48% residues in β structure and 2% in α-helices. The estimated secondary structure matches reasonably with the secondary structure (42% β-structure and 7% α-helices) in the recently resolved X-ray structure of the active region of BinA protein.[8]

One of the expected results of PEGylation is the altered thermal stability of target protein. The thermal stability of PEGylated BinA protein was investigated using CD and DSF experiments (Figures 4 and 5). An increase of nearly 6 °C in the denaturation temperature of secondary structure and an increase of nearly 3 °C in the transition temperature of tertiary structure, upon PEGylation in one of the isoform (m1BinA), clearly indicated the enhancement in the thermal stability of the modified BinA. The other isoform (m2BinA) though did not show any enhancement in T_m of the secondary structure; yet, its tertiary structure was more stable than the unmodified BinA protein by 3 °C. Since both the PEG-BinA conjugates appear to be mono-PEGylated, the obtained differences in the T_m values together with changes in DLS profile may suggest differences in the site of PEGylation. Also, the observed thermal stability may be independent of the site of PEG attachment.

The PEGylated BinA proteins were assayed for mosquito-larvicidal activity against third instar *Culex* larvae. The PEGylated BinA isoforms alone, without BinB, exhibited significantly enhanced median LC_{50} value of ~3.4 ng/mL, compared to LC_{50} values of 21.1 ng/mL for the recombinant (unmodified) BinA protein (Table 1). Nearly the same lethal concentration was observed for the two isoforms of PEGylated BinA and these compare favorably against the LC_{50} value for BinAB used worldwide. The LC_{50} values for recombinant BinA differ marginally from the previously[12] reported toxicity (LC_{50} 16.5 ng/mL), and may be due to the variation in mosquito population. The observed increase of nearly 6-fold in BinA mosquito-larvicidal activity is the first evidence of improving protein biopesticidal efficacy using PEGylation. The exact

DOI: 10.1021/acs.bioconjchem.6b00565
Bioconjugate Chem. 2017, 28, 410−418

mechanism for the increased larvicidal activity of polymer modified BinA is not clear, but may involve enhanced accumulation of active toxin due to higher stability and bioavailability.

The proteolytic stability analysis for recombinant BinA and PEGylated BinA isoforms, and mixture of pro-BinA and pro-BinB purified from spore-crystals against trypsin, chymotrypsin, and proteinase K, revealed that BinA is highly stable against trypsin and broad-spectrum proteinase K (Figure 6). BinB, however, was found to be susceptible to degradation by these proteases and can be adjudged to be unstable in complex proteolytic environments.

Taken together, these studies suggest that PEGylated BinA alone, without BinB, can be an effective biological control agent. BinA protein can be purified on a large scale through recombinant technology, and easily PEGylated. Field trials are required further to adjudge usefulness of this approach in control of mosquito vector diseases.

■ CONCLUSION

In conclusion, we discuss a new approach of increasing the mosquito-larvicidal activity of *L. sphaericus* binary toxin by PEGylating the toxic BinA protein component. The PEGylated protein isoforms shows a significant 6-fold enhancement in its activity against *Culex* larvae, which may not depend on the site of PEG attachment, and their toxicity matches favorably against the combination of BinA and BinB proteins. The homogeneously modified biopolymers also display preservation of the toxin structure with improved thermal stability.

■ EXPERIMENTAL PROCEDURES

Materials. The mPEG-ISC-750 for PEGylation reaction was from NANOCS, USA. Lysozyme, barium chloride, and IPTG were from SRL, Mumbai, India. The Ni-IDA matrix was from GE Healthcare, India. Other chemical reagents and culture media used were from Himedia, India. SYPRO Orange Protein Gel Stain and trypsin and chymotrypsin proteases were from Sigma, and proteinase K was from NEB. Anti-polyethylene glycol and anti-PEG (methoxy group) antibodies were procured from Genetex. Alkaline phosphatase conjugated anti-rabbit IgG and NBT/BCIP were procured from Sigma and Roche, respectively. Bovine serum albumin and anti-His antibodies were obtained from Sigma.

Methods. *Purification of Bin Proteins.* The active BinA with fused N-terminus poly-His tag was expressed and purified from BL21 (DE3) cells as described earlier[12] (shall be referred as recombinant BinA hereafter). Briefly, BL21(DE3) cells transformed with pET28a-*bin*A construct, carrying truncated *bin*A gene (1056 bp) coding for residues 15−366 of BinA that constitutes active protein, were grown in LB medium supplemented with kanamycin (50 μg/mL) at 37 °C. The culture was transferred to 18 °C when it achieved an OD600−0.2 and was grown further. The protein expression was induced with 0.25 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at OD600−0.7 and cultured further for 4 h. The cells were harvested by centrifugation at 10 000*g* for 10 min at 4 °C and were resuspended in the lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl, 1 mM PMSF, 15% sucrose, 5% glycerol, 2 mM DTT) containing lysozyme (2 mg/mL). The suspension was subsequently sonicated in pulse mode (5s ON/10s OFF, 15 s each) followed by centrifugation at 21 000*g* for 30 min at 4 °C. The supernatant was loaded onto the pre-equilibrated Ni-IDA (nickel-iminodiacetic acid) matrix. The matrix was washed with buffer (50 mM Tris pH 8.0, 100 mM NaCl, and 25 mM imidazole). The bound protein was eluted using imidazole linear gradient (25−1000 mM) in buffer (50 mM Tris pH 8.0, 100 mM NaCl) over six column volumes. The buffer of the eluted protein was exchanged to buffer A (50 mM Tris pH 8.0, 20 mM NaCl) by dialysis. The recombinant protein was treated with thrombin at 37 °C overnight to remove the poly-His tag. Removal of the tag was confirmed by Western blotting using anti-His antibodies. The protein was further purified by anion exchange chromatography (High Q column, BioRad). The column was pre-equilibrated with buffer A. The bound proteins were eluted using a linear gradient of NaCl (20−1000 mM) in buffer (50 mM Tris pH 8.0) over six column volumes. The eluted fractions were adjudged on 12% SDS-PAGE. The purified recombinant BinA protein was dialyzed to 1× PBS at pH 7.2 and 8.5 for subsequent reactions and stored at 4 °C.

The purification of BinAB mixture (pro-BinA and pro-BinB proteins) from bacterial spore-crystals was achieved using highly toxic *L. sphaericus* ISPC-8 having serotype 5a5b and phage type 3. The ISPC-8 strain was grown until sporulation and the spore pellet collected by centrifugation. The purification of BinAB from the spore pellet was carried out using ion-exchange and size exclusion chromatography as described by Hire et al.[23] The final purification of BinAB mixture was achieved by size exclusion chromatography using Superdex-200 column (GE Healthcare).

PEGylation of Recombinant BinA. Methoxy polyethylene glycol isocyanate of MW 750 Da (PEGylation reagent) was used for BinA modification. Isocyanate derivatives of PEGs react readily at pH range of 7.0 to 9.0 to form covalent bond with free amino groups on the protein. The PEGylation reagent (reconstituted in DMSO) was added in excess to the purified recombinant BinA in the molar ratio of 1:50 at pH 7.2 and 8.5 in 1× PBS buffer in independent reactions. The reactions were allowed to proceed overnight at 4 °C with continuous stirring. The PEGylated BinA proteins, m1BinA (PEG-BinA conjugation at pH 7.2) and m2BinA (BinA modification at pH 8.5), were stored at 4 °C.

Purification and Oligomeric Status of PEGylated BinA Proteins. The PEGylated BinA proteins were purified through size exclusion chromatography using Superdex-200 column in order to separate differentially modified BinA molecules and for complete removal of unreacted excess of PEGylation reagent. The column was calibrated using standard molecular marker proteins (Bovine thyroglobulin, 669 kDa; Apoferritin, 440 kDa; Bovine serum albumin, 66 kDa; Ovalbumin, 44 kDa; Carbonic anhydrase, 29 kDa). The column was pre-equilibrated with 1× PBS buffer (pH 7.2) for purification of m1BinA and with 1× PBS buffer (pH 8.5) for m2BinA purification. The protein was loaded with the help of a syringe and the eluted fractions were collected and adjudged on 12% SDS-PAGE. The protein fractions eluted corresponding to monomeric state of recombinant BinA (~42 kDa) were collected and were dialyzed against buffer A. The dialyzed proteins were loaded onto pre-equilibrated High Q AEC column. The bound proteins were eluted with linear gradient (20−1000 mM NaCl) over six column volumes. The eluted PEGylayted BinA proteins were dialyzed and stored in the buffer B (10 mM Tris pH 8.0, 10 mM NaCl) at 4 °C.

The recombinant and PEGylated BinA proteins in buffer B were analyzed by dynamic light scattering using about 0.3 mg/mL protein concentration at 25 °C with Malvern Zetasizer

DOI: 10.1021/acs.bioconjchem.6b00565
Bioconjugate Chem. 2017, 28, 410−418

Nano ZS instrument. Three sets of measurement with 12 acquisitions in each set were collected. The data analysis was done with the help of Malvern Zetasizer software suite (http://www.malvern.com).

Electrophoretic and MALDI-TOF Analyses. The PEGylated BinA proteins, after final purification through AEC, were loaded onto 15% reducing SDS-PAGE gel and 12% native PAGE gel along with the purified recombinant BinA protein. After electrophoresis, the gels were stained with Coomassie solution. To visualize PEGylated proteins, SDS-PAGE was completely destained overnight and stained using the barium iodide method from Kurfurst.[25] Briefly, the gel was soaked in 5% glutaraldehyde solution for 15 min at room temperature and was subsequently stained for the PEG moiety. First, the gel was maintained in 20 mL perchloric acid (0.1 M) for 15 min. Then 5 mL of 5% barium chloride solution and 2 mL of 0.1 M iodine solution were added according to the procedure as described by Skoog.[26]After about 15 min, the staining solution was replaced and gel was washed with water for another 15 min.

An unstained 15% SDS-PAGE gel containing recombinant BinA, m1BinA, and m2BinA proteins was also electro-blotted on nitrocellulose membrane and was challenged with anti-polyethylene glycol and anti-PEG (methoxy group) antibodies using the antibody concentrations recommended by the manufacturer. Alkaline phosphatase-conjugated anti-rabbit IgG was employed as a second antibody, and color development was performed using 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and nitro-blue tetrazolium (NBT).

The PEGylated and recombinant BinA proteins were also analyzed by MALDI-TOF to determine their molecular weights and extent of modification. The protein masses were estimated from the highest observed peak intensity in each spectrum.

Toxicity Assays. The toxicity assays were performed for BinAB purified from *L. sphaericus*, recombinant BinA, and PEGylated BinA proteins (m1BinA and m2BinA). Prior to the toxicity assay, protein concentrations were determined by modified Bradford assay.[30] Different concentrations of the protein samples were tested against the third instar larvae of *Culex quinquefasciatus*. The nucleus culture of *C. quinquefasciatus* was obtained from National Institute of Virology, Pune and was maintained at 28 ± 2 °C and 85% relative humidity in our laboratory. Eggs were allowed to hatch in plastic bowls containing 1 L tap water supplemented with 0.13 g sterilized larval food (13:6:1, wheat flour:chickpea flour:yeast extract).[31] Five concentrations each for the BinAB, recombinant BinA, and PEGylated BinA proteins were tested in 10 mL of sterile distilled water containing ten third instar larvae with three replicates for each concentration. The experiments were done twice for each concentration on different days with the same batch of larvae to minimize variation between the larval batches, due to origin and seasons. As a control for the effect of PEG, bovine serum albumin (BSA) was modified by mPEG-ISC-750 (SI Figure S4) and PEGylated BSA was assayed for toxicity against *Culex* larvae using concentration of 200 ng/mL. No toxicity was observed in the control experiments. The toxicity of m2BinA was also assayed against third instar *Aedes* larvae using protein concentration of 100 ng/mL. All the proteins tested for toxicity were constituted in a low salt buffer (10 mM Tris, pH 8, and 10 mM NaCl).

The larval mortality was recorded after 48 h and was corrected by Abbott's method[32] using mortality observed in water and buffer only control experiments, each having 10 larvae, with three replicates. Mortality data was analyzed using the Probit analysis[33] and the median lethal concentration (LC_{50}) along with the 95% confidence interval were calculated with the help of the R project for statistical computing (https://www.r-project.org/).[34]

Estimation of Secondary Structure and Thermal Stability by Circular Dichroism. The effect on the secondary structure of BinA protein upon PEGylation was investigated using far-UV circular dichroism (CD). The spectra were recorded for the PEGylated BinA proteins (10 μM) and recombinant BinA (10 μM) in buffer B on a JASCO spectrometer (J-815) equipped with Peltier thermostatic cell holder. The data were recorded in 200−260 nm range at 20 °C in 1 mm quartz cuvette. Each spectrum was obtained by averaging three individual scans. The spectrum of the buffer blank was subtracted from the sample CD spectrum. Deconvolution and analysis of the CD spectra were carried out using K2D2 software available online (http://www.ogic.ca/projects/k2d2).[27]

Further, thermal denaturation measurements of recombinant and PEGylated BinA were performed using ellipticity at 218 nm from temperature-wavelength scan in the temperature range 20−75 °C with a scan rate of 1 °C/min. The CD spectra during temperature-wavelength scans were recorded in 200−260 nm wavelength range at regular temperature intervals (every increase of 5 °C between 20−50 and 2 °C thereafter). The transition temperature (T_m) of secondary structures was determined from the first derivative of temperature-wavelength scan spectra at 218 nm, as suggested by Greenfield.[35] For recording thermal denaturation in 20 mM phosphate buffer (including 20 mM NaCl), ellipticity at 218 nm was recorded in simple temperature scans, without acquisition over the complete wavelength range.

Estimation of Thermal Stability by Differential Scanning Fluorimetry. Differential scanning fluorimetry measurements were carried out to study thermal denaturation of tertiary structure of the proteins as described by Niesen et al.[36] The recombinant and PEGylated BinA proteins (0.1 mg/mL in 1× PBS buffer) were assayed in a 96 well plate in duplicates. SYPRO orange dye solution (5000× concentrate in 100% DMSO) was added directly to the protein samples to a final concentration of 5×. The plate was properly sealed and data was recorded on Light Cycler 480 Instrument II (Roche). The protein samples were heated from 25 to 95 °C at ramp rate of 1.2 °C/min with 25 acquisitions at each point. Thermal denaturation of Hen Egg White (HEW) lysozyme was also probed in the same 96 well plate as an internal control. The fluorescence signal (λ_{ex}, 498 nm; λ_{em}, 580 nm) was recorded as a function of temperature. Experimental data were processed using the in-built software and the melting temperature for each protein was determined from the first derivative.

Assessment of Stability against Proteolytic Enzymes. Recombinant BinA, m1BinA, m2BinA, and BinAB mixture proteins were assayed in independent experiments for their stability against proteolytic enzymes, like trypsin, proteinase K, and chymotrypsin. The proteins were mixed with proteases in protease:protein ratio of 1:10 (w/w) for trypsin, and 1:50 for proteinase K and chymotrypsin. Lower concentrations were used for proteinase K and chymotrypsin as 185 and 232 cleavage sites, respectively, in BinA sequence were predicted using Peptide Cutter tool (http://web.expasy.org/peptide_cutter/[37]) for these proteases, compared to only 30 sites for trypsin enzyme. The protein−protease reaction mixtures were incubated at 37 °C. Aliquots were collected at two different time points of 1 h and overnight incubation. The reaction was

DOI: 10.1021/acs.bioconjchem.6b00565
Bioconjugate Chem. 2017, 28, 410−418

terminated with 1 mM phenylmethylsulfonyl fluoride (PMSF). The collected aliquots were adjudged on 12% SDS-PAGE gel for assaying proteolytic stability of the proteins.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.bioconjchem.6b00565.

Image of barium iodide stained gel, DLS profiles of BinA proteins, thermal denaturation of hen egg white lysozyme used as an internal control in DSF experiments, and PEGylation of BSA used as an control in toxicity assays (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: vinay@barc.gov.in. Tel: +91-22-25593761. Fax: +91-22-25505151.

ORCID

Vinay Kumar: 0000-0002-7619-6236

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the research funding from Board of Research in Nuclear Sciences (BRNS), Department of Atomic Energy, India, to Vinay Kumar. The funding source has no role in study design, in the collection, analysis, and interpretation of data, in the writing of the report, and in the decision to submit the article for publication. We are thankful to Dr. Ajay Saini, BARC, for the help with Differential Scanning Fluorimetry, and Drs. Ashok Varma and Kakoli Bose, ACTREC, Kharghar, Navi Mumbai for their help with MALDI-TOF and CD experimental facilities. We also thank Sujit Tota for the help in maintaining mosquito culture.

■ REFERENCES

(1) Berry, C. (2012) The bacterium, *Lysinibacillus sphaericus*, as an insect pathogen. *J. Invertebr. Pathol. 109*, 1−10.

(2) Lacey, L. A. (2007) Bacillus thuringiensis serovariety israelensis and Bacillus sphaericus for mosquito control. *J. Am. Mosq. Control Assoc. 23*, 133−163.

(3) Berry, C., Hindley, J., Ehrhardt, A. F., Grounds, T., De Souza, I., and Davidson, E. W. (1993) Genetic determinants of host ranges of Bacillus sphaericus mosquito larvicidal toxins. *J. Bacteriol. 175*, 510−518.

(4) Lacey, L. A., and Mulla, M. S. (1990) Safety of Bacillus thuringiensis (H-14) and Bacillus sphaericus to nontarget organisms in the aquatic environment, In *Safety of Microbial Insecticides* (Laird, M., Lacey, L. A., and Davidson, E. W., Eds.) pp 169−188, CRC Press, Boca Raton, FL.

(5) Lacey, L. A., and Siegel, J. P. (2000) Safety and ecotoxicology of entomopathogenic bacteria. In *Entomopathogenic Bacteria: From laboratory to field application* (Charles, J. F., Delecluse, A., and Nielsen-LeRoux, C., Eds.) pp 253−273, Kluwer Academic Publishers, Dordrecht, The Netherlands.

(6) Brown, M. D., Watson, T. M., Carter, J., Purdie, D. M., and Kay, B. H. (2004) Toxicity of VectoLex (Bacillus sphaericus) products to selected Australian mosquito and non target species. *J. Econ. Entomol. 97*, 51−58.

(7) Broadwell, A. H., and Baumann, P. (1987) Proteolysis in the gut of mosquito larvae results in further activation of the Bacillus sphaericus toxin. *Appl. Environ. Microbiol. 53*, 1333−1337.

(8) Colletier, J. P., Sawaya, M. R., Gingery, M., Rodriguez, J. A., Cascio, D., Brewster, A. S., Michels-Clark, T., Hice, R. H., Coquelle, N., Boutet, S., et al. (2016) De novo phasing with X-ray laser reveals mosquito larvicide BinAB structure. *Nature 539*, 43−47.

(9) Oei, C., Hindley, J., and Berry, C. (1992) Binding of purified Bacillus sphaericus binary toxin and its deletion derivatives to Culex quinquefasciatus gut: elucidation of functional binding domains. *J. Gen. Microbiol. 138*, 1515−1526.

(10) Nicolas, L., Nielsen-Leroux, C., Charles, J. F., and Delécluse, A. (1993) Respective role of the42- and 51-kDa components of the Bacillus sphaericus toxin overexpressed in Bacillus thuringiensis. *FEMS Microbiol. Lett. 106*, 275−280.

(11) Hire, R. S., Hadapad, A. B., Dongre, T. K., and Kumar, V. (2009) Purification and characterization of mosquitocidal Bacillus sphaericus BinA protei. *J. Invertebr. Pathol. 101*, 106−111.

(12) Kale, A., Hire, R. S., Hadapad, A. B., D'souza, S. F., and Kumar, V. (2013) Interaction between mosquito-larvicidal *Lysinibacillus sphaericus* Binary toxin components: analysis of complex formation. *Insect Biochem. Mol. Biol. 43*, 1045−1054.

(13) Federici, B. A., Park, H. W., Bideshi, D. K., Wirth, M. C., Johnson, J. J., Sakano, Y., and Tang, M. (2007) Developing recombinant bacteria for control of mosquito larvae. *J. Am. Mosq. Control Assoc. 23*, 164−175.

(14) Shen, H., Brandt, A., Bissinger, B. W., Gunnoe, B., and Roe, R. M. (2009) Novel insecticide polymer chemistry to reduce the enzymatic digestion of a protein pesticide, trypsin modulating oostatic factor (TMOF). *Pestic. Biochem. Physiol. 93*, 144−152.

(15) Boccù, E., Velo, G. P., and Veronese, F. M. (1982) Pharmacokinetic properties of polyethylene glycol derivatized superoxide dismutase. *Pharmacol. Res. Commun. 14*, 113−120.

(16) Roberts, M. J., Bentley, M. D., and Harris, J. M. (2002) Chemistry for peptide and protein PEGylation. *Adv. Drug Delivery Rev. 54*, 459−476.

(17) Abuchowski, A., McCoy, J. R., Palczuk, N. C., van Es, T., and Davis, F. F. (1977) Effect of covalent attachment of polyethylene glycol on immunogenicity and circulating life of bovine liver catalase. *J. Biol. Chem. 252*, 3582−3586.

(18) Harris, J. M., and Chess, R. B. (2003) Effect of pegylation on pharmaceuticals. *Nat. Rev. Drug Discovery 2*, 214−221.

(19) Harris, J. M., Martin, N. E., and Modi, M. (2001) Pegylation: a novel process for modifying pharmacokinetics. *Clin. Pharmacokinet. 40*, 539−551.

(20) Alconcel, S. N. S., Baas, A. S., and Maynard, H. D. (2011) FDA-approved poly (ethylene glycol)-protein conjugate drugs. *Polym. Chem. 2*, 1442−1448.

(21) An, Q., Lei, Y., Jia, N., Zhang, X., Bai, Y., Yi, J., Chen, R., Xia, A., Yang, J., Wei, S., et al. (2007) Effect of site-directed PEGylation of trichosanthin on its biological activity, immunogenicity, and pharmacokinetics. *Biomol. Eng. 24*, 643−649.

(22) Jeffers, L. A., Shen, H., Khalil, S., Bissinger, B. W., Brandt, A., Gunnoe, T. B., and Roe, R. M. (2012) Enhanced activity of an insecticidal protein, trypsin modulating oostatic factor (TMOF), through conjugation with aliphatic polyethylene glycol. *Pest Manage. Sci. 68*, 49−59.

(23) Hire, R. S., Hadapad, A. B., Vijayalakshmi, N., and Dongre, T. K. (2010) Characterization of highly toxic indigenous strains of mosquitocidal organism Bacillus sphaericus. *FEMS Microbiol. Lett. 305*, 155−161.

(24) Grimsley, G. R., Scholtz, J. M., and Pace, C. N. (2008) A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Sci. 18*, 247−251.

(25) Kurfürst, M. M. (1992) Detection and molecular weight determination of polyethylene glycol-modified hirudin by staining after sodium dodecyl sulfate-polyacrylamide gel electrophoresi. *Anal. Biochem. 200*, 244−248.

(26) Skoog, B. (1979) A method for qualitative and/or quantitative detection of polyethylene glycols in biological fluids. *Vox Sang. 37*, 345−349.

DOI: 10.1021/acs.bioconjchem.6b00565
Bioconjugate Chem. 2017, 28, 410−418

(27) Perez-Iratxeta, C., and Andrade-Navarro, M. A. (2008) K2D2: estimate of protein secondary structure from circular dichroism spectra. *BMC Struct. Biol. 8*, 25.
(28) Freskgaard, P., Maartensson, L., Jonasson, P., Jonsson, B. H., and Carlsson, U. (1994) Assignment of the Contribution of the Tryptophan Residues to the Circular Dichroism Spectrum of Human Carbonic Anhydrase II. *Biochemistry 33*, 14281−14288.
(29) World health organisation (WHO) official website; http://www.who.int/en/ (accessed on August 2016).
(30) Bradford, M. 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.
(31) Hadapad, A. B., Vijayalakshmi, N., Hire, R. S., and Dongre, T. K. (2008) Effect of ultraviolet radiation on spore viability and mosquitocidal activity of an indigenous ISPC-8 Bacillus sphaericus Neide strai. *Acta Trop. 107*, 113−116.
(32) Abbott, W. S. (1925) A method of computing the effectiveness of an insecticide. *J. Econ. Entomol. 18*, 265−267.
(33) Finney, D. J., Ed. (1952) *Probit Analysis*, Cambridge University Press, Cambridge, England.
(34) *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, https://www.r-project.org/.
(35) Greenfield, N. J. (2007) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc 1*, 2527−2535.
(36) Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc. 2*, 2212−2221.
(37) Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook* (Walker, J. M., Ed.) pp 571−607, Humana Press Inc, Totowa, NJ.

DOI: 10.1021/acs.bioconjchem.6b00565
Bioconjugate Chem. 2017, 28, 410−418

Journal of Invertebrate Pathology 122 (2014) 44–47

Contents lists available at ScienceDirect

Journal of Invertebrate Pathology

journal homepage: www.elsevier.com/locate/jip

Regular Article

An oligomeric complex of BinA/BinB is not formed *in-situ* in mosquito-larvicidal *Lysinibacillus sphaericus* ISPC-8

Ramesh S. Hire [a], Mahima Sharma [b], Ashok B. Hadapad [a], Vinay Kumar [b,*]

[a] Nuclear Agriculture & Biotechnology Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400085, India
[b] Solid State Physics Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400085, India

ARTICLE INFO

Article history:
Received 29 June 2014
Accepted 26 August 2014
Available online 6 September 2014

Keywords:
Lysinibacillus sphaericus
Binary protein
Surface layer protein
Culex quinquefasciatus

ABSTRACT

Binary toxin of *Lysinibacillus sphaericus* is composed of two polypeptides; receptor binding BinB and toxic BinA. Both the polypeptides are required for maximal toxicity. It has been suggested that binary toxin exerts toxicity as a heterotetramer constituted by two copies of each of the component polypeptides. It has also been observed that oligomers consisting of two copies of BinA and BinB are pre-formed in *L. sphaericus* spore–crystals. However, recombinant proteins from *Escherichia coli* expression system elute individually as monomers. We purified the likely oligomeric complex from the spore–crystals of highly toxic *L. sphaericus* ISPC-8 strain and probed it with proteomic tools. The analysis showed that the high molecular mass complex in the toxic spore–crystals is composed of only surface layer protein (SlpC). The purified SlpC from the local isolate exists as a dimer and also showed poor mosquito-larvicidal activity.

1. Introduction

Lysinibacillus sphaericus is a gram positive soil bacterium which produces mosquitocidal MTX, Bin and Cry toxins. The Bin toxin is produced by highly active strains and these strains are widely used as biopesticide for mosquito control (Berry, 2012). The Bin toxin is produced during sporulation and is composed of two polypeptides; receptor binding BinB and toxic BinA. After ingestion, the Bin toxin proteins solubilize in the alkaline midgut and are converted to active BinA and active BinB. It has been observed that the active form of BinB binds to a cell surface receptor in midgut *microvilli* of larvae and the toxic BinA binds to this receptor-BinB complex (Silva-Filha et al., 1999). The heterotetramer (BinA2.BinB2) of Bin toxin formed onto cell receptor, a GPI-anchored α-glucosidase, inserts into the cell membrane forming membrane pores (Cokmus et al., 1997; Schwartz et al., 2001; Smith et al., 2005; Opota et al., 2011) and subsequently causes cell death.

It has been suggested that a stable oligomeric form of Bin toxin may exist in spore–crystals of *L. sphaericus* 2362 and this oligomer may be playing role in toxicity towards *Culex* larvae (Smith et al., 2005). Such a pre-formed oligomer was not reported for the recombinant proteins co-expressed in *Escherichia coli*, nor was this observed in a stoichiometric mixture of recombinant proteins *in vitro* (Kale et al., 2013). Interestingly, a covalent complex of active forms of BinA and BinB, synthesized by glutaraldehyde crosslinking, showed maximal activity reported for any combination of BinA and BinB proteins (Kale et al., 2013). It has also recently been shown that Bin toxin has the potential to be developed as a chemotherapeutic agent, as it induces apoptosis of human tumor cell (Luo et al., 2014). It is thus important that *L. sphaericus* components expressing toxic activity be characterized properly.

We probed for a pre-formed oligomer in locally isolated and highly toxic strain of *L. sphaericus* ISPC-8. We did not find any evidence of a pre-formed oligomer of BinA and BinB proteins. Rather, the expected protein complex, purified to homogeneity, was confirmed to be a surface layer protein that displayed poor mosquito-larvicidal activity.

2. Materials and methods

2.1. Bacterial strain and growth conditions

L. sphaericus ISPC-8 having serotype 5a5b and phage type 3 was isolated from dead larvae of *Culex fatigans*. The bacterial culture was grown till sporulation as described earlier (Hire et al., 2010). The pellet containing spore–crystals was washed twice with 1 M NaCl and 10 mM EDTA (pH 7.0) and twice with double-distilled water. The pellet was suspended in sterile distilled water and sonicated for 5 min with 3 s pulses at 30% amplitude in a Branson

* Corresponding author. Fax: +91 22 25505151.
E-mail addresses: vinay@barc.gov.in, vkbhatia@gmail.com (V. Kumar).

http://dx.doi.org/10.1016/j.jip.2014.08.005
0022-2011/

digital sonifier (Model 250, Branson Ultrasonics Corporation, CT). The sonicated suspension was centrifuged at 15,000 Xg for 30 min. The supernatant was discarded and the spore–crystal pellet was used for purification of the target proteins.

2.2. Purification of proteins

The purification of proteins was carried out with slight modification of the protocol described by Smith et al. (2005). The spore–crystal pellet was suspended in 50 mM NaOH and the suspension was kept shaking on ice for 3 h. The suspension was centrifuged at 15,000 Xg for 20 min at 4 °C. The supernatants containing the solubilized binary toxins were dialyzed overnight against buffer A (25 mM Tris–HCl, 10 mM NaCl, 2 mM DTT, pH 9.0). The dialyzed suspension was centrifuged at 15,000 Xg for 20 min at 4 °C and the supernatant was loaded onto Q-Sepharose column (Bio-Rad laboratories, Hercules, CA). The bound proteins were eluted with a linear gradient of 10–1000 mM NaCl in buffer A over six-column volume. The fractions of high molecular mass proteins were pooled and dialyzed extensively against buffer A. After dialysis, the pooled fractions were concentrated to about 2 mg/ml.

2.3. Estimation of molecular size

The purified proteins were loaded on the Superdex™ 200 10/300 GL column (Superdex-200; GE Healthcare) for further purification as well as for molecular weight determination. Each eluted peak from Superdex-200 molecular sieve column was further loaded onto Superdex-75 column. The Superdex-200 and Superdex-75 columns were calibrated with gel filtration molecular weight markers (Sigma–Aldrich; Cytochrome C, 12.4 kDa; Carbonic anhydrase, 29 kDa; Ovalbumin, 44.3 kDa; Albumin, 67 kDa; Apoferritin, 443 kDa; Thyroglobulin, 669 kDa). The eluted peak of each independent gel filtration experiment was resolved on SDS–PAGE. The purified proteins were concentrated to 5–7 mg/ml and were subjected to dynamic light scattering experiments at 25 °C with Malvern Zetasizer NanoZS instrument. Three sets of measurements with 18 acquisitions in each set were collected. Data were analyzed with Malvern Zetasizer software suite (http://www.malvern.com).

2.4. MALDI-TOF mass spectrometry

The mass spectrometry analysis was performed at the Adelaide Proteomics Centre, University of Adelaide, Australia. Lyophilized protein sample was sonicated in 100 µl of 10% acetonitrile (ACN)/0.1% trifluoroacetic acid (TFA) to dissolve the protein. About 1 µl of concentrated sample was electrophoresed in 5% acrylamide gel (Bis–Tris buffer). The gel was cut in two and one section was stained with silver and the other stained with Commassie G-250. The closely spaced high molecular weight (approximately 125 kDa) bands seen in the G-250 stained gel were excised, destained and digested with 100 ng of trypsin per sample according to the "low salt" protocol (Bruker Guide to MALDI Sample Preparation). One microlitre of each sample was applied to a 600 µm AnchorChip (Bruker Daltonik GmBH, Bremen, Germany) to which 400 ng of α-cyano-4-hydroxycinnamic acid had been applied. MALDI-TOF mass spectra were acquired using a Bruker ultraflex III MALDI-TOF/TOF mass spectrometer (Bruker Daltonik GmBH) operating in reflection mode under the control of the flex Control software (version 3.0, Bruker Daltonik GmBH). Between three and six of the most highly abundant sample ions (i.e. non-trypsin and non-keratin) were selected as precursors for MS/MS analysis. MALDI-TOF/TOF was performed in the LIFT mode using the same spot on the target.

The peptide mass fingerprinting search was carried using 21 mass peaks observed for Trypsin digested proteins and MASCOT search engine (http://www.matrixscience.com/). The specifications were: Taxonomy: eubacteria, databases: NCBInr 20140323 & SwissProt 2014_04, Enzyme: Trypsin, Fixed modifications: Carbamidomethyl (C), Variable modifications: Oxidation (M), peptide mass tolerance and charge state of 1.2 Da and 1+, and missed cleavages: 1. The MOWSE and probability scores calculated by the software were used as the criteria for protein identification.

2.5. Toxicity analysis of the high molecular mass protein

The susceptibility of third-instar *Culex quinquefasciatus* larvae was checked against *L. sphaericus* spore–crystal mixture (1×10^5 spores/ml) in 10 ml of distilled water in triplicate each containing ten larvae. Total mortality was scored after 48 h. The toxicity of the high molecular mass protein (purified by molecular sieve column chromatography) was tested against the same colony of *C. quinquefasciatus* larvae. Different concentrations of the purified protein were tested in 10 ml of distilled water containing ten third-instar *C. quinquefasciatus* larvae. Two replicates for each concentration (10, 50, 100, 150 µg/ml) were tested. The larval mortality was recorded after 48 h.

3. Results and discussion

Formation of high-molecular weight precursors (~125 kDa) of Bin toxin in *L. sphaericus* 2362 spores and their mosquitocidal activity has been intensely debated. These precursors migrate corresponding to proteins of molecular masses of 110 kDa and 125 kDa in reducing as well as non-reducing SDS–PAGE, and hydrodynamic radius of 5.6 ± 1.2 nm has been estimated for these (Baumann et al., 1985; Broadwell and Baumann, 1987; Smith et al., 2005). Attempts have also been made to probe these precursor proteins with biophysical methods and with antisera raised against BinA and BinB proteins. The earlier observed cross-reactivity of precursor proteins with the antiserum of BinA and BinB polypeptides was subsequently rationalized to be due to contamination of precursor peptides in BinA and BinB preparations used for raising antisera (Bowditch et al., 1989).

We purified to homogeneity the expected high molecular-weight precursor proteins and Bin proteins from spore–crystals of highly toxic local strain of *L. sphaericus* ISPC-8 using ion-exchange and molecular-sieve column chromatography. High molecular weight components (Mr, ~125 kDa), BinB (51.4 kDa) and BinA (41.9 kDa) were found to co-elute on ion-exchange column (Fig. 1A, Lane 2). The high molecular weight precursors and Bin proteins were clearly resolved by molecular sieve column chromatography using Superdex-200 (Figs. 1A and 2). This suggests that high molecular weight proteins and Bin components may not form a stable complex. The peaks were individually loaded onto Superdex-75 column. Expectedly, the high molecular weight precursor protein eluted in the void volume of Superdex-75 column. Also, BinA and BinB proteins were further resolved on this molecular sieve column. The molecular masses of the proteins were determined based on their elution profile on the molecular sieve Superdex-200 and Superdex-75 columns. The molecular mass of the high molecular weight protein was estimated to be 282 kDa. In contrast, the BinA and BinB proteins elute as monomers from Superdex-200 and Superdex-75 columns with estimated molecular mass of 40 and 56 kDa, respectively (Fig. 2). The purified high molecular weight precursor and Bin proteins were resolved on 12% SDS–PAGE (Fig. 1A). Two bands of approximately 125 and 110 kDa, in agreement with the published reports, were observed for the high molecular-weight proteins when resolved on 5% SDS–PAGE (Fig. 1B). The high molecular weight precursor proteins were further probed with dynamic light scattering experiments and by mass spectrometry.

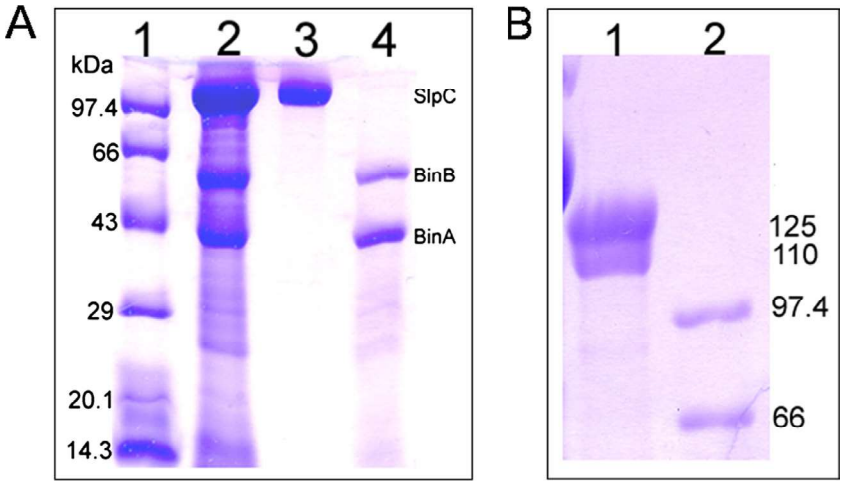


Fig. 1. Purification of proteins from the spores of *L. sphaericus* ISPC-8. (A) 12% SDS–PAGE analysis of proteins purified from the spore–crystal mixture using ion-exchange and molecular sieve column chromatography. Lane-1, molecular weight markers; Lane-2, spore–crystal proteins purified using anion-exchange chromatography; Lanes-3&4, High molecular weight precursor (SlpC) and Bin proteins purified using Superdex-200 column chromatography, respectively. (B) 5% SDS–PAGE gel showing two distinct bands at 125 and 110 kDa (Lane-1) of high molecular weight precursors; molecular weight marker (Lane-2).

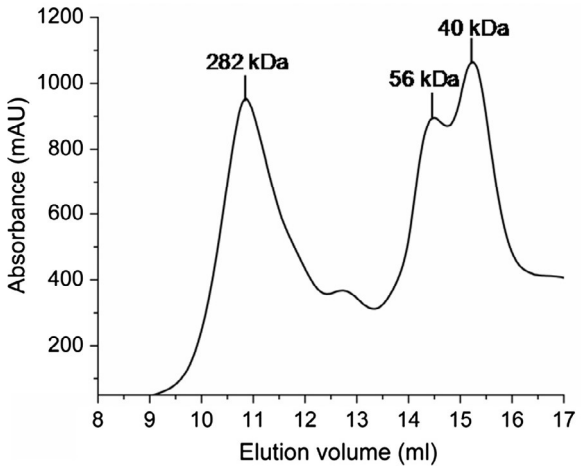


Fig. 2. Elution profile from Superdex-200 gel-filtration chromatography of proteins purified from spore–crystals of *L. sphaericus*. The high molecular weight protein elutes from Superdex-200 column corresponding to mass of 282 kDa. BinA and BinB proteins (corresponding to proteins of masses of 40 and 56 kDa) are well resolved in the elution profiles from Superdex-75 molecular sieve column.

The analysis of the major peak observed in the dynamic light scattering experiments showed that purified high molecular weight proteins were polydispersive (percentage polydispersity ~24.5) and had a molecular weight of about 244 ± 59 kDa corresponding to hydrodynamic radius of about 5.85 nm. The peptide mass fingerprinting and MS/MS analyses revealed with high confidence (MASCOT score 158, Expect score 3.9E–09) that the two SDS–PAGE bands of high molecular weight proteins correspond to surface layer protein (SlpC; GenBank ID: AAA50256.1) of *L. sphaericus*. All the 21 peptide masses used in MASCOT search matched with sequence of SlpC with the total protein sequence coverage of 23% (Fig. 3). With the whole genome sequence of at least three *L. sphaericus* species being available in databases, the present MASCOT analysis provides unambiguous identification of high molecular-weight protein of *L. sphaericus* spores as SlpC protein.

The observation of proteins of 110 and 125 kDa on SDS–PAGE in the cell extracts of *Synechococcus*, *Bacillus thuringiensis* and *E. coli* expressing *L. sphaericus binA* and *binB* genes (Sangthongpitag et al., 1997; Yuan et al., 2001; Shanmugavelu et al., 1998) were considered as an evidence of these high molecular-weight proteins to be composed of BinA and BinB components (Smith et al., 2005). BLAST search against NCBI database with *L. sphaericus* SlpC query sequence revealed the presence of *slpC*-like genes in *B. thuringiensis* (GenBank IDs: WP_021727980.1, ADU04484.1, AAY28601.1) and in *Synechococcus* (GenBank Id: YP_001226520.1). However, a homologous *slpC*-like gene was not detected in the genome of *E. coli*.

The SlpC protein of *L. sphaericus* (Mr, 125.2 kDa) is predicted to be secreted with a leader peptide that is removed between Ala-30 and Ala-31 of the primary translation product and is a precursor of the 110 kDa peptide (Bowditch et al., 1989). The molecular weight and hydrodynamic size estimations by molecular sieve column chromatography and dynamic light scattering suggest that a dimer of SlpC constitutes a major fraction in the solution form, though higher oligomers may also be possible. This is in contrast to earlier studies which suggested that high-molecular-weight SDS–PAGE bands were mono-disperse and had a molecular weight of 186 ± 38 kDa (Smith et al., 2005).

The mosquitocidal toxicity of SlpC protein from several *L. sphaericus* strains has been studied. The purified 110 kDa protein of *L. sphaericus* 2362 was found to be toxic to the larvae of *Culex pipiens* with LC_{50} values of 115 ng/ml (Broadwell and Baumann, 1986). The purified surface layer protein (10–80 μg) of wild type *L. sphaericus* C3-41 and recombinantly expressed in *E. coli* BL21 (100 μg) were not toxic to third-instar *Culex* larvae (Hu et al., 2008). Whereas S-layer proteins from *L. sphaericus* strains OT4b25, OT4b26 and III(3)7 have recently been reported to exert mild toxicity towards third-instar *Culex* larvae (LC_{50} 0.68 – 24 μg/ml; Lozano et al., 2011). The spore–crystal mixture of *L. sphaericus* ISPC-8 showed 100% mortality of *Culex* larvae after 48 h. Further, the toxicity of SlpC (high molecular weight protein) purified from spore–crystal of *L. sphaericus* ISPC-8 was assayed against third-instar larvae up to the concentration of 150 μg/ml. The purified SlpC protein was not toxic to *Culex* mosquitoes up to 100 μg/ml protein concentration but showed poor toxicity (20% mortality) at 150 μg/ml, as compared to control experiments without the SlpC protein. Intriguingly, the SlpC protein is highly conserved in different strains of *L. sphaericus* (Hu et al., 2008). Whether the observed differences in SlpC toxicity are due to dose of the protein tested or due to processing of 125 kDa protein needs further investigation.

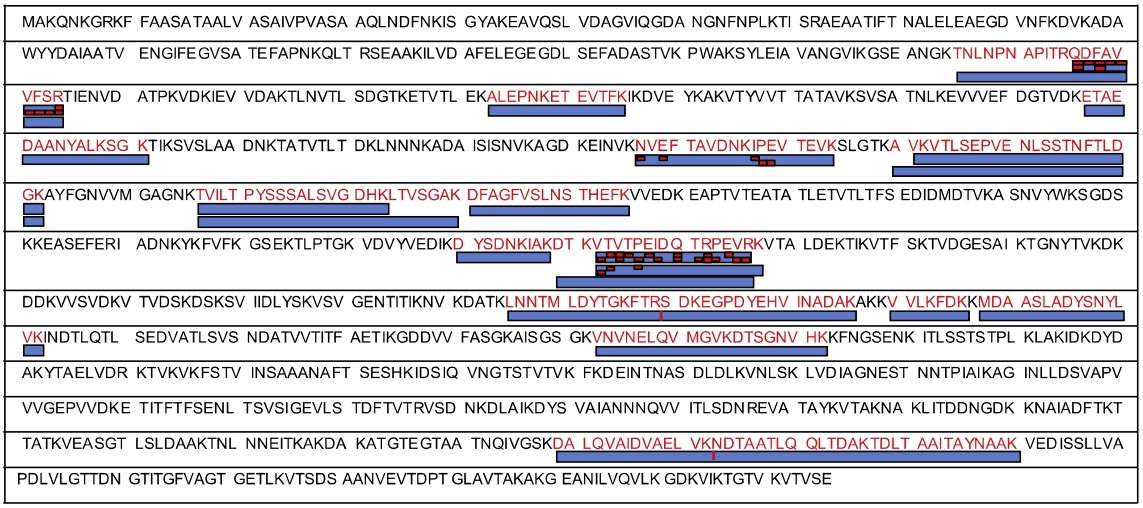


Fig. 3. Sequence coverage diagram in the MASCOT analysis. Blue boxes indicate peptides matched by the peptide mass fingerprinting with the red boxes indicating *b*- and *y*-ions from the MS/MS spectra. The amino acids shown in red constitute the trypsin generated peptide fragments those were detected in MALDI-TOF experiments within ± 1.2 Da. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In conclusion, this work provides the evidence that the high molecular weight (110 and 125 kDa) proteins of *L. sphaericus* spore–crystals, which usually co-purify with the Bin toxin, do not represent oligomeric binary toxin. These are products of highly conserved S-layer protein of *L. sphaericus*. The purified S-layer protein of ISPC-8 strain showed poor toxicity against the third instar of *Culex* larvae.

Acknowledgments

This work was supported by the research funding from Board of Research in Nuclear Sciences (BRNS), Department of Atomic Energy, India, to Vinay Kumar and Ramesh Hire. The funding source has no role in study design, in the collection, analysis and interpretation of data, in the writing of the report, and in the decision to submit the article for publication. The authors wish to thank Head, NA&BTD, BARC for encouragement and support. The authors are thankful to Dr. Gagan D. Gupta for the help with gel-filtration and DLS experiments. The authors are also thankful to Drs. Ravindra Makde and Sahayog Jamdar for useful discussions.

References

Berry, C., 2012. The bacterium, *Lysinibacillus sphaericus*, as an insect pathogen. J. Invertebr. Pathol. 109, 1–10.
Baumann, P., Unterman, B.M., Baumann, L., Broadwell, A.H., Abbene, S.J., Bowditch, R.D., 1985. Purification of the larvicidal toxin of *Bacillus sphaericus* and evidence for high-molecular-weight precursors. J. Bacteriol. 163, 738–747.
Bowditch, R.D., Baumann, P., Yousten, A.A., 1989. Cloning and sequencing of the gene encoding a 125-kilodalton surface-layer protein from *Bacillus sphaericus* 2362 and of a related cryptic gene. J. Bacteriol. 171, 4178–4188.
Broadwell, A.H., Baumann, P., 1986. Sporulation–associated activation of *Bacillus sphaericus* larvicide. Appl. Environ. Microbiol. 52, 758–764.
Broadwell, A.H., Baumann, P., 1987. Proteolysis in the gut of mosquito larvae results in further activation of the *Bacillus sphaericus* toxin. Appl. Environ. Microbiol. 53, 1333–1337.
Cokmus, C., Davidson, E.W., Cooper, K., 1997. Electrophysiological effects of *Bacillus sphaericus* binary toxin on cultured mosquito cells. J. Invertebr. Pathol. 69, 197–204.
Hire, R.S., Hadapad, A.B., Vijayalakshmi, N., Dongre, T.K., 2010. Characterization of highly toxic indigenous strains of mosquitocidal organism *Bacillus sphaericus*. FEMS Microbiol. Lett. 305, 155–161.
Hu, X., Li, J., Hansen, B.M., Yuan, Z., 2008. Phylogenetic analysis and heterologous expression of surface layer protein SlpC of *Bacillus sphaericus* C3-41. Biosci. Biotechnol. Biochem. 72, 1257–1263.
Kale, A., Hire, R.S., Hadapad, A.B., D'souza, S.F., Kumar, V., 2013. Interaction between mosquito-larvicidal *Lysinibacillus sphaericus* Binary toxin components: analysis of complex formation. Insect Biochem. Mol. Biol. 43, 1045–1054.
Lozano, L.C., Ayala, J.A., Dussan, J., 2011. *Lysinibacillus sphaericus* S-layer protein toxicity against *Culex quinquefasciatus*. Biotechnol. Lett. 33, 2037–2041.
Luo, W., Liu, C., Zhang, R., He, J., Han, B., 2014. Anticancer activity of binary toxins from *Lysinibacillus sphaericus* IAB872 against human lung cancer cell line A549. Nat. Prod. Commun. 9, 107–110.
Opota, O., Gauthier, N.C., Doye, A., Berry, C., Gounon, P., Lemichez, E., Pauron, D., 2011. *Bacillus sphaericus* binary toxin elicits host cell autophagy as a response to intoxication. PLoS one. 6, e14682.
Sangthongpitag, K., Penfold, R.J., Delaney, S.F., Rogers, P.L., 1997. Cloning and expression of the *Bacillus sphaericus* 2362 mosquitocidal genes in a non-toxic unicellular cyanobacterium, *Synechococcus* PCC6301. Appl. Microbiol. Biotechnol. 47, 379–384.
Schwartz, J.L., Potvin, L., Coux, F., Charles, J.F., Berry, C., Humphreys, M.J., Jones, A.F., Bernhart, I., Dalla Serra, M., Menestrina, G., 2001. Permeabilization of model lipid membranes by *Bacillus sphaericus* mosquitocidal binary toxin and its individual components. J. Membr. Biol. 184, 171183.
Shanmugavelu, M., Rajamohan, F., Kathirvel, M., Elangovan, G., Dean, D.H., Jayaraman, K., 1998. Functional complementation of nontoxic mutant binary toxins of *Bacillus sphaericus*1593M generated by site directed mutagenesis. Appl. Environ. Microbiol. 64, 756–759.
Silva-Filha, M.H., Nielsen-LeRoux, C., Charles, J.F., 1999. Identification of the receptor for *Bacillus sphaericus* crystal toxin in the brush border membrane of the mosquito *Culex pipiens* (Diptera: Culicidae). Insect Biochem. Mol. Biol. 29, 711–721.
Smith, A.W., Cámara-Artigas, A., Brune, D.C., Allen, J.P., 2005. Implications of high molecular-weight oligomers of the binary toxin from *Bacillus sphaericus*. J. Invertebr. Pathol. 88, 27–33.
Yuan, Z., Rang, C., Maroun, R.C., Juárez-Pérez, V., Frutos, R., Pasteur, N., Vendrely, C., Charles, J.F., Nielsen-Leroux, C., 2001. Identification and molecular structural prediction analysis of a toxicity determinant in the *Bacillus sphaericus* crystal larvicidal toxin. Eur. J. Biochem. 268, 2751–2760.