Title
Structural and Biochemical Studies of Cell-Grown Cry11Ba Insecticidal Protein Crystals

Author
Schibrowsky, Natalie Alice

Publication Date
2023

UNIVERSITY OF CALIFORNIA

Los Angeles

Structural and Biochemical Studies of Cell-Grown Cry11Ba Insecticidal Protein Crystals

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy in Biochemistry, Molecular, and Structural Biology

by

Natalie Alice Schibrowsky

2023

ABSTRACT OF THE DISSERTATION

Structural and Biochemical Studies of Cell-Grown Cry11Ba Insecticidal Protein Crystals

by

Natalie Alice Schibrowsky

Doctor of Philosophy in Biochemistry, Molecular, and Structural Biology

University of California, Los Angeles, 2023

Professor Jose Alfonso Rodriguez, Chair

Nature self-assembles protein structures for various functions, including: storage, protection, and fortification. These self-assemblies range from filaments to full three-dimensional crystals and are pervasive across the tree of life. They include the granules present in immune system cells, the packing of hormones in the pancreas, the storage of proteins in plants, carboxysomes, viruses, and cell-grown crystals in microbes. To accomplish this, a better understanding of how certain organisms are able to naturally self-assemble macromolecules and how to recreate these in living cells must be achieved. *Bacillus thuringiensis* subsp. *israelenesis*' (*Bti*) crystalline inclusions are of exceptional interest, since they naturally package a single protein into a crystalline inclusion through a life cycle process called sporulation. Compared to classical macromolecular crystallography that takes a plethora of variables to exhaustion can still yield no crystals. The laborious process could be prevented; however, by better understanding Bt and these crystalline inclusions' cellular self-assembly process. Cry11Ba is a protein packed into these crystalline (Cry) inclusions and found to be one of the most toxic pesticidal proteins. These crystals are then ingested by their host and switch from their packaged toxin crystal to their inactive protoxin at the high pH within their gut. My interests have been in elucidating the macromolecular structure from in vivo produced crystals, further understanding the ambiguous mode of action to gain better perspective for other δ-endotoxins, and probing the self-assembly of the crystalline inclusions in vivo throughout the sporulation process. I have studied Cry11Ba with structural analysis, solubility & toxicity assays, mutational studies, and imaging capabilities developed in cryo-EM to successfully in solve a de novo structure via in vivo crystalline inclusions, charting the pH sensitivity of the crystals, identifying key residues for stability and toxicity, analyzed the monomeric and multimeric particles in alkaline environments to understand Cry11Ba's mode of action, and visualized previously unobserved sporulation stages for *Bti*.

The dissertation of Natalie Alice Schibrowsky is approved.

Steven G. Clarke

Robert P. Gunsalus

Margot E. Quinlan

Jose Alfonso Rodriguez, Committee Chair

University of California, Los Angeles

2023

DEDICATION

To my family and friends who have shown me unconditional love and support during this process, all of my mentors who have invested their time and knowledge in me, and to scientists who continue to actively participate, promote, and pursue diversity, equity, and inclusivity in the scientific community.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: Introduction

Chapter 2: Serial femtosecond structural determination and biochemical analysis of *in vivo* Cry11Ba crystalline inclusions

Chapter 3: Mutational and Biochemical Analysis of Cry11Ba

Chapter 4: Electron microscopy of Cry11Ba crystalline inclusions at biologically relevant alkaline pH

Appendix A: Supplementary Information for Chapter 2: Serial femtosecond structural determination and biochemical analysis of *in vivo* Cry11Ba crystalline inclusions

Appendix C: Self-assembly of *Bacillus thuringiensis* subsp. *israelenesis* during sporulation harboring Cry11Ba crystalline inclusions

LIST OF VIDEOS

Chapter 3: Mutational and Biochemical Analysis of Cry11Ba

Chapter 4: Electron microscopy of Cry11Ba crystalline inclusions at biologically relevant alkaline pH

LIST OF TABLES

Chapter 2: Serial femtosecond structural determination and biochemical analysis of *in vivo* Cry11Ba crystalline inclusions

Chapter 3: Mutational and Biochemical Analysis of Cry11Ba

Appendix A: Supplementary Information for Chapter 2: Serial femtosecond structural determination and biochemical analysis of *in vivo* Cry11Ba crystalline inclusions

ACKNOWLEDGEMENTS

The completion of my PhD would not have been possible without the small army of individuals that helped me see this to the end.

I first need to thank the Rodriguez lab that without this thesis would not have been possible. First and foremost, Prof. Jose Rodriguez, who took me on as a student and provided me with the opportunity to conduct this fascinating research. His insight, guidance, vision, and independence were key and allowed not only the project, but myself as a scientist, to evolve, adapt, and grow in different ways, which I will always be eternally grateful. Next, I would be remiss to not thank all of the members of the Rodriguez lab who make this an amazing and inspiring place to work everyday. This includes Logan, Maria, Ambarneil, Sam, Becky, Jess, Lily, Niko, Aldo, Nima, Eric, Nina, Fred, Gibran, Andres, and Jeff with a special thanks to Dr. Marcus Gallagher-Jones and Shervin who helped me develop the idea and data processing for Cry11Ba's tomography. To my esteemed committee members, Dr. Margot Quinlan, Dr. Steve Clarke, and Dr. Robert Gunsalus, whose approval, guidance, advice, and support were insightful and enabled me to overcome some of my most difficult obstacles. The amazing instrumentation facility managers, Dr. Duilio Cascio, Dr. Mark Arbing, Genesis Falcon, Alex Lisker, Dr. Yu Chen, Dr. Dan Anderson, Dr. Martin Phillips, and Dr. Will Silkworth, who were all instrumental (no pun intended) in my training and without would not have been able to have conduct, produce, or analyze my data. The mentors that have invested their precious time and effort in training me, Dr. Mike Grasso, Dr. Adam Olia, and Dr. Michael (Mikey) Sawaya, thank you for making me the painstakingly thorough, diligent, and hard-working scientist I am today. To the PIs I have had the pleasure to work for, Dr. Lloyd Stark, Dr. Hong Sun, and Dr. Ronen Marmorstein, thank you for taking me on as a student and allowing me to experience this amazing world of scientific research in your labs. Lastly thank you to all of my collaborators, Dr. Brian Federici, Dr. Hyun-Woo Park, Dr. Dennis Bideshi, and Sarah-Jane Rudd from UC-Riverside; Dr. Jacques-Phillipe Colletier and Guillaume Tetreau from University of Grenoble Alpes; Dr. Trevor Moser and Dr. James Evan’s from PNNL; and the LCLS XFEL team; you all provided me your expertise, passing of knowledge, and invested in me and I hope to pay it forward with the next generation of scientists as you have done for me and have my deepest gratitude.

At UCLA, I would also like to thank all of my fellow BMSB PhD fellows, as this collaborative, engaging, and welcoming community we have cultured would not be possible, especially my 2017 cohort. Special thanks go to Maeve, John, and Orlando who convinced me during recruitment to come to UCLA and I never regretted the decision. I would also like to thank the people I am ecstatic to call colleagues, but luckier to call friends as they helped me in so many ways to reach the light at the end of this amazing journey. Hannah, Cindy, Nicole, Josh, Carolyn, Bryan, Dylan, QiaoQiao, Cheylene, and Eileen thank you for all of that, watching my pets, and our conversations that made me feel normal in our hectic grad school lives. A special thanks to the students I have had the pleasure to teach and TA for, thank you for allowing me to be part of your scientific journey and helping me become a better passer of scientific knowledge. Thank you to all of the wonderful individuals that I have worked with in gBSA, MPSC, and CNSI who helped me run events, create outreach, and make this an inclusive place. These are the relationships that I have made and know will last a lifetime.

I especially want to thank all of my friends, KC, Sonny, George, Sam, Danny, Zybrell, Eduardo, Andrea, Kennedy, Maddy, Terry, Reggie, Charlotte, David, Jake, Dani, Sterling, Nate, Adrian, and Claudia, just to name a few back home in Las Vegas and my “Neon Fam” from streaming that made me feel whole and complete despite the distance during lockdown. Thank you for talking all of my late-night calls, staying up with me, and confirming that we are not crazy, just night owls. My past teachers that have inspired me throughout my life, Anissa Cole, Cathi Vuckovic, Cynthia Pechman, MaryJo Viscardo, Susan McGuan, Bob McKee, Luther Bohanan, Dr. Maria Jerinic, Dr. Marta Meana, Dr. Gail Ball, Dr. Jimmy Peltier, Dr. Anne Hong-Hermesdorf, Dr. Keith Neidermeier, Dr. David Allen, Dr. Ralph Conti, Dr. Emily Fant, and Dr. Nadia Pomirleanu, thank you for always being a beacon of light when I needed it most through my education and many life lessons. My pets, Anastasia, Christian, Elric, Kazahana, Mahō Shōjo, and Raijū who have stayed up countless nights with me and listened to way too many scientific talks when they only know the sounds of their names and food bags.

Lastly, thank you to my AMA family, my actual family and especially my parents, Angel and Jack, who have always believed in me, never doubted me, and never questioned my craziest ideas, even when I declared I would get a PhD in Zoology at the ripe age of 9. All I can say is just…thank you, from the bottom of my heart for never giving up on me and giving me every opportunity to succeed even when the odds were trying to stack against me. Hopefully, I have finally proven that I am “The Smartest Kid” you know.

The work presented here was supported in part by the Cellular and Molecular Biology Pre-Doctoral Training Grant (NIH T32), the DOE Grant DE-FC02-02ER63421, and the NIH NIGMS Grant R35 GM128867.

Without further ado, let the science commence!

Chapter 2 in this thesis is presented (Tetreau, G. et al., 2022) with permission from the authors.

(1) Tetreau, G.‡; Sawaya, M. R.‡; De Zitter, E.‡; Andreeva, E. A.§; Banneville, A.-S.§; Schibrowsky, N. A. §; Coquelle, N.; Brewster, A. S.; Grünbein, M. L.; Kovacs, G. N.; Hunter, M. S.; Kloos, M.; Sierra, R. G.; Schiro, G.; Qiao, P.; Stricker, M.; Bideshi, D.; Young, I. D.; Zala, N.; Engilberge, S.; Gorel, A.; Signor, L.; Teulon, J.-M.; Hilpert, M.; Foucar, L.; Bielecki, J.; Bean, R.; de Wijn, R.; Sato, T.; Kirkwood, H.; Letrun, R.; Batyuk, A.; Snigireva, I.; Fenel, D.; Schubert, R.; Canfield, E. J.; Alba, M. M.; Laporte, F.; Després, L.; Bacia, M.; Roux, A.; Chapelle, C.; Riobé, F.; Maury, O.; Ling, W. L.; Boutet, S.; Mancuso, A.; Gutsche, I.; Girard, E.; Barends, T. R. M.; Pellequer, J.-L.; Park, H.-W.; Laganowsky, A. D.; Rodriguez, J.; Burghammer, M.; Shoeman, R. L.; Doak, R. B.; Weik, M.; Sauter, N. K.; Federici, B.; Cascio, D.; Schlichting, I.; Colletier, J.-P. De Novo Determination of Mosquitocidal Cry11Aa and Cry11Ba Structures from Naturally-Occurring Nanocrystals. *Nat Commun* **2022**, *13* (1), 4376. https://doi.org/10.1038/s41467-022-31746-x.

CURRICULUM VITAE

Education

University of Nevada, Las Vegas, 2010 - 2015

Bachelor of Science in Biochemistry, *Cum Laude*

University Honors and Department Honors

Honors Thesis: “The Key Factor in Desiccation Tolerance Success for *Timmiella crassinervis*"

University of Pennsylvania, 2015 - 2017

Masters of Chemical Science in Biological Chemistry

Capstone Thesis: “Identification of a copper binding site in MEK1 that promotes MAPK signaling”

Honors and Awards

2023	Ralph and Charlene Bauer Graduate Award
2021	SoCal cryo-EM Symposium Poster Award
2018 – 2020	Cellular and Molecular Biology Training Grant Fellowship

Publications

Resulting from this dissertation research

(1) Tetreau G ‡, Sawaya MR ‡, de Zitter E ‡, Banneville A-S §, Andreeva EA §, **Schibrowsky NA** §, Coquelle, Brewster AS, Grunbein M-L, Nass-Kovacs G, Hunter MS, Kloos M, Sierra RG, Schiro G, Qiao P, Stricker M, Bideshi D, Young ID, Zala N, Engilberge S, Gorel A, Signor L, Teulon J-M, Hilpert M, Foucar L, Bielicki J, Bean R, de Wijn R, Sato T, Kirkwood H, Letrun R, Batyuk A, Snigireva I, Fenel D, Schubert R, Canfield EJ, Alba MM, Laporte F, Després L, Bacia M, Roux-Gossart A, Chapelle C, Riobe’ F, Maury O, Ling WL, Boutet S, Mancuso A, Gutsche I, Girard E, Barends TRM,

Pellequer J-L, Park H-W, Laganowsky AD, Rodriguez JA, Burghammer M, Shoeman RL, Doak, RB, Weik M, Sauter NK, Federici B, Cascio D, Schlichting I, Colletier, J-P. (2022) De novo SAD phasing of serial femtosecond nano-crystallography data sheds light on the *in vivo* structures of mosquitocidal Cry11Aa and Cry11Ba. **Nat Commun.** *13*(1): 4376

‡ and § denotes equal contributions

Outside of this dissertation research

(2) Richards, L. S.; Flores, M. D.; Zink, S.; **Schibrowsky, N. A.**; Sawaya, M. R.; Rodriguez, J. A. Cryo-EM Structure of a Human LECT2 Amyloid Fibril Reveals a Network of Polar Ladders at Its Core. bioRxiv February 11, 2023, p 2023.02.08.527771. https://doi.org/10.1101/2023.02.08.527771

(3) Grasso, M.; Bond, G. J.; Kim, Y.-J.; Boyd, S.; Matson Dzebo, M.; Valenzuela, S.; Tsang, T.; **Schibrowsky, N. A.**; Alwan, K. B.; Blackburn, N. J.; Burslem, G. M.; Wittung-Stafshede, P.; Winkler, D. D.; Marmorstein, R.; Brady, D. C. The Copper Chaperone CCS Facilitates Copper Binding to MEK1/2 to Promote Kinase Activation. *Journal of Biological Chemistry* **2021**, *297* (6), 101314. DOI:10.1016/j.jbc.2021.101314.

Chapter 1

Introduction

1.1 Background and Significance

While structure determination techniques have continued to advance and solve more complex structures, there are still processes that have not been able to be completely studied. One gap in knowledge is the self-assembly process of proteins and cellular structures within a native or *in vivo* environment. Previously, these data would have been unobtainable, but with the advancements made in structural biology for resolution and determining ultrastructures, these processes can now be studied in-depth *in vivo*.

1.2 Macromolecular Protein Self-Assembly

Nature self-assembles proteins for various functions, including storage, protection, and fortification. These self-assemblies range from filaments to three-dimensional crystals and are pervasive across the tree of life. Even with all of these discoveries of different inclusion bodies, protein self-assembly within a cell is still a poorly understood concept.[1] The general self-assembly of biological macromolecules has been understood energetically as spontaneous upon synthesis of the polypeptide chain with entropy of the environment or system driving the protein folding process.[2,3] There are different types of self-assemblies, either static, where the molecules are in close enough contact to form a structure, or dynamic, where the molecules are colliding rapidly and as a result form different structures. With both static and dynamic assemblies, the molecules can then undergo co-assembly.[4] Co-assembly is the process of two different species assembling at the same time either hierarchically, where one interaction occurs and acts as a building block for another interaction, or directed, when molecules are affected by the environment to localize and induce close-field interactions. The packing, directionality, selectivity, attraction, and connectivity of the molecules' contacts can all drastically affect these mechanisms.[5–7] While mechanically this does address the question of why these biological macromolecules do fold, it does not address the active selection for particular protein folds from secondary to quaternary structure.

Cyrus Levinthal began to discuss this topic and came to Levinthal's paradox, where the degrees of freedom a polypeptide chain possesses will consider an immense, potentially infinite, amount of possible protein structures in its search for selecting the best minimum energy configuration.[8] Proteins; however, fold spontaneously and rapidly, which led Christian Anfinsen to suggest that there is a "pathway" to the correct, low-energy, native protein state. Currently, the widely accepted rationale for protein fold selection is Anfinsen's dogma's "Thermodynamic Hypothesis," where as previously mentioned the driving force is energetic where a polypeptide chain will select the lowest energetic state until it reaches stability and is unable to overcome another barrier into a new local free energy minimum.[9] With both of these hypotheses revolving around how the primary structure can give way to the tertiary structure at the rapid, spontaneous speed, Ken Dill came to a "folding funnel hypothesis," which begins to address these questions with a funnel-like energy landscape plot,[10] where the native state will reach a low energy configuration for its native state, with multiple local minima configurations along the way during its molten globular form.[11–14] This further followed Anfinsen's dogma and addressed Levinthal's paradox where the folding of each secondary structure of the molten globular protein would reach a new local minima intermediate state and proceed down a kinetic path to a native fold.[12] This has prompted further discussion with Alan Fersht's Φ-value analysis as a way to quantitatively characterize protein folding by mutating a certain residue and seeing its effects on the protein's folding ability/barrier.[15] The Φ-value addresses the two major hypotheses of protein folding: three-step mechanism or framework model. The three-step mechanism considers that short range, residues close in sequence, fold their secondary structures first, while the framework model suggests that the secondary and tertiary structures form simultaneously, with long range interactions forming first.[16,17] To probe these, the transition states of proteins undergoing folding Φ-values were calculated via algorithmic models or trapped transition states to begin elucidating the folding path it follows, while the sampling component is still unclear to confirm either hypothesis.[18]

1.3 Protein Crystallization Self-Assembly

In order to determine the structures of these macromolecules and understand how they fold and function, protein crystallization became one of the top structural biology techniques. This technique makes aqueous proteins solid by exploiting the packaging of macromolecules into ordered crystal lattices that utilize X-rays to produce diffraction patterns that are unique to the protein of interest and determine the position of the residues at atomic resolution.[19] In order to have the aqueous, soluble native proteins become amenable to crystallization, the protein is placed in a mixture containing protein, buffer, and precipitant, which through vapor diffusion increase the protein concentration to the super-saturation threshold (**Fig. 1.1a,b**).[19,20] By the protein slowly reaching super-saturation, it begins to nucleate a crystalline lattice, this continues to pack within the precipitation range and form mature macromolecular crystal. The precipitant is the key component of the mixture that will vary different factors, including: pH, temperature, chemical additives, and concentrations, to favor the macromolecular crystallization mechanics.[19] While many crystallization techniques have been developed and variables manipulated to cause crystallization, the exact crystallization parameters for each protein vary and have never been quite understood why they worked over others. For this reason, protein crystallography has maintained a technique barrier of trial-and-error with many protein studies never producing protein crystals or ones of sufficient quality to determine a high-resolution structure. While a wide array of structure determination values ensure that the data is not biased or over-fit when building/solving the structure, other issues may arise from the crystal itself.[19] One issue is whether the crystallized structures are true representations of the native proteins within the environment due to the manipulations that proteins undergo in order to be crystallized. This is a valid point-of-view that many protein crystallographers have addressed with crystal artifacting, which include the multimer state due to crystal packing and symmetry and contaminated crystallization due to the protein of interest's purification process.[21] As with any technique, there are many pros and cons, but more than 50 years after the solving of the first protein structures, X-ray crystallography has remained a strong pillar of structural biology and provided countless insight into proteins' structures for more informed experiments to probe functionality and structurally-targeted drug design.[22,23]

1.4 *in vitro* and *in vivo* Macromolecular Self-Assembly

Though there have been several *in vitro* studies to determine values for protein folding, there have been few to study *in vivo* folding and able to compare them. The inability to compare the two is mostly due to the lack of techniques to efficiently collect quantitative or qualitative data to determine a mechanism in both environments. To combat this, there have been a number of techniques developed and optimized to investigate these complex cellular environments, ie. interferometric PhotoActivated Localization Microscopy (iPALM), STochastic Optical Reconstruction Microscopy (STORM), Fluorescence Resonance Energy Transfer (FRET), and 3D-modeling, which will be immense aids to transition from *in vitro* to *in vivo* studies. Due to their high-sensitivity levels, these techniques have been utilized to study *in vivo* protein interactions along with how the environments affect the protein expression levels.[24] For example, alkaline pHs, more than pH 7.9, were found to decrease the presence of certain nutrients and prevent the ability of certain intramolecular interactions to occur for proteins to properly fold.[25]

Delving deeper into macromolecular self-assembly is the ability of certain cells to form crystalline inclusions. Biomineralization is the formation of crystals in biological systems and can range in size from diatomic to macromolecular molecules.[26,27] This process is not only affected by the cellular environment, but also the presence of chaperoning proteins in the crystal assembly process. While most organisms that form these crystalline inclusions are fungus and bacteria cells from genus *Candida*, *Pseudomonas*, and *Schizosaccharomyces*, that can form diatomic crystals in vivo,[26] animal cells have also been discovered to form crystals within different tissues. One example is the discovery of rod-like cytoplasmic inclusions called Reinke crystals within Leydig cells and have only been observed in adult humans and wild bush rats with unknown function.[28–30]

One species of high interest is the *Bacillus* genus as they have evolved to package insecticidal toxin proteins into a crystal as a survival mechanism when their soil environment indicates a low nutrient stress.[26] It is suspected, but not well understood which of the approaches the *Bacillus* cells utilizes to form their crystalline inclusions that form during stressed growth cycles. This is of interest, since classical crystallography techniques take a wide range of variables into consideration, eg. pH, salts, concentration, etc., when trying to form macromolecular crystals; however, these cells can undergo this self-assembly process naturally. Many questions arise about this self-assembly, such as do these cells producing the ideal environment, are the proteins constructed in a way to induce this alone or are there other structures involved within the cell to aid in this inclusion formation by way of a macromolecular scaffold, that are not clearly understood since this aspect of the cells has been understudied.

1.5 *Bacillus thuringiensis*' Production of *in vivo* Toxin Crystals via Sporulation

Bacillus thuringiensis (*Bt*) is a soil-dwelling, rod-shaped bacteria that are sensitive to the nutrient levels in their environment.[31] They were first discovered in a silkworm colony in Japan in what was first called "diseased," but was a result of this colony ingesting the *Bt*[32–35] with more research being conducted in Germany upon its discovery there.[36] Upon their discovery, they were first believed to responsible through a parasitic relationship within the host, ie. silkworms, but was truly a relationship through the lysis and hibernated death of the *Bt* that will then harm their vector their targeted pesticidal proteins. They can vary in dimensions with sizes ranging 0.5 to 2.5 µm thick by 1.2 to 10 µm long. These bacteria have been observed to undergo two growth phases, vegetative and sporulation.[37,38] Vegetative cells are cells with a plentiful amount of nutrients available and thus are constantly undergoing binary fission; however, when the resources are depleted in the area, *Bt* will signal to itself that it must conserve nutrients and the cells will undergo a process of sporulation. The cell then begins to split the DNA for normal binary fission with a membrane invagination that forms a coat around one side with the DNA. The other half of the DNA is then degraded, and a spore coat is formed around the membrane-isolated DNA. This spore coat provides protection of the DNA from the environment and can signal to the cell to leave the dormant state when conditions are favorable for growth. The spore coat continues to mature until the cell naturally lyses and releases the protected DNA. During sporulation, *Bt* maintain the rod-shape while containing an endospore (non-activated, dormant) and a toxin crystal[39,40] (**Fig. 1.2a-c**).

Bt strains are of interest due to their ability to natively produce functional proteins that crystallize as a product of the sporulation phase. These toxin crystals contain dormant delta-endotoxins (cytolytic and crystal proteins) and are utilized to reduce insect-transmitted diseases, eg. mosquitoes[41], via these endotoxin's natural insecticidal properties. The insecticidal properties are specific to different insects depending on the subspecies of *Bt*. These *Bt*-produced delta-endotoxins are either crystalline (Cry) or cytolytic (Cyt) parasporins.[42] The Cyt proteins display cytolytic or hemolytic activity upon interacting the membrane lipids and causing the host's cells to burst displaying these toxin properties broadly during *in vitro* studies and during the larva stage from *in vivo* studies. Cry proteins become active toxins once they are ingested by the insect and bind to the posterior gut where they form a pore within the stomach's midgut and preventing the insect from digesting nutrients.[37] While both proteins have toxic (insecticidal) properties, Cry and Cyt proteins share limited sequence homology, which has been shown previously to have different activation pathways.[43] While not confirmed, these toxin activation pathways could be related to the proteins' structure from their natural crystalline assembly since the Cyt and Cry proteins can pack differently. These proteins are of interest since they naturally self-assemble as crystals and could be a candidate to study *in vivo* self-assembly and the possibility of another crystallography technique in the future.

1.6 X-ray Free Electron Lasers and Serial Femtosecond Crystallography Make Breakthroughs from Crystal Size Limitations

An X-ray Free Electron Laser (XFEL) utilizes electrons that are pushed through an accelerator to a state where they produce hard (high energy) X-rays. The electrons are then skirted off, leaving the hard X-rays to hit the crystal and produce diffraction. With high-quality focusing optics different foci sizes available, allowing the beam to be focused for nanocrystal diffraction.[44] Since X-rays harm and destroy the crystals normally, Serial Femtosecond Crystallography (SFX) uses a pump injector to produce a stream of crystals that capitalize on the "diffraction-before-destruction" method where they will be hit by the X-rays, diffract, and then be destroyed by radiation damage.[45] All of the diffraction patterns are collected and processed to produce a data set that can then be used for structure determination. Partiality models with and without refinement exist of crystalline Cry and Cyt toxin proteins due to the small size of the cell-grown crystals, until recent successes in X-ray crystallography. The structural study of Cry3A displayed a proof of concept for determining the structure from *in vivo* (*Bt*) produced crystals and was determined by XFEL crystallography.[46] Cry3A's structure; however, was previously known and thus not a novel structure.[47,48] These *in vivo* produced crystals have been pursued further, including *Trypanosoma brucei* cathepsin B grown in insect cells[49] and the BinA/B subunits[50], which took the proof-of-concept study of Cry3A and solved *de novo* structures via XFEL studies. Cry11A/B toxins were first discovered together during a toxicity study to characterize new toxin proteins/crystals in *Bt susp. israelensis (Bti)*[51], but can also be produced in *Bt susp. jegathesan* and *subsp. medellin*. They are two individual toxins, thus different isoforms, that function as a binary toxin; however, Cry11B was 10-fold more toxic alone than its counterpart, Cry11A. Toxicity levels have been observed to have a 50% lethal concentration [LC_{50}] = 1.7 ng/mL[51] against fourth-instar *Culex quinquefasciatus* the highest mosquitocidal protein produced by *Bti*. Cry11Ba is around 72 kDa and target mosquitoes of *Aedes aegypti* and *Aedes albopictus*, both of which are vectors for Zika and Dengue viruses.[43] Biochemically, little is known about these proteins and the mechanism by which they crystallize *in vivo* or activate upon dissolution in the gut. While an unusual process, *Bt* have evolved to develop these crystals over thousands of years to target the specific insects, but have been able to prevent the pest's from evolving a resistance against the toxin.[52]

1.7 Electron Microscopy and Ultrastructures in Cellular Biology

As the field of cellular biology discovers more intricate details about how cells work, the technology and techniques available become the limiting factor in obtaining new data and results. With its variety of usages, light microscopy is one of the most utilized techniques for cell studies. Histochemical staining of cross-sections or fluorescence detection of tagged species are examples of different characterization studies that can be performed. The main obstacle that arises from light microscopy is limited resolution due to it being dependent on the wavelength of light passing through. With the current advancements being made in the microscopy field, these obstacles are being overcome. Both scanning and transmission electron microscopy, SEM and TEM respectively, have provided the ability to overcome this obstacle with higher resolution images that allow cellular structures to be distinguished and identified. SEM has been useful in imaging extracellular structural features, while TEM penetrate the cell to view features, like the nucleus and its interaction with DNA, specific organelles, and various pathways, in the intracellular space. The limitations of TEM are dependent upon sample thickness, where higher energy electron sources can penetrate thick samples further, but samples that are too thick, must undergo additional modification, ie. Focused Ion Beam (FIB) milling, to combat the thickness limitation and thin the sample to allow the passthrough of electrons.[53–55] Tomography continues to propel the field forward by being able to image through solid samples and produce a cross-section representation.[53] There are two imaging methods for tomography, TEM and scanning transmission electron microscopy (STEM), that pass electrons through the sample and provide information of these regions as well. One of STEM's biggest advantages is producing clearer images for thick specimens composing these cross-sections with a large amount of information and for this reason have been discussed and conducted within biological samples. STEM also utilizes dynamic focus, which allows for the scanning beam to focus flexibly to increase defocus in biological and cell samples that display distinct levels of variability.[56] This is an exceptionally powerful technique in identifying new interactions and differentiating minute morphological differences within the cell. While a relatively new and evolving technique for biological field, tomography provides a unique opportunity to determine more *in vivo* ultrastructures, which will aid in a deeper understanding about cells' architecture, assembly, and cellular structures' functions.

With the recent developments in structure determination, cryo-electron tomography (cryo-ET) can study cellular interactions and solve these complicated protein assemblies, eg. crystals, within the cell. This technique also provides resolution at a crucial range, eg. 1-10 µm, which is at the limitation limits for many techniques.[57] Cryo-ET is the process of freezing a (biological) sample within vitrified ice which freezes so rapidly that it preserves the native (cellular) structures and tilted within the microscope to capture images at different angles. This tilt-series of images can be reconstructed to view the sample computationally and produce a 3D tomogram. By studying these structures, more information can be obtained about determining stabilized crystalline conditions and the metabolic and chemical environment that facilitates self-assembly.[58]

1.8 Figures

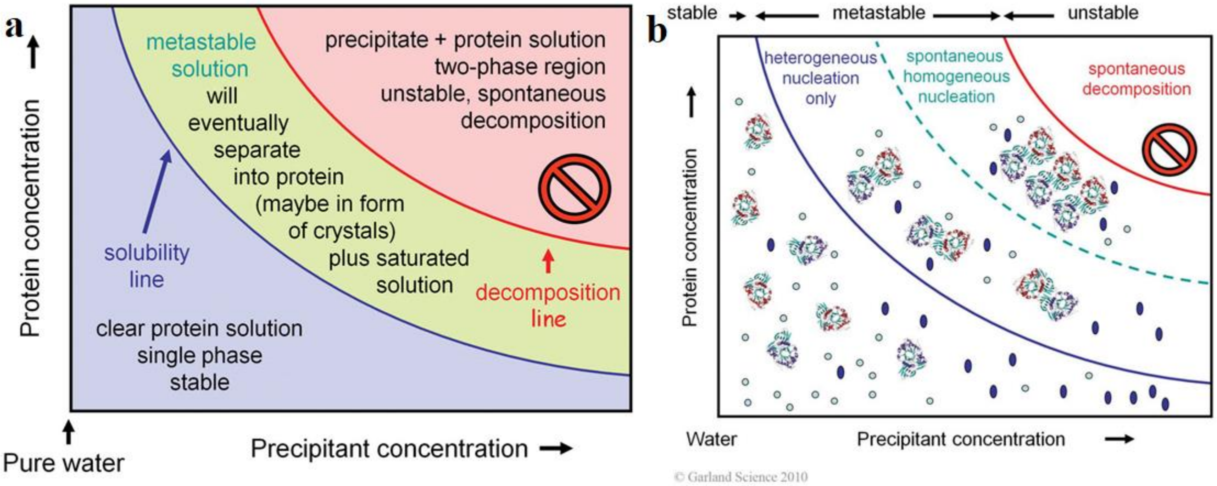


Figure 1.1 Graphical display of protein crystallization mechanics in solution. [20] (**a**) The different solution zones and states of the protein dependent upon concentration. The lower concentration passes the solubility line to reach a metastable solution which is the ideal region for crystallization. Going past the degradation line will cause the solution's protein aggregate and precipitate out of solution due to high concentration. (**b**) A graphical display showing how individual protein monomers interact in differing concentration solutions. The ideal regions are the heterogeneous and spontaneous homogeneous nucleation concentration regions.

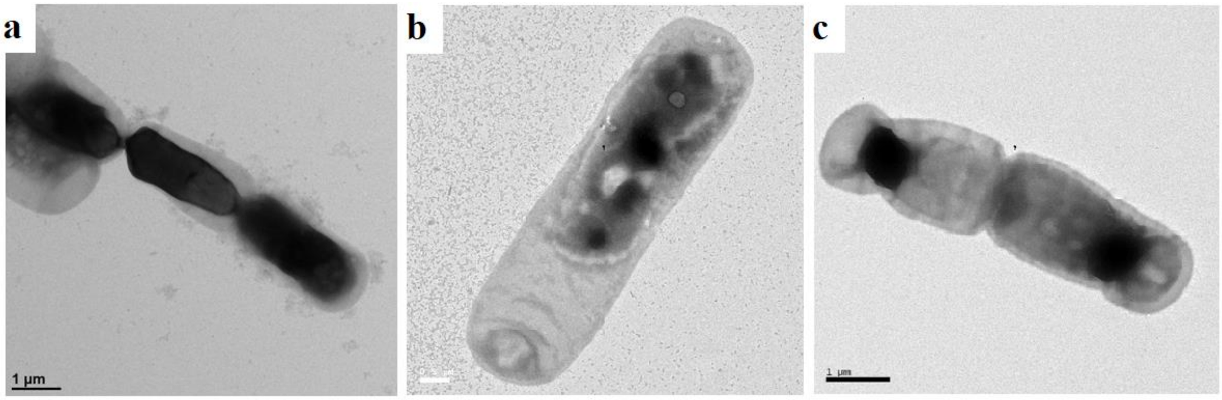


Figure 1.2 Overview of sporulation stages of *Bacillus thuringiensis* subsp. *israelensis*. (**a**) At the beginning of sporulation, the *Bti* cells make a dense mesh for the nucleoid of the parasporal body. (**b**) As the parasporal body continues to form, asymmetric division and development of the crystalline inclusion and endospore occurs on opposite poles of the cells. (**c**) As the crystalline inclusion and endospore reach maturation, the parasporal body membrane begins to detach from the cell membrane before clear definition of crystal and endospore and natural lysis occurs.

1.9 References

(1) Johnson, J. L.; Craig, E. A. Protein Folding In Vivo: Unraveling Complex Pathways. *Cell* **1997**, *90* (2), 201–204. https://doi.org/10.1016/S0092-8674(00)80327-X.

(2) Sun, P. D.; Foster, C. E.; Boyington, J. C. Overview of Protein Structural and Functional Folds. *Curr Protoc Protein Sci* **2004**, *35* (1). https://doi.org/10.1002/0471140864.ps1701s35.

(3) Perham, R. N. Self-Assembly of Biological Macromolecules. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* **1975**, *272* (915), 123–136. https://doi.org/10.1098/rstb.1975.0075.

(4) Sorokina, I.; Mushegian, A. Modeling Protein Folding in Vivo. *Biol Direct* **2018**, *13* (1), 13. https://doi.org/10.1186/s13062-018-0217-6.

(5) Bai, Y.; Luo, Q.; Liu, J. Protein Self-Assembly via Supramolecular Strategies. *Chem Soc Rev* **2016**, *45* (10), 2756–2767. https://doi.org/10.1039/C6CS00004E.

(6) Ozin, G. A.; Hou, K.; Lotsch, B. V; Cademartiri, L.; Puzzo, D. P.; Scotognella, F.; Ghadimi, A.; Thomson, J. Nanofabrication by Self-Assembly. *Materials Today* **2009**, *12* (5), 12–23. https://doi.org/https://doi.org/10.1016/S1369-7021(09)70156-7.

(7) Cademartiri, L.; Bishop, K. J. M.; Snyder, P. W.; Ozin, G. A. Using Shape for Self-Assembly. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2012**, *370* (1969), 2824–2847. https://doi.org/10.1098/rsta.2011.0254.

(8) Levinthal, C. Are There Pathways for Protein Folding? *Journal de Chimie Physique* **1968**, *65*, 44–45. https://doi.org/10.1051/jcp/1968650044.

(9) Anfinsen, C. B. Principles That Govern the Folding of Protein Chains. *Science (1979)* **1973**, *181* (4096), 223–230. https://doi.org/10.1126/science.181.4096.223.

(10) Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. The Protein Folding Problem. *Annu. Rev. Biophys.* **2008**, *37*, 289–316.

(11) Dill, K. A. The Stabilities of Globular Proteins. *Protein Eng* **1987**, *1*, 187–192.

(12) Dill, K. A.; Chan, H. S. From Levinthal to Pathways to Funnels. *Nat Struct Biol* **1997**, *4* (1), 10–19. https://doi.org/10.1038/nsb0197-10.

(13) Dill, K. A. Polymer Principles and Protein Folding. *Protein Science* **1999**, *8* (6), 1166–1180. https://doi.org/10.1110/ps.8.6.1166.

(14) Dill, K. A.; MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science (1979)* **2012**, *338* (6110), 1042–1046. https://doi.org/10.1126/science.1219021.

(15) Fersht, A. R.; Leatherbarrow, R. J.; Wells, T. N. C. Structure-Activity Relationships in Engineered Proteins: Analysis of Use of Binding Energy by Linear Free Energy Relationships. *Biochemistry* **1987**, *26* (19), 6030–6038.

(16) Ozkan, S. B.; Bahar, I.; Dill, K. A. Transition States and the Meaning of Φ-Values in Protein Folding Kinetics. *Nat Struct Biol* **2001**, *8* (9), 765–769. https://doi.org/10.1038/nsb0901-765.

(17) Wako, H.; Abe, H. Characterization of Protein Folding by a Φ-Value Calculation with a Statistical-Mechanical Model. *Biophys Physicobiol* **2016**, *13* (0), 263–279. https://doi.org/10.2142/biophysico.13.0_263.

(18) Varela, A. E.; England, K. A.; Cavagnero, S. Kinetic Trapping in Protein Folding. *Protein Engineering, Design and Selection* **2019**, *32* (2), 103–108. https://doi.org/10.1093/protein/gzz018.

(19) McPherson, A.; Gavira, J. A. Introduction to Protein Crystallization. *Acta Crystallogr F Struct Biol Commun* **2014**, *70* (1), 2–20. https://doi.org/10.1107/S2053230X13033141.

(20) Rupp, B. *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*; Garland Science, 2009.

(21) Niedzialkowska, E.; Gasiorowska, O.; Handing, K. B.; Majorek, K. A.; Porebski, P. J.; Shabalin, I. G.; Zasadzinska, E.; Cymborowski, M.; Minor, W. Protein Purification and Crystallization Artifacts: The Tale Usually Not Told. *Protein Science* **2016**, *25* (3), 720–733. https://doi.org/10.1002/pro.2861.

(22) Mitchell, S. D.; Gronenborn, A. M. After Fifty Years, Why Are Protein X-Ray Crystallographers Still in Business? *Br J Philos Sci* **2017**.

(23) Dyakova, Yu. A.; Kovalchuk, M. V. Protein Self-Assembly in Crystals and Films. *Crystallography Reports* **2022**, *67* (5), 772–790. https://doi.org/10.1134/S1063774522050030.

(24) Feng, R.; Gruebele, M.; Davis, C. M. Quantifying Protein Dynamics and Stability in a Living Organism. *Nat Commun* **2019**, *10* (1), 1179.

(25) Mackenzie, C. G.; Mackenzie, J. B.; Beck, P. The Effect of PH on Growth, Protein Synthesis, and Lipid-Rich Particles of Cultured Mammalian Cells. *J Cell Biol* **1961**, *9* (1), 141–156.

(26) Cuéllar-Cruz, M. Synthesis of Inorganic and Organic Crystals Mediated by Proteins in Different Biological Organisms. A Mechanism of Biomineralization Conserved throughout Evolution in All Living Species. *Progress in Crystal Growth and Characterization of Materials* **2017**, *63* (3), 94–103. https://doi.org/10.1016/j.pcrysgrow.2017.07.001.

(27) De Yoreo, J. J.; Gilbert, P. U. P. A.; Sommerdijk, N. A. J. M.; Penn, R. L.; Whitelam, S.; Joester, D.; Zhang, H.; Rimer, J. D.; Navrotsky, A.; Banfield, J. F.; Wallace, A. F.;

Michel, F. M.; Meldrum, F. C.; Cölfen, H.; Dove, P. M. Crystallization by Particle Attachment in Synthetic, Biogenic, and Geologic Environments. *Science (1979)* **2015**, *349* (6247). https://doi.org/10.1126/science.aaa6760.

(28) Young, B.; O'Dowd, Geraldine.; Woodford, Phillip. *Wheater's Functional Histology : A Text and Colour Atlas.*, 6th ed.; Churchill Livingston/Elsevier: Philadelphia, 2014.

(29) Zhang, J.; Young, R. H.; Arseneau, J.; Scully, R. E. Ovarian Stromal Tumors Containing Lutein or Leydig Cells (Luteinized Thecomas and Stromal Leydig Cell Tumors)—A Clinicopathological Analysis of Fifty Cases. *International Journal of Gynecological Pathology* **1982**, *1* (3), 270–285. https://doi.org/10.1097/00004347-198203000-00004.

(30) Mesa, H.; Gilles, S.; Smith, S.; Dachel, S.; Larson, W.; Manivel, J. C. The Mystery of the Vanishing Reinke Crystals. *Hum Pathol* **2015**, *46* (4), 600–606. https://doi.org/10.1016/j.humpath.2015.01.004.

(31) Sansinenea, E. Discovery and Description of Bacillus Thuringiensis. *Bacillus thuringiensis biotechnology* **2012**, 3–18.

(32) Aida, J.; Takubo, K.; Ishikawa, N.; Matsuda, Y.; Arai, T.; Ishiwata, T. Distinction Between Ductal Involvement and Invasive Nests in Esophageal Squamous Cell Carcinoma. In *LABORATORY INVESTIGATION*; NATURE PUBLISHING GROUP 75 VARICK ST, 9TH FLR, NEW YORK, NY 10013-1917 USA, 2018; Vol. 98, pp 240–241.

(33) Ishiwata, S. On a Kind of Severe Flacherie (Sotto Disease). *Dainihon sanshi kaiho* **1901**, *114*, 1–5.

(34) Ishiwata, S. Examination of Sotto Disease Bacillus. *Kyoto Sangyo Koshujo Sanji Hokoku* **1902**, *2*, 346–347.

(35) Ishiwata, S. Sotto Bacillus Study. *Japenese). Dainihon Sanshi Kaiho* **1905**, *14*, 24–28.

(36) Berliner, E. Die" Schlaffsucht" Der Mehlmottenraupe. *Z. Gesamte Getreidewes* **1911**, *3*, 63–70.

(37) Bravo, A.; Gill, S. S.; Soberón, M. Mode of Action of Bacillus Thuringiensis Cry and Cyt Toxins and Their Potential for Insect Control. *Toxicon* **2007**, *49* (4), 423–435.

(38) Bechtel, D. B.; Bulla, L. A. Electron Microscope Study of Sporulation and Parasporal Crystal Formation in *Bacillus Thuringiensis*. *J Bacteriol* **1976**, *127* (3), 1472–1481. https://doi.org/10.1128/jb.127.3.1472-1481.1976.

(39) Federici, B. A.; Park, H.-W.; Sakano, Y. Insecticidal Protein Crystals of Bacillus Thuringiensis. *Inclusions in prokaryotes* **2006**, 195–236.

(40) Pigott, C. R.; Ellar, D. J. Role of Receptors in Bacillus Thuringiensis Crystal Toxin Activity. *Microbiology and molecular biology reviews* **2007**, *71* (2), 255–281.

(41) Soberón, M.; López-Díaz, J. A.; Bravo, A. Cyt Toxins Produced by Bacillus Thuringiensis: A Protein Fold Conserved in Several Pathogenic Microorganisms. *Peptides (N.Y.)* **2013**, *41*, 87–93.

(42) Xu, C.; Wang, B.-C.; Yu, Z.; Sun, M. Structural Insights into Bacillus Thuringiensis Cry, Cyt and Parasporin Toxins. *Toxins (Basel)* **2014**, *6* (9), 2732–2770.

(43) Park, H.-W.; Bideshi, D. K.; Federici, B. A. Recombinant Strain of Bacillus Thuringiensis Producing Cyt1A, Cry11B, and the Bacillus Sphaericus Binary Toxin. *Appl Environ Microbiol* **2003**, *69* (2), 1331–1334.

(44) Spence, J. C. H. XFELs for Structure and Dynamics in Biology. *IUCrJ* **2017**, *4* (4), 322–339.

(45) Johansson, L. C.; Stauch, B.; Ishchenko, A.; Cherezov, V. A Bright Future for Serial Femtosecond Crystallography with XFELs. *Trends Biochem Sci* **2017**, *42* (9), 749–762.

(46) Sawaya, M. R.; Cascio, D.; Gingery, M.; Rodriguez, J.; Goldschmidt, L.; Colletier, J.-P.; Messerschmidt, M. M.; Boutet, S.; Koglin, J. E.; Williams, G. J. Protein Crystal Structure Obtained at 2.9 Å Resolution from Injecting Bacterial Cells into an X-Ray Free-Electron Laser Beam. *Proceedings of the National Academy of Sciences* **2014**, *111* (35), 12769–12774.

(47) Loseva, O. I.; Tiktopulo, E. I.; Vasiliev, V. D.; Nikulin, A. D.; Dobritsa, A. P.; Potekhin, S. A. Structure of Cry3A δ-Endotoxin within Phospholipid Membranes. *Biochemistry* **2001**, *40* (47), 14143–14151.

(48) Li, J.; Carroll, J.; Ellar, D. J. Crystal Structure of Insecticidal δ-Endotoxin from Bacillus Thuringiensis at 2.5 Å Resolution. *Nature* **1991**, *353* (6347), 815–821.

(49) Duszenko, M.; Redecke, L.; Mudogo, C. N.; Sommer, B. P.; Mogk, S.; Oberthuer, D.; Betzel, C. In Vivo Protein Crystallization in Combination with Highly Brilliant Radiation Sources Offers Novel Opportunities for the Structural Analysis of Post-Translationally Modified Eukaryotic Proteins. *Acta Crystallogr F Struct Biol Commun* **2015**, *71* (8), 929–937.

(50) 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. De Novo Phasing with X-Ray Laser Reveals Mosquito Larvicide BinAB Structure. *Nature* **2016**, *539* (7627), 43–47.

(51) Park, H.-W.; Delécluse, A.; Federici, B. A. Construction and Characterization of a Recombinant Bacillus Thuringiensis Subsp. Israelensis Strain That Produces Cry11B. *J Invertebr Pathol* **2001**, *78* (1), 37–44.

(52) Bideshi, D. K.; Park, H.-W.; Hice, R. H.; Wirth, M. C.; Federici, B. A. Highly Effective Broad Spectrum Chimeric Larvicide That Targets Vector Mosquitoes Using a Lipophilic Protein. *Sci Rep* **2017**, *7* (1), 11282.

(53) Yakushevska, A. E.; Lebbink, M. N.; Geerts, W. J. C.; Spek, L.; Van Donselaar, E. G.; Jansen, K. A.; Humbel, B. M.; Post, J. A.; Verkleij, A. J.; Koster, A. J. STEM Tomography in Cell Biology. *J Struct Biol* **2007**, *159* (3), 381–391.

(54) Lučić, V.; Leis, A.; Baumeister, W. Cryo-Electron Tomography of Cells: Connecting Structure and Function. *Histochem Cell Biol* **2008**, *130*, 185–196.

(55) Lučić, V.; Rigort, A.; Baumeister, W. Cryo-Electron Tomography: The Challenge of Doing Structural Biology in Situ. *Journal of Cell Biology* **2013**, *202* (3), 407–419.

(56) Aoyama, K.; Takagi, T.; Hirase, A.; Miyazawa, A. STEM Tomography for Thick Biological Specimens. *Ultramicroscopy* **2008**, *109* (1), 70–80.

(57) Quentin, D.; Raunser, S. Electron Cryomicroscopy as a Powerful Tool in Biomedical Research. *J Mol Med* **2018**, *96* (6), 483–493. https://doi.org/10.1007/s00109-018-1640-y.

(58) Kirkwood, J.; Hargreaves, D.; O'Keefe, S.; Wilson, J. Analysis of Crystallization Data in the Protein Data Bank. *Acta Crystallogr F Struct Biol Commun* **2015**, *71* (10), 1228–1234.

(59) Whiteley, H. R.; Schnepf, H. E. The Molecular Biology of Parasporal Crystal Body Formation in Bacillus Thuringiensis. *Annual Reviews in Microbiology* **1986**, *40* (1), 549–576.

(60) Feng, R.; Gruebele, M.; Davis, C. M. Quantifying Protein Dynamics and Stability in a Living Organism. *Nat Commun* **2019**, *10* (1), 1179.

(61) Johnston, K.; Joglekar, A.; Hori, T.; Suzuki, A.; Fukagawa, T.; Salmon, E. D. Vertebrate Kinetochore Protein Architecture: Protein Copy Number. *Journal of Cell Biology* **2010**, *189* (6), 937–943.

(62) Schönherr, R.; Rudolph, J. M.; Redecke, L. Protein Crystallization in Living Cells. *Biol Chem* **2018**, *399* (7), 751–772.

(63) Khanna, K.; Lopez-Garrido, J.; Zhao, Z.; Watanabe, R.; Yuan, Y.; Sugie, J.; Pogliano, K.; Villa, E. The Molecular Architecture of Engulfment during Bacillus Subtilis Sporulation. *Elife* **2019**, *8*, e45257.

(64) Kushner, D. J. Self-Assembly of Biological Structures. *Bacteriol Rev* **1969**, *33* (2), 302–345.

(65) Hingorani, K. S.; Gierasch, L. M. Comparing Protein Folding in Vitro and in Vivo: Foldability Meets the Fitness Challenge. *Curr Opin Struct Biol* **2014**, *24*, 81–90.

(66) Pryor Jr, A.; Yang, Y.; Rana, A.; Gallagher-Jones, M.; Zhou, J.; Lo, Y. H.; Melinte, G.; Chiu, W.; Rodriguez, J. A.; Miao, J. GENFIRE: A Generalized Fourier Iterative Reconstruction Algorithm for High-Resolution 3D Imaging. *Sci Rep* **2017**, *7* (1), 10409.

(67) Milne, J. L. S.; Subramaniam, S. Cryo-Electron Tomography of Bacteria: Progress, Challenges and Future Prospects. *Nat Rev Microbiol* **2009**, *7* (9), 666–675.

(68) Heimpel, A. M. A Critical Review of Bacillus Thuringiensis Var. Thuringiensis Berliner and Other Crystalliferous Bacteria. *Annu Rev Entomol* **1967**, *12* (1), 287–322.

(69) Lopez-Garrido, J.; Ojkic, N.; Khanna, K.; Wagner, F. R.; Villa, E.; Endres, R. G.; Pogliano, K. Chromosome Translocation Inflates Bacillus Forespores and Impacts Cellular Morphology. *Cell* **2018**, *172* (4), 758–770.

(70) Guerchicoff, A.; Delécluse, A.; Rubinstein, C. P. The Bacillus Thuringiensis Cyt Genes for Hemolytic Endotoxins Constitute a Gene Family. *Appl Environ Microbiol* **2001**, *67* (3), 1090–1096.

(71) Hirata, E.; Yukinaga, H.; Kamioka, Y.; Arakawa, Y.; Miyamoto, S.; Okada, T.; Sahai, E.; Matsuda, M. In Vivo Fluorescence Resonance Energy Transfer Imaging Reveals Differential Activation of Rho-Family GTPases in Glioblastoma Cell Invasion. *J Cell Sci* **2012**, *125* (4), 858–868.

(72) Bates, M.; Jones, S. A.; Zhuang, X. Stochastic Optical Reconstruction Microscopy (STORM): A Method for Superresolution Fluorescence Imaging. *Cold Spring Harb Protoc* **2013**, *2013* (6), pdb-top075143.

(73) Diaz-Mendoza, M.; Bideshi, D. K.; Federici, B. A. A 54-Kilodalton Protein Encoded by PBtoxis Is Required for Parasporal Body Structural Integrity in Bacillus Thuringiensis Subsp. Israelensis. *J Bacteriol* **2012**, *194* (6), 1562–1571.

(74) Berry, C.; O'Neil, S.; Ben-Dov, E.; Jones, A. F.; Murphy, L.; Quail, M. A.; Holden, M. T. G.; Harris, D.; Zaritsky, A.; Parkhill, J. Complete Sequence and Organization of PBtoxis, the Toxin-Coding Plasmid of Bacillus Thuringiensis Subsp. Israelensis. *Appl Environ Microbiol* **2002**, *68* (10), 5082–5095.

(75) Cabeen, M. T.; Jacobs-Wagner, C. Skin and Bones: The Bacterial Cytoskeleton, Cell Wall, and Cell Morphogenesis. *J Cell Biol* **2007**, *179* (3), 381.

(76) Gustafsson, M. G. L. Nonlinear Structured-Illumination Microscopy: Wide-Field Fluorescence Imaging with Theoretically Unlimited Resolution. *Proceedings of the National Academy of Sciences* **2005**, *102* (37), 13081–13086.

(77) Jeong, H.; Park, S.-H.; Choi, S.-K. Genome Sequence of the Acrystalliferous Bacillus Thuringiensis Serovar Israelensis Strain 4Q7, Widely Used as a Recombination Host. *Genome Announc* **2014**, *2* (2), e00231-14.

(78) Ito, A.; Sasaguri, Y.; Kitada, S.; Kusaka, Y.; Kuwano, K.; Masutomi, K.; Mizuki, E.; Akao, T.; Ohba, M. A Bacillus Thuringiensis Crystal Protein with Selective Cytocidal Action to Human Cells. *Journal of Biological Chemistry* **2004**, *279* (20), 21282–21286.

(79) Ammons, D.; Toal, G.; Roman, A.; Rojas-Avelizapa, L. I.; Ventura-Suárez, A.; Rampersad, J. Cry-like Genes, in an Uncommon Gene Configuration, Produce a Crystal That Localizes within the Exosporium When Expressed in an Acrystalliferous Strain of Bacillus Thuringiensis. *FEMS Microbiol Lett* **2016**, *363* (4), fnw010.

(80) Kittisopikul, M.; Virtanen, L.; Taimen, P.; Goldman, R. D. Quantitative Analysis of Nuclear Lamins Imaged by Super-Resolution Light Microscopy. *Cells* **2019**, *8* (4), 361.

(81) Dye, N. A.; Shapiro, L. The Push and Pull of the Bacterial Cytoskeleton. *Trends Cell Biol* **2007**, *17* (5), 239–245.

(82) Dadd, R. H. Alkalinity within the Midgut of Mosquito Larvae with Alkaline-Active Digestive Enzymes. *J Insect Physiol* **1975**, *21* (11), 1847–1853.

(83) Palermo, V.; Schwartz, E.; Finlayson, C. E.; Liscio, A.; Otten, M. B. J.; Trapani, S.; Müllen, K.; Beljonne, D.; Friend, R. H.; Nolte, R. J. M. Macromolecular Scaffolding: The Relationship between Nanoscale Architecture and Function in Multichromophoric Arrays for Organic Electronics. *Advanced Materials* **2010**, *22* (8), E81–E88.

(84) Krissinel, E.; Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J Mol Biol* **2007**, *372* (3), 774–797.

Chapter 2

Serial femtosecond structural determination and biochemical analysis of *in vivo* Cry11Ba crystalline inclusions

This chapter is based on the published article "*De novo* determination of mosquitocidal Cry11Aa and Cry11Ba structures from naturally-occurring nanocrystals." [1]

2.1 Summary

In this article, the structures of Cry11Aa and Cry11Ba are presented via *de novo* phasing along with biochemical assays conducted to probe the interactions that were postulated to stabilize and/or promote the crystalline lattice that these naturally form during sporulation. These interactions were previously undetermined for Cry11Aa and Cry11Ba, even though the previously studied crystalline (Cry) and cytolytic (Cyt) pesticidal proteins found have displayed 1 of 2 structural homologies, most likely due to the residues driving crystal formation being unique for each pesticidal protein. The other challenge that was posed with these crystals was their size. Since the crystalline inclusions produced during sporulation are 5 µm or less, this pushes the limit of canonical X-ray crystallography techniques and increases the difficulty due to the aforementioned size affecting the protein monomer to crystal ratio. For example, peptide crystals of this size have been solved, but this is due to the size of the peptide monomers being very small and thus robust throughout the crystal. For these crystals of the same size, there should be significantly less with a protein monomer 10-100 times the number of residues in the peptide monomer. Despite many attempts at determining the Cry11Aa and Cry11Ba structures prior, this time proved fruitful as heavy atom soaking with caged-terbium compound (Tb-Xo4) allowed sufficient phasing data for single-wavelength anomalous dispersion (SAD) of Cry11Aa and then lead to a molecular replacement (MR) structure for Cry11Ba. Furthermore, solving the structure of these proteins from their crystalline inclusion state was of the utmost importance to gain more insight into what is naturally packaged into the crystals and reduce any perturbations that can occur when utilizing X-ray crystallography, ie. crystal packing interaction artifacts and interactions with buffers that can enhance or prevent interactions for multimerization. From their structures, point mutations targeting the sites of inter- and intra-faces were selected and displayed sufficient ability to change the crystal morphology and stability in a gradation for both factors. While the visible changes to the crystalline inclusions were often significant, the solubility assays elucidated that the crystal stability had been altered and furthermore shifted the pH trigger for both Cry11Aa and Cry11Ba. By observing changes from these predicted sites of importance, the differences in protein scaffolding of the Cry11Aa and Cry11Ba crystals could be seen and Cry11Aa was found to have a weaker tetramer-tetramer interface interaction, while Cry11Ba has a stronger tetramer-tetramer interaction. Cry11Ba's stronger interaction causes two subunits within the asymmetric unit and could be what increases Cry11Ba's toxicity 7-fold increase than Cry11Aa against *A. aegypti* and *A. stephensi* and 37-fold increase against *C. pipiens*. With these results, crystalline inclusion morphology, protein packing, and stability have been further illuminated and will allow more informed studies on the crystalline inclusion formation *in vivo* during sporulation with further understanding the mode of activation for these pesticidal toxin proteins.

(1) Tetreau, G.; Sawaya, M. R.; De Zitter, E.; Andreeva, E. A.; Banneville, A.-S.; Schibrowsky, N. A.; Coquelle, N.; Brewster, A. S.; Grünbein, M. L.; Kovacs, G. N.; Hunter, M. S.; Kloos, M.; Sierra, R. G.; Schiro, G.; Qiao, P.; Stricker, M.; Bideshi, D.; Young, I. D.; Zala, N.; Engilberge, S.; Gorel, A.; Signor, L.; Teulon, J.-M.; Hilpert, M.; Foucar, L.; Bielecki, J.; Bean, R.; de Wijn, R.; Sato, T.; Kirkwood, H.; Letrun, R.; Batyuk, A.; Snigireva, I.; Fenel, D.; Schubert, R.; Canfield, E. J.; Alba, M. M.; Laporte, F.; Després, L.; Bacia, M.; Roux, A.; Chapelle, C.; Riobé, F.; Maury, O.; Ling, W. L.; Boutet, S.; Mancuso, A.; Gutsche, I.; Girard, E.; Barends, T. R. M.; Pellequer, J.-L.; Park, H.-W.; Laganowsky, A. D.; Rodriguez, J.; Burghammer, M.; Shoeman, R. L.; Doak, R. B.; Weik, M.; Sauter, N. K.; Federici, B.; Cascio, D.; Schlichting, I.; Colletier, J.-P. De Novo Determination of Mosquitocidal Cry11Aa and Cry11Ba Structures from Naturally-Occurring Nanocrystals. *Nat Commun* **2022**, *13* (1), 4376. https://doi.org/10.1038/s41467-022-31746-x.

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-022-31746-x OPEN

De novo determination of mosquitocidal Cry11Aa and Cry11Ba structures from naturally-occurring nanocrystals

Guillaume Tetreau[1,18], Michael R. Sawaya[2,18], Elke De Zitter[1,18], Elena A. Andreeva[1,3], Anne-Sophie Banneville[1], Natalie A. Schibrowsky[2,4], Nicolas Coquelle[5], Aaron S. Brewster[6], Marie Luise Grünbein[3], Gabriela Nass Kovacs[3], Mark S. Hunter[7], Marco Kloos[3,8], Raymond G. Sierra[7], Giorgio Schiro[1], Pei Qiao[9], Myriam Stricker[3], Dennis Bideshi[10,11], Iris D. Young[6], Ninon Zala[1], Sylvain Engilberge[1], Alexander Gorel[3], Luca Signor[1], Jean-Marie Teulon[1], Mario Hilpert[3], Lutz Foucar[3], Johan Bielecki[8], Richard Bean[8], Raphael de Wijn[8], Tokushi Sato[8], Henry Kirkwood[8], Romain Letrun[8], Alexander Batyuk[7], Irina Snigireva[12], Daphna Fenel[1], Robin Schubert[8], Ethan J. Canfield[13], Mario M. Alba[14], Frédéric Laporte[15], Laurence Després[15], Maria Bacia[1], Amandine Roux[16], Christian Chapelle[17], François Riobé[16], Olivier Maury[16], Wai Li Ling[1], Sébastien Boutet[7], Adrian Mancuso[8], Irina Gutsche[1], Eric Girard[1], Thomas R. M. Barends[3], Jean-Luc Pellequer[1], Hyun-Woo Park[10,11], Arthur D. Laganowsky[9], Jose Rodriguez[2,4], Manfred Burghammer[12], Robert L. Shoeman[3], R. Bruce Doak[3], Martin Weik[1], Nicholas K. Sauter[6], Brian Federici[10], Duilio Cascio[2], Ilme Schlichting[3] & Jacques-Philippe Colletier[1]✉

Cry11Aa and Cry11Ba are the two most potent toxins produced by mosquitocidal *Bacillus thuringiensis* subsp. *israelensis* and *jegathesan*, respectively. The toxins naturally crystallize within the host; however, the crystals are too small for structure determination at synchrotron sources. Therefore, we applied serial femtosecond crystallography at X-ray free electron lasers to in vivo-grown nanocrystals of these toxins. The structure of Cry11Aa was determined *de novo* using the single-wavelength anomalous dispersion method, which in turn enabled the determination of the Cry11Ba structure by molecular replacement. The two structures reveal a new pattern for in vivo crystallization of Cry toxins, whereby each of their three domains packs with a symmetrically identical domain, and a cleavable crystal packing motif is located within the protoxin rather than at the termini. The diversity of in vivo crystallization patterns suggests explanations for their varied levels of toxicity and rational approaches to improve these toxins for mosquito control.

A full list of author affiliations appears at the end of the paper.

The most commonly used biological insecticides for controlling mosquito and black fly vector populations are produced by the bacterium *Bacillus thuringiensis* subsp. *israelensis* (*Bti*), discovered in Israel in 1976[1]. These products target the larval stage of a wide variety of vectors, and due to their high efficacy and environmental safety, have replaced broad spectrum synthetic chemical insecticides in many vector control programs. These include Anopheles gambiae and related species that transmit malaria, as well as numerous Culex and Aedes species that spread viruses such as those that cause West Nile Encephalitis and Yellow Fever. *Bti* products are also used in Africa to regulate black fly species responsible for vectoring the filarial worms that cause River Blindness. Aside from vector populations, they are used to control nuisance mosquitoes in the Rhine Valley in Germany, in the Camargue in southern France, and throughout the U.S., Asia, and Latin and South America, with thousands of tons applied over the past 30 years.

The highly potent mosquitocidal activity of *Bti* is due to three nanocrystalline forms of four protoxins, viz. Cyt1Aa, Cry11Aa, and co-crystallized Cry4Aa and Cry4Ba. These are produced during sporulation and are remarkably stable in a variety of conditions, but dissolve after ingestion under the high alkaline pH levels characteristic of the larval mosquito midgut[2]. Solubilized protoxins are activated by insect gut proteases enabling binding to gut cell membranes, subsequent oligomerization, and ultimately gut cell lysis leading to larval death[2]. *Bti* toxins are environmentally safe because they are much more specific for target mosquitoes than broad-spectrum chemical larvicides.

The most potent of the four *Bti* toxins is Cry11Aa, but its activation and mechanism of toxicity are poorly understood, in large part because unlike Cry4Aa, Cry4Ba, and Cyt1Aa, its structure is unknown. A related toxin produced by *Bt* subsp. *jegathesan* (*Btj*) is Cry11Ba, which is seven to thirty-seven times more toxic than Cry11Aa against major mosquito vector species belonging to the genera *Aedes*, *Anopheles*, and *Culex*[3], and in some bacterial hosts appears to form slightly larger crystals. Cry11Ba's structure is also unknown, although it has been used as a replacement for Cry11Aa in recombinant strains of *Bti* to improve mosquitocidal activity significantly[3,4]. Thus, our goal was to determine the structures of Cry11Aa and Cry11Ba protoxins to help understand how they achieve formation of robust crystals labile only at alkaline pH, and to obtain structural insights for increasing the efficacy of these proteins for mosquito control.

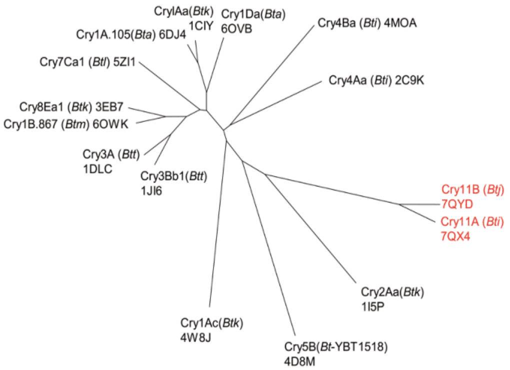


Fig. 1 Phylogentic tree of delta-endotoxins. Tree plot showing structural relatedness of Cry11Aa and Cry11Ba to the family of 13 delta-endotoxins reported in the PDB. Structural similarity is represented on a phylogenetic tree plot. Shorter connecting lines signify closer structural similarity. Cry11Aa and Cry11Ba are structurally similar to each other, but distant from the other 13 delta-endotoxins. The closest structural homolog of Cry11Aa and Cry11Ba is Cry2Aa. The PDB ID codes are reported next to the toxin name. Species of origin are annotated in parentheses. *Bta*, *Bti*, *Btj*, *Btk*, *Btm* and *Btt* abbreviate insectidal *Bacillus thuringiensis* subspecies *aizawai*, *israelensis*, *jegathesan*, *kurstaki*, *medellin* and *tenebrionis*, respectively. *Bt*-YBT1518 is a strain displaying nematocidal activity.

Structure determination of Cry11Aa and Cry11Ba protoxins from natural nanocrystals requires cutting-edge technology. Conventional crystallography is limited to projects in which crystals are sufficiently large to mount and oscillate individually in a synchrotron X-ray beam. In the past, crystals of activated Cry4Aa[5], Cry4Ba[6] and Cyt1Aa[7] attained sufficient size by growing these in vitro from toxins dissolved from natural nanocrystals and activating the toxins enzymatically. However, Cry11Aa and Cry11Ba do not recrystallize in vitro from dissolved nanocrystals[8]. Moreover, enzymatic activation is unwanted since our goal is to understand the pH-controlled mechanism of natural crystal dissolution. To observe the protoxin state in natural nanocrystals produced in bacterial cells, we applied serial femtosecond crystallography (SFX) at X-ray free electron lasers (XFEL)[9–11]. In the SFX experiment, high brilliance XFEL beam pulses, each lasting only ~10–50 fs, intercept a series of nanocrystals, one pulse-per-crystal, eliciting the strongest possible diffraction signal from each tiny crystal before it vaporizes, and producing a series of diffraction snapshots, later assembled into a full data set. Feasibility of this strategy had been demonstrated by the recent elucidation of the full bioactivation cascade of Cyt1Aa[12].

Our success in determining the structures of Cry11Aa and Cry11Ba protoxins highlights the capability of XFEL sources to overcome limits of small crystal size. We relied on de novo phasing of the native SFX data because all attempts at molecular replacement (MR) failed despite detectable sequence similarity with thirteen structurally-determined members of the three-domain Cry δ-endotoxin family (Fig. 1)[13–15]. We opted to derivatize our Cry11Aa nanocrystals with a recently-introduced phasing-agent, a caged-terbium compound, Tb-Xo4[16,17]. The phases obtained from single-wavelength anomalous dispersion (SAD) were sufficient to reveal the Cry11Aa protoxin structure at 2.6 Å resolution and subsequently enable phasing of the Cry11Ba protoxin structure at 2.4 Å resolution by molecular replacement. In hindsight, we attribute the failure of early MR attempts to three extra β-strands in domain II which alter the relative orientation of the three domains in Cry11 toxins.

Our studies of Cry11Aa and Cry11Ba crystals reveal a new paradigm of molecular packing among Cry δ-endotoxins reported thus far. In particular, the cleavable peptides that constitute important crystal contacts are located near the middle of the toxin sequence, rather than at the termini. Molecules pack in tetramer units, exhibiting D2 symmetry; these tetramers in turn pack in a body centered pattern (like a 3-dimensional brick-wall in which successive rows are offset by half a brick). To achieve this pattern, each of the three domains in a Cry11 molecule packs with an identical domain from a symmetry related molecule: domain I packs with domain I, II with II, and III with III. Thus, each Cry11 domain fulfills two biological roles: a dimer interface manifested in the crystalline state, and a functional role manifested in the soluble state: target recognition (domain II), oligomerization (domain III) and pore formation (domain I)[18]. Differences in the size and composition of the three packing interfaces explain shape and size differences between Cry11Aa and Cry11Ba nanocrystals. Structure-guided site-directed mutagenesis verifies which residues affect crystal size, pH sensitivity of the crystal, and toxin folding. Our results elucidate the Cry11Aa and Cry11Ba bioactivation

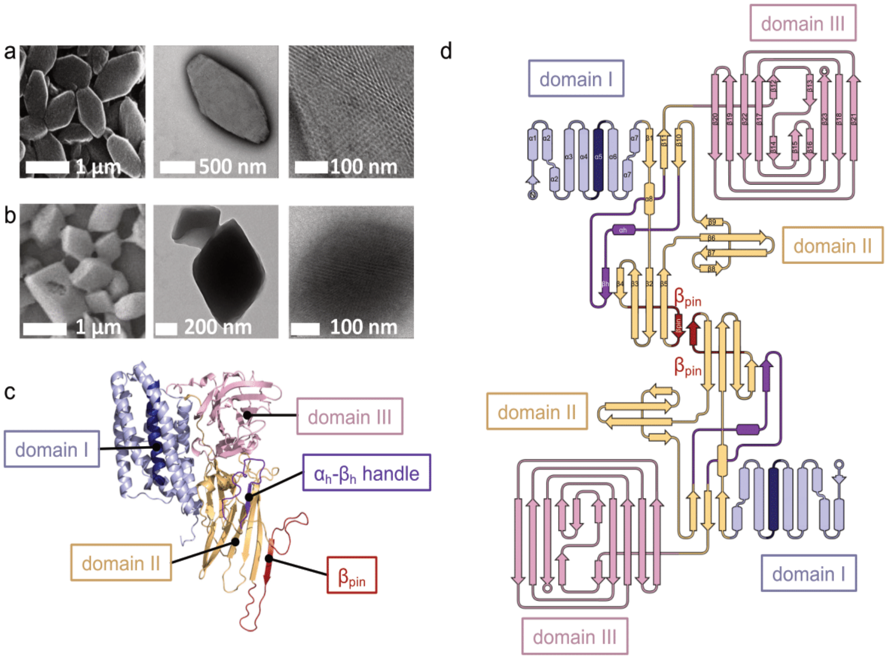


Fig. 2 Crystals and overall fold of Cry11 toxins. **a**, **b**, Scanning (left; SEM) and transmission (middle, right; TEM) electron micrographs of gold plated and negatively-stained Cry11Aa (**a**) and Cry11Ba (**b**) crystals, respectively. The right panels show a close-up view of the crystal surface. SEM and TEM experiments on Cry11Aa were repeated 7 and 2 times, respectively, while those on Cry11Ba were repeated 3 and 16 times, respectively. **c** Cry11Aa crystal structure, depicted as cartoon. Domain I is shown in light blue, except for central helix $\alpha5$ which is shown in dark blue; domain II is shown in orange except for the $\alpha_h\beta_h$-handle and β_{pin} which are shown in purple and red, respectively; domain III is shown in pink. **d** Topology diagram of a Cry11Aa dimer with similar color code as in (**c**) and with labeling of secondary structure elements in one of the two monomers. The two monomers in a dimer assemble via the β_{pin}, resulting in the formation of a large β-sheet. The short helices $\alpha9$, $\alpha10$ and $\alpha11$, respectively located in the $\beta11$-$\beta12$, $\beta15$-$\beta16$ and $\beta21$-$\beta22$ loops, are not shown in the diagram.

cascade and enable development of new, rational strategies for improved mosquito control.

Results

De novo phasing of Cry11Aa and Cry11Ba structures by SFX.

In vivo-grown crystals of Cry11Aa and Cry11Ba protoxins exhibit distinct morphologies, which initially concealed a surprising conservation of their crystal packing patterns. Cry11Aa crystallizes as hexagonal plates and Cry11Ba crystallizes as larger bipyramidal crystals (Fig. 2a-b) as reported earlier[4]. These morphological distinctions cannot be attributed to differences in crystallization mechanisms in their parent organisms, *Bti* and *Btj*, since both protoxins were recombinantly produced in the same host organism, an acrystalliferous strain of *Bti* (4Q7). Cry11Aa and Cry11Ba protoxins are expected to share structural resemblance to each other since the two sequences share 54% identity; however, 46% non-identity at the molecular level could easily produce large differences at the macroscopic level of crystal morphology. Moreover, the sequence of Cry11Ba is extended by 77-residues at its C-terminus, potentially also affecting differences in crystal packing (Supplementary Fig. 1). Interestingly, this extension has been identified as a low complexity region (LCR) by both CAST[19] and SEG[20] computational methods, which implicates the extension in the mechanism of crystal nucleation. At this point in our studies, the balance of evidence suggested that sequence divergence was likely to have erased the crystal packing pattern that early ancestors of today's Cry11Aa and Cry11Ba presumably once shared.

Our diffraction experiments yielded the first hint that Cry11Aa and Cry11Ba shared a conserved crystal packing pattern. We collected diffraction data from Cry11Aa and Cry11Ba nanocrystals injected in the vacuum chamber of the CXI-SC3 micro-focused beamline at the Stanford Linear Accelerator Center (SLAC) Linac Coherent Light Source (LCLS)[21] using a microfluidic electrokinetic sample holder (MESH)[22] (Cry11Ba crystals) or a gas-dynamic virtual nozzle (GDVN)[23] (Cry11Aa crystals). The underlying similarity in the packing of Cry11Aa and Cry11Ba became evident when their diffraction patterns were indexed, revealing similarly sized unit cells ($a \sim 58$; $b \sim 155$; $c \sim 171$ Å; $\alpha = \beta = \gamma = 90°$), albeit belonging to two different space groups: $I222$ and $P2_12_12$, respectively (Table 1). Conservation of unit cell parameters hinted that this crystal packing pattern is special, evolved to perform a function more intricate than just storing protein.

To gain further insight into Cry11Aa and Cry11Ba crystal packing, we depended on de novo methods to solve the crystallographic phase problem. Initial attempts to acquire phases from homologous structures by molecular replacement (MR) failed, suggesting Cry11Aa and Cry11Ba contained novel features, not present in the PDB. Our search models included structures of Cry δ-endotoxins homologs (exhibiting up to 26% sequence identity to our two targets) and homology models produced using Robetta[24] (http://robetta.bakerlab.org/) and SwissProt[25] (https://www.ebi.ac.uk/uniprot/). After MR failed, we turned to de novo phasing methods. We soaked Cry11 nanocrystals with conventional heavy atom derivatives including gadolinium (Gd),

Table 1 Data collection and refinement statistics of Cry11Aa and Cry11Ba.

	Cry11Aa pH 7	**Cry11Aa-TBXO4 pH7**	**Cry11Ba pH 6.5**	**Cry11Ba pH 10.4**
PDB ID	**7QX4**		**7QYD**	**7R1E**
Data collection				
Space group	*I* 2 2 2	*I* 2 2 2	P 21 21 2	P 21 21 2
Cell dimensions (Å)	57.64 ± 0.19	57.64 ± 0.15	168.18 ± 0.19	167.50 ± 0.29
	155.69 ± 0.80	156.29 ± 0.73	158.45 ± 0.26	157.99 ± 0.47
	171.14 ± 0.54	170.75 ± 0.40	57.51 ± 0.08	57.43 ± 0.14
Wavelength (Å)	1.27	1.27	1.30	1.30
X-ray beam focus (μm)	5	5	1	1
No. collected frames	792623	558747	813133	990643
No. indexed frames	48652	77373	19708	15689
No. merged crystals	50613	88511	19708	15689
Resolution range (Å)	33.55–2.60	33.51–2.55	42.06–2.40	35.72–2.65
	(2.66–2.60)	(2.61–2.55)	(2.44–2.40)	(2.70–2.65)
No. observations	8253629 (365007)	14069217 (640046)	3541082 (51048)	3482475 (67984)
No. unique reflections	24198 (1583)	48634 (3297)	61141 (2980)	45243 (2204)
⟨I/σ (I)⟩	9.50 (1.16)	11.23 (1.62)	4.73 (0.90)	3.98 (1.02)
R_{split} (%)	10.73 (95.40)	7.97 (70.58)	14.5 (84.9)	22.4 (84.0)
$CC_{1/2}$ (%)	99.3 (37.7)	99.6 (67.5)	99.1 (35.9)	98.5 (15.2)
Completeness (%)	99.9 (100.0)	100.0 (100.0)	100.0 (100.0)	99.4 (100.0)
Multiplicity	341.09 (230.58)	289.29 (194.13)	57.9 (17.1)	77.0 (30.9)
Anomalous data				
Completeness (%)		100.0 (100.0)		
CCano		0.26 (0.00)		
CRDano		1.35 (1.01)		
Refinement				
Resolution range (Å)	33.55–2.60		42.10–2.40	35.72–2.65
No. reflections	24196		55060	45228
R_{work}/R_{free}[a]	17.2/24.1		18.7/23.1	23.7/24.7
No. atoms				
Protein	5080		10083	9900
Water	261		623	119
B-factors (Å^2)				
Main chain	50.47		49.0/48.7[b]	46.4/43.7
Side chain	51.44		53.3/53.1	47.0/44.8
Water	46.17		48.4	36.5
R.m.s.d.				
Bonds lengths (Å)	0.004		0.008	0.001
Bonds angles (°)	0.633		1.324	0.408

[a] R_{free} is calculated using 5 or 10% (Cry11Aa or Cry11Ba, respectively) of random reflections excluded from refinement.
[b] Average B-factor for chain A / chain B.

gold (Au), platinum (Pt), and mercury (Hg) salts, but they failed to produce interpretable isomorphous or anomalous difference Patterson peaks. Finally, a recently introduced caged-terbium compound[16,17], Tb-Xo4, produced a successful derivative of Cry11Aa (after a 30 h soak at 10 mM concentration), and phases were determined by the single-wavelength anomalous dispersion (SAD) method at 2.55 Å resolution (using anomalous signal up to 3.5 Å). Two Tb-Xo4 molecules were identified bound to the single Cry11Aa molecule in the asymmetric unit (isomorphous peaks at 23 and 9 σ, and anomalous peaks at 33 and 8.1 σ, respectively; Supplementary Fig. 2a). The success of Tb-Xo4 can be partly ascribed to the dramatically high anomalous dispersion signal (i.e., f' and f") of terbium, but likely also stems from stronger binding of Tb-Xo4 to the protein owing to presence of an organic cage; indeed, f' and f" of Gd and Tb are similar at the X-ray energy used for data collection (9 keV). Regardless, phases were of sufficient quality to reveal all Cry11Aa residues from N13 to the C-terminal K643.

The Cry11Ba structure was thereafter phased successfully by MR using the Cry11Aa structure as a search model, revealing *a posteriori* that the Gd, Pt and Au ions had successfully bound to the crystalline Cry11Ba, despite anomalous and isomorphous signals being too weak to enable phasing (Supplementary Fig. 2b-c and Methods section). Our MR-phase 2.4 Å resolution map reveals two Cry11Ba molecules in the asymmetric unit. All residues are visible except for the N-terminus (residues M1-I11), two loops (residues S332-C335, and G354-S359) and the C-terminal extension (residues S659-K724). The lack of order in this extension is not surprising given the low complexity of its sequence.

Cry11 domain organization is similar to δ-endotoxins, but exhibits some non-canonical features. Cry11Aa and Cry11Ba structures maintain the three-domain organization characteristic of Cry δ-endotoxins[13,26] (Fig. 2c and Supplementary Fig. 3). Domain I is involved in formation of a pore in the target membrane. Like in other Cry δ-endotoxins, it forms a seven-α-helix bundle; at the center of the bundle is α5 (residues 146–170), surrounded by the remaining six helices. Domain II is involved in the recognition of mosquito-specific receptors. It forms a β-prism composed of three-β-sheets, wherein the first two β-sheets (β4-β3-β2-β5 and β8-β7-β6-β9) each adopts a Greek-key topology while the third β-sheet is three-stranded (β1-β10-β11). Domain III is involved in oligomerization. It forms a β-sandwich of two antiparallel five-stranded β-sheets (viz. β20 – β19 – β22 – β17 – β12/β14 and β15 – β13/β16 – β23 – β18 – β21) forming a jelly-roll topology, whereby β12/β14 and β13/β16 are interrupted β-strands contributed by two non-consecutive shorts β-strands, which appose and intercalate one after the other onto β17 and between β15 and β23, respectively (Fig. 2d).

The closest homolog of known structure to Cry11 toxins is *Bt kurstaki* (*Btk*) Cry2Aa (PDBid: 1i5p), with a sequence identity of 26.6 and 23.6 % and main-chain r.m.s.d. of 3.7 and 4.0 Å, with respect to Cry11Aa and Cry11Ba, respectively (Fig. 1). As with Cry2Aa, the Cry11 toxins feature a long insert (27 residues in Cry2Aa; 21 residues in the Cry11 toxins) between strands β10 and β11, which together with domain-I β1, form the third β-sheet of the domain-II β-prism. This insert, which features a short α-helix (α_h) and a β-strand (β_h), folds like a handle, and is therefore referred to as the $\alpha_h\beta_h$-handle, throughout the manuscript (Fig. 2c, Supplementary Fig. 3). The $\alpha_h\beta_h$-handle fastens domain II onto domain III through direct (e.g., in Cry11Aa, D443(OD2)-R502(NH2); D443(O)-R502(NH1); L447(N)-S503(O)) and water mediated H-bonds (T446(OG1)/T448(O)-Wat869(O)-R502(N); T448(OG1)/V499(O)-Wat744(O)-D501(OD1); T448(N)/L447(N)-Wat774(O)-S503(OG)/(O)) (Supplementary Fig. 4), and enables the burying of domain-II α8 at an interface formed by $\alpha_h\beta_h$, α6-α7 (domain I), β10-β11 (domain II), β15 and the β13-β14 and β15-β16 loops (domain III), and the α9 helix connecting domain II and domain III (D469-K478 in Cry11Aa). The firm hold of α8 enables the three domains to be more tightly packed in Cry2Aa and Cry11 toxins than in other Cry toxins (e.g., *Bt tenebrionis* (*Btt*) Cry3Aa or *Btk* Cry1Ac). Additionally, strand β_h lays aside strand β4 thereby expanding – and consequently, stabilizing—the first β-sheet of domain II (β_h-β4-β3-β2-β5). Also, alike Cry2Aa, the Cry11 toxins feature a smaller β-prism due to deletions in the second constitutive β-sheet, namely between β7 and β8 (6 and 10 residues missing in Cry2Aa and Cry11 toxins, respectively), and between β9 and β10 (14 and 15 residues missing in Cry2Aa and Cry11 toxins, respectively; Supplementary Fig. 3).

The Cry11 toxin structures are distinguished by a 36 to 38 residue insertion that is observed between strands β4 and β5. The insertion appends a β-strand at the edge of the first β-sheet of domain II – hereafter referred to as the β_{pin} (Fig. 2c). The β_{pin} forms the center of a two-fold symmetric dimer interface with the β_{pin} of another toxin molecule. The interface features approximately twelve backbone hydrogen bonds, merging two β-sheets into a large, antiparallel, intermolecular β-sheet (β_h-β4-β3-β2-β_{pin} – β_{pin}-β2-β3-β4-β_h) which fastens chain A to C and B to D (interface #3, see below) and stabilizes the tetramer (Fig. 3b, e). We note that the buried surface area (BSA) at the tetramerization interface is 33% lower in Cry11Ba, pointing to higher flexibility; this hypothesis is supported by the absence of interpretable electron density for residues at the N-terminus (332–335) and C-terminus (354–359) of the β_{pin} in the Cry11Ba structure. Also noteworthy is that Cry11 toxins feature a conserved N/D-DDLGITT insertion between β21 and β22, and deletions (>3 residues) between α3 and α4 (−5 and −8 residues with respect to *Btk* Cry2Aa and *Btt* Cry3Aa), and β20 and β21 (−10 and −9 residues with respect to *Btk* Cry2Aa and *Btt* Cry3Aa). Altogether, these changes render Cry11 toxins uniquely large from the structural standpoint, with predicted radii of gyration of 27.5 and 26.7 Å for Cry11Aa and Cry11Ba, compared to 25.0 and 25.6 Å for *Btk* Cry2Aa and *Btt* Cry3Aa, respectively.

All domains engage in producing the in vivo crystal lattice. Examination of packing interfaces reveal that all three domains are involved in the formation and stabilization of Cry11Aa and Cry11Ba nanocrystals. The in vivo crystallization pathway can be best trailed from Cry11Aa crystals, which feature a single monomer per asymmetric unit and build on six packing interfaces burying a cumulated surface area (BSA) of 3515 Å^2, corresponding to 13.1 % of the total protein surface area. The main building block of Cry11Aa crystals consists of a tetramer with a total BSA of 9663 Å^2 and a predicted binding energy of −12.5 kcal.mol^{-1} at pH 7 by PISA[27] (Fig. 3a-b).

The tetramer comprises two principle dimers: dimer A-B, and dimer C-D. A two-fold symmetry axis (vertical in Fig. 3b) relates chain A to B and chain C to D. At the dimer interface domain II contacts domain II ($\alpha_h\beta_h$-handle residues P433-P457 and strand β4) and domain III contacts domain III (interface #1; Fig. 3b). Perpendicular to this axis, another two-fold symmetry axis relates dimer A-B to dimer C-D creating A-C and B-D interfaces. These interfaces involve β_{pin} strands in domains II as mentioned in the previous paragraph (interface #3; Fig. 3b). The tetramer is further stabilized by minor contacts between apices of domain II (interfaces AD or BC; interface #6). Crystals grow by packing such tetramers in a brick-wall fashion via face-to-back contacts between domains I (interface #2; Fig. 3c). Cry11Aa crystals are further cemented by two additional minor interfaces. The first involves the apex of the second β-sheet of domain II (interface #5) from monomers in each dimer of the tetramer (AD or BC). The second occurs between the α3-α4 loop of domain I in one tetramer and the apex of the second β-sheet of domain II in another tetramer (interface #4).

Cry11Ba crystals assemble from tetrameric building blocks analogous to those in Cry11Aa crystals, as judged by the similarity of their crystal packing patterns (Fig. 3d, e, f). However, the tetramer in Cry11Ba is not as readily identified as an autonomous unit by PISA as it was in Cry11Aa crystals. Our measurements of BSA in the crystal packing interfaces suggest an explanation. We assign crystal packing forces to two types: those that associate monomers into dimers and tetramers and those that assemble tetramers into a crystal. The BSA which associates monomers into a dimer (interface #1) is 38% lower than the homologous interface in Cry11Aa (1009 Å^2 Fig. 2e; vs. 1631 Å^2; Fig. 3g), explaining PISA's failure to identify the tetramer. However, the packing of tetramers into crystals is 53% higher in Cry11Ba than in Cry11Aa (1429 Å^2 at interface #2; Fig. 3f vs. 934 Å^2; Fig. 3g). Thus, the relative contributions of the two types of crystal packing interfaces differ between Cry11Aa and Cry11Ba, but the sum of the contributions is nearly the same. Cry11Aa exhibits only slightly more BSA per monomer (3515 Å^2) than Cry11Ba (3385 Å^2), corresponding to 13.1% of the total protein surface area of Cry11Aa and 12.6 % of Cry11Ba. The mechanism by which Cry11Ba evolved stronger tetramer-tetramer interfaces could have exploited the extra degrees of freedom afforded by having two molecules in the asymmetric unit rather than just one as does Cry11Aa. Moreover, the emphasis on lattice-forming associations (rather than associations within a tetramer unit) could explain the larger crystal size achieved by Cry11Ba. Regardless, the Cry11 toxins structures shows that each domain functions to assemble and stabilize in vivo-grown nanocrystals. These functions must have evolved alongside domain specific functions: pore formation (domain I), receptor-recognition and membrane-insertion (domain II), and oligomerization and stabilization of the toxic pore conformation (domain III)[26].

Drastic conformational changes drive crystal dissolution. We sought to characterize the conformational changes that ensue pH elevation, preceding dissolution of the crystals in the mosquito larvae gut[28]. As the crystals are naturally labile at pH 11, we aimed at collecting data from crystals soaked at slightly lower pH than 11, hypothesizing that early conformational changes would show but the crystal packing still hold. In the case of Cry11Aa crystals, diffraction quality was decreased

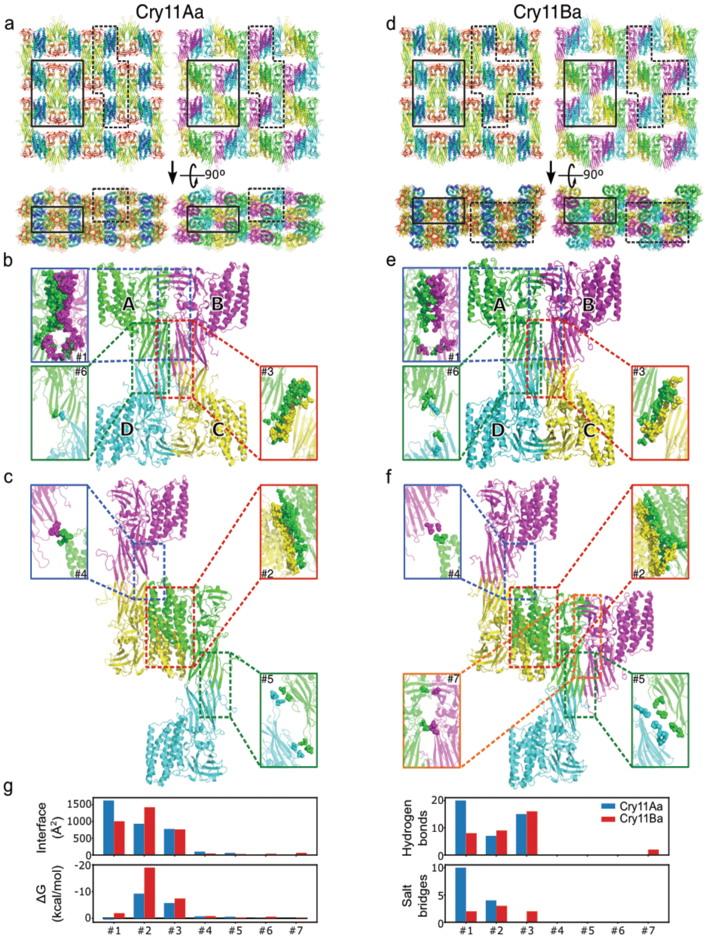


Fig. 3 Monomer interactions in Cry11Aa and Cry11Ba. a Cry11Aa crystal packing, colored according to sequence (from blue to red) indicating the domain-based assembly; and colored according to tetramer assembly (see panel (**b**)). The highlighted areas indicate the regions shown in (**b**) (full line) and (**c**) (dashed line). **b** Cry11Aa tetramer with zoom on each of the three interfaces identified by PISA (interface #1, #3 and #6), with the involved residues depicted as spheres. Supplementary Fig. 5 shows the residues involved in hydrogen bonds and salt bridges. **c** Cry11Aa crystal assembly by interactions between neighboring tetramers, formed by interface #2, #4 and #5, visualized as in (**b**). **d**, Cry11Ba crystal packing, colored as in (**a**). **e** Cry11Ba tetramer with zoom on the interfaces as in (**b**). **f** Cry11Ba crystal assembly, visualized as in (**c**). As compared to Cry11Aa, Cry11Ba crystals contain an additional interface #7 between an A-B pair from two neighboring tetramers. **g** Interface statistics as identified by PISA for Cry11Aa (blue) and Cry11Ba (red).

dramatically at pH values of 9.5 (CAPS buffer, glycerol 30%) and above, preventing collection of a sufficiently large number of diffraction patterns to produce a high-pH dataset. Hence, large conformational changes occur in Cry11Aa at pH as low as 9.5, opposing diffraction quality, despite crystals dissolving as of pH 11 only (Fig. 4a). In the case of Cry11Ba, 2.65 Å diffraction was preserved up to pH 10.4 (Table 1). Comparison between the refined 'pH10.4' and "pH6.5" structures points to large interdomain rearrangements induced by pH increase. Detailed analysis of structural changes at the side chain level was yet prevented by the non-isomorphism of the "pH6.5" and "pH10.4" datasets. A 1 % unit-cell contraction, and hence tighter crystal packing, was observed in the "pH10.4" crystals in comparison to the "pH6.5" crystals. However, because a higher glycerol concentration was used for injection of Cry11Ba crystals at pH 10.4, we cannot exclude that unit cell contraction might be caused by crystal dehydration.

Crystals are made of full-sized monomers of Cry11 protoxins. In both Cry11Aa and Cry11Ba toxins, the β_{pin} (residues E339-Q350 and I341-Y350, respectively) is a ~10-residue long β-strand that hydrogen-bonds with a twofold related symmetry mate, contributing the interface that assembles dimers (AC and BD) into tetramers. This strand is bordered on each side by the only two loops that have disordered electron density in Cry11Ba (missing residues S332-C335 and G354-S359) and are comparatively difficult to interpret in Cry11Aa (F330-D334 and Q350-E355), respectively. As Cry11Aa N335-Y349 and Cry11Ba I341-N351 regions match the enzymatic cleavage site known to generate the two activated fragments of ~32 and ~36 kDa[29,30] upon proteolytic activation in the mosquito larvae gut, we asked whether disorder in the F330-D334 (G330-E340) and Q350-E355 (D352-I358) loops serves the purpose of enabling facilitated access of proteases to Cry11Aa (Cry11Ba) cleavage sites or if each monomer occurs in natural crystals as two polypeptide chains

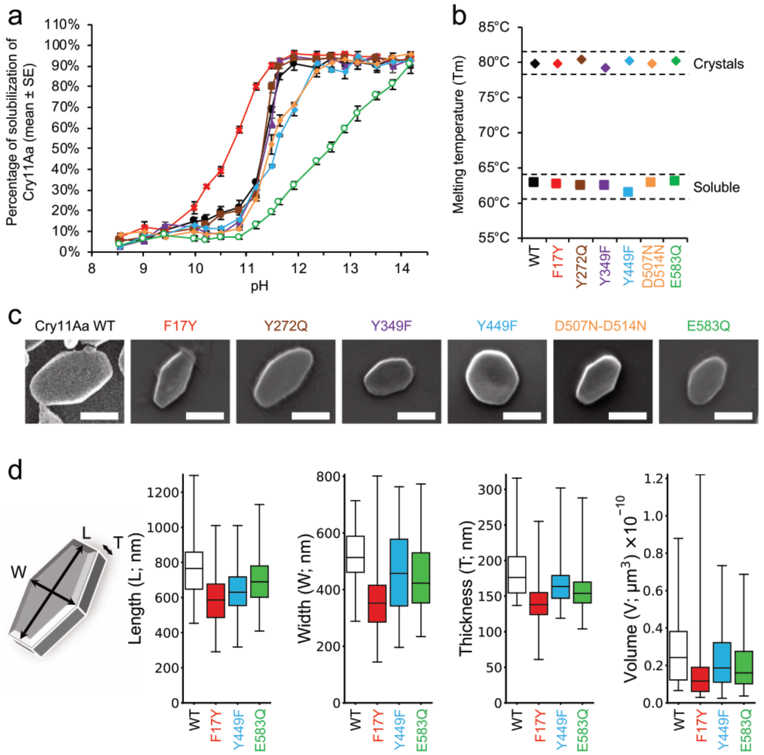


Fig. 4 Point-mutations of Cry11Aa affect the shape, size and pH-sensitivity of in vivo-grown nanocrystals. a Crystals from mutants exhibit similar sigmoidal patterns of crystal solubilization as a function of pH, except F17Y and E583Q that are more and less sensitive to pH, respectively ($n = 3$ independent measurements, data are presented as mean values ± SEM; Cry11Aa WT: black circles, F17Y: red crosses, Y272Q: brown squares, Y349F: purple triangles, Y449F: blue diamonds, D507N-D514N: orange plus symbols and E583Q: empty green circles). **b** Cry11Aa WT and mutants exhibit similar heat stability. As expected, toxins are more stable (+17.5 ± 0.3 °C) in their crystalline (diamonds) than soluble form (squares), irrespective of the mutation. The dashed lines are a guide to the eye. **c** Visualization of a representative crystal for Cry11Aa WT and mutants F17Y, Y272Q, Y349F, Y449F, D507N-D514N and E583Q by SEM (scale bar = 500 nm). **d** Crystals of F17Y ($n = 93$ crystals), Y449F ($n = 60$ crystals) and E583Q ($n = 94$ crystals) imaged by AFM were all smaller in length (L), width (W), thickness (T) and volume than WT ($n = 45$ crystals) highlighting a perturbation of the intrinsic crystal organization induced by these mutations. In each graph, the boxes represent the lower and upper quartiles around the median. The whiskers indicate the minimum and maximum values.

cleaved prior or during crystal formation. SDS-PAGE analysis of Cry11Aa (12% gels, heating at 95 °C for 5 min, presence of DTT and SDS; Supplementary Fig. 6) resulted in a major band ~70 kDa, in line with previous reports[31–33]. As the denaturing treatment would have broken any disulfide-bridge or non-covalent interactions that could maintain cleaved fragments together, this result suggests that Cry11Aa occurs in crystals as a full monomer. We further verified this hypothesis by use of MALDI TOF mass spectrometry. In MALDI mass spectra collected after direct solubilization of the natural crystals in sinapinic acid matrix in presence or absence of DTT, we observed main peaks at *m/z* of 72246 and 72235 (mass error: ± 100 Da) and 36154 and 36129 Da, respectively, in agreement with expected molecular masses for singly- and doubly- charged ions of a full-size monomer (expected mass: 72.349 kDa) (Uniprot accession number: P21256; Supplementary Fig. 7). However, because proteolytic activation is as well expected to yield a 36 kDa fragment, in addition to a 32 kDa fragment for which a minor peak was present in the MALDI-TOF mass spectra, we resorted to native mass spectrometry to assert that the ~72.240 and ~36.140 kDa peaks originated from the same species – rather than being indicative of the crystallization of proteolytic products. With this approach, we could confirm that upon dissolution of Cry11Aa crystals, a 72.345 kDa fragment is released, corresponding to the full-size monomer (Supplementary Fig. 8a). Moreover, both incubation of solubilized toxin at room temperature (RT) for 2 h (Supplementary Fig. 8b) and use of increased collision energy (Supplementary Fig. 8c-d) failed at yielding a signature for the two polypeptides that would have been generated if cleavage at position 329 had occurred. We conclude that natural crystals of Cry11Aa, and possibly Cry11Ba (Supplementary Figs. 6b and 7c), grow from the addition of full-size monomers, and that disorder in the F330-D334 (G330-E340) and Q350-E355 (D352-I358) loops could serve the purpose of enabling facilitated access of proteases to Cry11Aa (Cry11Ba) cleavage sites. Considering proteinase K as a surrogate analogue for mosquito larvae gut proteases[34], one would expect the β_{pin} to be released upon proteolytic activation, suggesting that the role of the latter is to promote in vivo crystallization, and that its cleavage would ensure irreversibility of crystal dissolution (Supplementary Fig. 3). We note that other cleavage sites are predicted, which would release the first six residues and last two β-strands (β22-β23), as well as rescind the covalent association between domain I and domains II and III, thereby leaving non-covalent interactions surfaces as the sole glue between them.

Mutagenesis to alter crystal formation and dissolution. We proposed earlier that the packing of Cry11Ba into slightly larger crystals than Cry11Aa could stem from differences in the extent and nature of the interfaces which support dimerization, tetramerization and packing of tetramers into crystals (Fig. 3). Considering recent evidence linking LCR regions with diverse functions including chaperoning[35] and reversible oligomerization, we further asked whether or not presence of the 77-residue LCR region at the C-terminus of Cry11Ba plays a complementary role in the promotion of crystal formation. A chimera was therefore designed, coined C11AB, wherein the LCR of Cry11Ba was fused to the C-terminal end of Cry11Aa (Methods section; Supplementary Fig. 9a). C11AB was produced at the expected size but at a lower yield than Cry11Aa WT (Supplementary Fig. 9b). Atomic force micrographs (AFM) revealed the presence of multiple needle-like inclusions in the parasporal envelope encasing the crystals, suggesting that presence of Cry11Ba-LCR at the C-terminal end of Cry11Aa favors nucleation, but not crystal growth (Supplementary Fig. 9c).

Seven Cry11Aa mutants and four Cry11Ba were additionally designed with the aim to probe the involvement of intra- and inter-molecular interfaces in toxin stability, crystal formation and dissolution. Each mutant was designed to challenge a specific interface and served as a coarse proxy to evaluate its pH sensitivity and putative participation in the crystal dissolution mechanism. First, we asked whether the intra-chain stabilization of α8 at an interface contributed by the three domains (namely, $\alpha_h\beta_h$, α6-α7, α9, β10, β11, β15 and the β13-β14 and β15-β16 loops) could play a role in crystal dissolution. Residues central to this interface in Cry11Aa are Y272, D514 and D507, which H-bond to one another and to Y203, R222, T249, S251 through direct and water-mediated interactions (W253 and W267), connecting the three domains (Supplementary Fig. 10a, Supplementary Table 1a). Upon pH elevation, Y272, D514 and D507 are all expected to be deprotonated, which should result in electrostatic repulsion and thence dissociation of the three domains. All these residues and their interactions are conserved in Cry11Ba (viz. Y273, D518, D511, Y203, R222, T249, S251, W253 and W268), suggesting that a similar mechanism could be at play. To test the hypothesis, we first produced three Cry11Aa mutants intended to eliminate pH sensitivity of the above-described H-bonds. Neither did the Y272Q nor D507N-D514N mutations impact the overall stability of the toxin, in the soluble or crystalline form (Fig. 4b), but their combination in the triple mutant Y272Q-D507N-D514N resulted in an unexpected abolishment of the ability of Cry11Aa to form crystals in vivo – possibly due to improper folding (Supplementary Fig. 11). The Y272Q mutation had no effect on the pH sensitivity of Cry11Aa crystals, while only a minor effect was seen with the D507N-D514N mutant (Fig. 4a). Thus, the increased stabilization of α8 at the interface between the three domains does not result in an increased tolerance of Cry11Aa crystals to pH elevation. We therefore probed the opposite question in Cry11Ba, and disrupted the hydrogen bond between Y273(OH) and D518(O) by engineering the Y273F mutation (Supplementary Fig. 12a). We found Cry11Ba-Y273F crystals to dissolve at a lower pH, indicating that destabilization of α8 at the interface between the three domains effects in increasing pH sensitivity (Supplementary Fig. 13). Thus, we could increase the pH sensitivity of Cry11Ba crystals by tampering with interactions between α8 and the three domains, suggesting that dissociation of domains is an important step in the pH induced dissolution of Cry11 crystals. However, decreasing the pH sensitivity of Cry11Aa crystals by stabilization of this region was not successful.

We then focused on Cry11Aa E583, a residue sitting at the intramolecular interface between domain I and domain III. This β21 residue, condemned to be anionic at higher pH, takes part in the water-mediated hydrogen bond network that connects α6 and α7 from domain I with domain III (Supplementary Fig. 10b, Supplementary Table 1b). We asked whether or not suppression of the pH sensitivity of the network would stabilize the monomer at high pH, thereby increasing the tolerance of crystals to pH elevation. This was indeed the case, with an SP_{50} (pH at which 50% of crystals are dissolved) of 12.6 ± 1.0 for crystals of the E583Q mutant, compared to 11.2 ± 1.0 for WT Cry11Aa crystals (Fig. 4a). The dissolution profile was also altered, showing a reduced slope and no visible plateau up to pH 14. Thus, the alteration of the protonation state of residues and water molecules at the intramolecular interface between domain I and domain III may be involved in the early step of Cry11Aa crystal dissolution. In Cry11Ba, which displays a similar SP_{50} of 11.9 ± 1.0 (Supplementary Fig. 13), this residue is substituted for glycine (G587) suggesting a different mechanism of pH-induced intramolecular separation of domain I and domain III – or at least the involvement of additional residues at the interface. We tested this hypothesis by engineering the Y241F mutation in Cry11Ba, effectively suppressing the H-bond between this residue, at the junction between domain I and domain II, and domain III residue D590 (Supplementary Fig. 12b). A reduced SP_{50} value was observed, confirming that the interface between domain I and domain III is central to the tuning of pH sensitivity in Cry11Ba. Considering that Y241 and D590 are strictly conserved in Cry11Aa (Y241 and D586, respectively), this conclusion could be valid for the two toxins.

Crystal contacts were also investigated. We first tampered with the interface enabling the brick-wall packing of Cry11Aa tetramers (Fig. 3c, interface #2) by introducing a F17Y substitution, intended to induce electrostatic repulsion with the negatively charged D180 (distance D180(OD1)—F17(CZ) of 3.3 Å) due to deprotonation of its hydroxyl group upon pH increase (Supplementary Fig. 10c). As expected, crystals of the F17Y mutant were found to be more sensitive to increases in pH, with crystals starting to dissolve at pH as low as ~9.5 and an SP_{50} of 10.6 ± 1.0 (Fig. 4a). The dissolution profile of F17Y crystals is again characterized by a reduced slope, as compared to WT crystals, explaining that the plateau is nonetheless reached at the same pH (~pH 11.6). Nevertheless, the result suggests that dissolution of Cry11Aa crystals can be accelerated by separation of the tetramers associated through interface #2. The F17Y mutation was also found to challenge crystal formation, yielding crystals far smaller than their WT counterparts. We note that F17, D180 and the H-bond between them are strictly conserved in Cry11Ba; hence, the importance of interface 2 for crystal formation and dissolution could be extendable to crystals formed by Cry11Ba.

Next, we challenged the role of the dimerization interface (Fig. 3b, interface #1) by mutating a residue positioned in the central part of the interface, viz. Y449 in Cry11Aa, corresponding to Y453 in Cry11Ba. Recall that the two toxins differ greatly at this interface contributed by domain III from facing monomers, with only 8 hydrogen bonds and 2 salt bridges to support the interface in Cry11Ba, compared to 20 hydrogen bonds and 10 salt bridges in Cry11Aa, and a 38% lower BSA in Cry11Ba than in Cry11Aa. Y449 is not involved in direct H-bonding to other protein residues but supports a large H-bond network that interconnects waters

Table 2 Data collection and refinement statistics of the Cry11Aa mutants.

PDB ID	Cry11Aa-F17Y pH 7 7QX7	Cry11Aa-Y449F pH 7 7QX5	Cry11Aa-E583Q pH 7 7QX6
Data collection			
Space group	*I* 2 2 2	*I* 2 2 2	*I* 2 2 2
Cell dimensions (Å)	57.72 ± 0.35 155.39 ± 1.49 171.66 ± 0.64	57.73 ± 0.24 155.55 ± 1.21 171.52 ± 0.57	57.76 ± 0.24 155.51 ± 0.98 171.51 ± 0.58
Wavelength (Å)	1.33	1.33	1.33
X-ray beam focus (μm)	1.3	1.3	1.3
No. collected frames	3150500	5993679	3523741
No. indexed frames	28227	104359	21833
No. merged crystals	28811	111014	22760
Resolution range (Å)	23.17–3.40 (3.48–3.40)	23.78–3.10 (3.17–3.10)	23.50–3.30 (3.38–3.30)
No. observations	2908715 (141787)	20279640 (1092683)	3210163 (154933)
No. unique reflections	10990 (707)	14447 (950)	12014 (787)
$\langle I/\sigma(I) \rangle$	6.31 (1.67)	9.95 (1.35)	5.64 (1.52)
R_{split} (%)	19.74 (76.86)	11.79 (89.56)	21.11 (80.18)
$CC_{1/2}$ (%)	95.9 (20.7)[a]	99.8 (59.3)	98.7 (30.7)
Completeness (%)	99.6 (100.0)	99.7 (100.0)	99.6 (100.0)
Multiplicity	265.7 (200.5)	1403.7 (1150.2)	267.2 (196.8)
Refinement			
Resolution range (Å)	23.17–3.40	23.18–3.10	23.08–3.30
No. reflections	10986	14442	12008
R_{work}/R_{free}[b]	21.2/25.1	22.4/25.2	21.5/25.4
No. atoms			
Protein	4970	4965	4970
Water	5	13	6
B-factors (Å^2)			
Main chain	54.6	43.1	45.4
Side chain	54.2	42.7	45.3
Water	52.9	59.3	36.0
R.m.s.d.			
Bonds lengths (Å)	0.002	0.002	0.003
Bonds angles (°)	0.448	0.441	0.489

[a] The low value for $CC_{1/2}$ for the outer resolution shell can be justified by the other favorable data quality statistics at this resolution.
[b] R_{free} is calculated using 5% of random reflections excluded from refinement.

and residues from facing monomers in the dimer (Supplementary Fig. 10d, Supplementary Table 1c). Furthermore, the two facing Y449 engage in edge-to-edge aromatic-dimer[36] interactions (centers of the rings are 5.5 Å apart, and the angle between them is 88°; Supplementary Fig. 10d). Contrastingly, the facing Y453 are 10.4 Å apart (center-to-center distance) in the Cry11Ba dimer, but the hydroxyl oxygen is H-bonded to T318(O) (Supplementary Fig. 12c). The Y449F mutation only had a minor effect on Cry11Aa crystal dissolution (Fig. 3a), indicating that deprotonation of its hydroxyl does not play a major role in the dissolution mechanism. Nonetheless, the mutation was detrimental to the protein stability (Fig. 3b) and resulted in the growth of crystals of different size and shape (Fig. 3c), likely owing to destabilization of the H-bond network at the dimer interface. The Y453F mutation, however, resulted in Cry11Ba crystals that dissolved at a lower pH (Supplementary Fig. 13), demonstrating the importance of this H-bond for the stability of the Cry11Ba dimer and the pH-sensing properties of the crystals.

Finally, we introduced a Y349F mutation in the β_{pin}, hypothesizing that suppression of its pH sensitive H-bond to E295(OE1) in the adjacent strand β2 would disturb the β_{pin} fold and destabilize the tetramerization interface (Fig. 3b, interface #3, Supplementary Fig. 10e, Supplementary Table 1d), thereby increasing sensitivity to pH elevation. This expected effect was not observed, with crystals of the mutant displaying the same pH-induced dissolution profile as those of the WT. Nonetheless, smaller crystals were observed that were characterized by a decreased thermal stability (Fig. 4 and Supplementary Fig. 14), indicating that reduced stabilization of the turn preceding the β_{pin} not only impacts folding and stability of the toxin, but as well its packing into crystals – probably due to reduced tetramerization. As Y349 is conserved in Cry11Ba, where it H-bonds to P362(O) (Supplementary Fig. 12d) due to the shorter side chain of D296 compared to Cry11Aa E295, we produced the analogous Y350F mutant and found that it solubilizes at a lower pH (Supplementary Fig. 13). Hence, in absence of a glutamic acid facing the tyrosine hydroxyl (E295), the expected effect on crystal dissolution is seen. This observation suggests that while Cry11Aa E295 and Y349 likely serve the same goal of inducing electrostatic repulsion upon deprotonation, pH sensing mostly depends on the that of E295.

Of all the single and double Cry11Aa mutants we investigated, the Y349F mutation is that which results in the smallest crystals, closely followed by F17Y and E583Q. The Y449F mutant, however, exhibits the most noticeable change in shape compared to WT Cry11Aa. To evaluate the significance of these changes, we characterized the distribution in size of crystals of Cry11Aa-WT, Y449F, F17Y and E583Q using AFM (Fig. 4d). All three mutants had a significantly reduced volume compared to WT Cry11Aa, due to a reduced thickness of the crystals.

Probing crystalline order of the Cry11Aa mutants by SFX. The presence of crystals does not necessarily infer that molecules are well ordered within them. We therefore used SFX to assess the level of crystalline order in crystals of the mutants that displayed modified solubilization or shape. Data were collected at the SPB/SFX beamline of the EuXFEL (Hamburg, Germany) from crystals delivered across the X-ray beam using a liquid microjet focused through a gas-dynamic virtual nozzle GDVN[23] (Table 2). All crystals were kept in water at pH 7 for the GDVN injection, and pulses were delivered at the MHz repetition rate (1.1 MHz)[37,38] using 10 Hz trains of 160 pulses, with a spacing of 880 ns apart. Data was collected on the AGIPD detector at its maximum rate of 3.52 kHz[39]. With the notable exception of Y349F, crystals of all four single point mutants diffracted, yet unequal amounts of data were collected from each, and none from WT crystals, due to technical difficulties that arose during the experiment. This impeded a thorough comparison of the diffraction power of the various mutants, and prevented structure determination for the Y272Q mutant. The structures of the other three mutants were determined, using the WT structure as a molecular replacement model for the phasing of diffraction data. We found that neither overall packing, tertiary structure nor interface formation is affected in the tested mutants at neutral pH (Supplementary Fig. 15). Of important note, these data demonstrate the feasibility of macromolecular nano-crystallography at MHz pulse rate using the brilliant micro-focused beam available at the SPB/SFX beamline of the EuXFEL.

The needle shape inclusions formed by C11AB were also investigated by SFX and found to present some crystalline order, as evidenced by diffraction rings up to ~6 Å resolution in the powder diagram calculated from the maximum projection of 395,656 hits (Supplementary Fig. 9d). It is clear, however, that a high-resolution structure is not readily practicable with these crystals, either because their small size makes them unsuitable for diffraction using a micro-focused XFEL beam or due to intrinsic disorder.

Discussion

We here report the structures of Cry11Aa and Cry11Ba, the two most potent Cry δ-endotoxins expressed by mosquitocidal *Bti* and *Btj*, respectively. Both toxins occur as natural nanocrystals that are produced during the sporulation phase of the bacteria, and dissolve upon elevation of pH in the mosquito larvae gut. Proteolytic activation enables binding to their specific receptors[40], including a membrane embedded alkaline phosphatase[41] but as well the co-delivered Cyt1Aa[12,42–44], triggering insertion in gut cell membranes and subsequent oligomerization into pores that will eventually kill the cells. Both toxins are of industrial interest due to their environmental safety, explained by the multi-step activation outlined above, and to their high stability as crystals. Our results shed light on the mechanisms of in vivo crystallization, pH-induced dissolution and proteolytic activation, and on structural features that support the toxins specificity with respect to other Cry toxins. Thereby, our work offers a foundation for further improvement of the toxic activity or crystal size by rational design. Additionally, we demonstrate the feasibility of de novo structure determination of a previously-unknown protein-structure by SFX, from nanocrystals featuring only ~75,000 unit-cells, using a single caged-terbium (Tb-Xo4) derivative. Below, we recapitulate these findings and discuss their implications.

In vivo crystallization pathway of Cry11 toxins. The building block of Cry11Aa and Cry11Ba crystals is a tetramer formed by the interaction of two dimers, via their domain II. The dimers are themselves assembled from the interaction of two monomers, via their domains II and III. Crystals form from the brick-wall packing of tetramers, as enabled by the face-to-back interaction of domain I from symmetry-related tetramers (Fig. 3). Thus, all three domains are involved in the in vivo crystal packing of Cry11 toxins, each contributing a twofold axis. This observation contrasts with other toxin structures determined from in vivo-grown crystals, wherein either propeptide(s) (e.g., *Lysinibacillus sphaericus* BinAB[28] and *Bti* Cyt1Aa[12]) or a specific domain (e.g., domain I in *Btt* Cry3Aa[45,46]) serves as the major contributor to crystallization. Expanding to previously determined Cry δ-endotoxins[5–7,13–15,46,47] structures, solved from in vitro-grown macrocrystals obtained following dissolution of the natural crystals at high pH, the same trend is observed—i.e., crystallization mostly depends on a dedicated portion of the protein, either it be a N-terminal and/or C-terminal propeptide (e.g., the ~650 C-terminal residues in *Btk* Cry1Ac) or a specific domain (e.g., domain II in *Btk* Cry2Aa). Thus, the Cry11Aa and Cry11Ba structures illustrate a yet unobserved pathway for in vivo crystallization, wherein all domains act on a specific step of the coalescence process, viz. dimerization (domains II and III from two Cry11 monomers), tetramerization (domain II from two Cry11 dimers) and tetramer packing (domain I in each tetramer). With Cry11Aa featuring a larger dimerization interface, and Cry11Ba a larger interface between piled tetramers, the two structures underline different levels of tradeoff between packing *into* tetramers and packing *of* the tetramers.

The difference in thickness of Cry11Aa and Cry11Ba crystals is of interest. Considering that all crystals were produced in *Bti*, we could exclude the possibility that the slightly larger size of Cry11Ba crystals originates from a more efficient crystallization machinery in *Btj* than in *Bti*. Puzzled by the presence of a 77-residue long low complexity region at the C-terminus of Cry11Ba (LCR-Cry11Ba), which is absent in Cry11Aa, we asked whether or not a C-terminal fusion of LCR-Cry11Ba with Cry11Aa would result in larger crystals. LCR regions have indeed been shown to support a variety of functions, including chaperoning[35] and reversible oligomerization[48,49] so that the role of the LCR-Cry11Ba could be found either in crystal nucleation or in crystal growth. Support of the first, but not the second hypothesis was obtained. Indeed, the C11AB chimera, consisting of a fusion of LCR-Cry11Ba to the C-terminus of Cry11Aa, yields smaller crystals that poorly diffract, even when exposed to high intensity XFEL pulses. This observation is in line with previous results which showed that substitution of Cry11Ba domain III for that of Cry11Aa leads to limited expression and comparatively small inclusions[50]. Thus, the LCR region of Cry11Ba is unlikely to account for the difference in size between Cry11Aa and Cry11Ba crystals. Instead, we favor the hypothesis that it is the larger surface of interaction between piled tetramers that accounts for the larger size of the Cry11Ba crystals. Given the absence of electron density for LCR-Cry11Ba residues in the Cry11Ba structure, and the abundance of needle-like inclusions in the parasporal body enveloping the C11AB crystals, it is reasonable to assume that they do not engage in structurally important interactions with functional domains, but rather favor nucleation of crystals. This aid-to-nucleation would be required for Cry11Ba, which features a reduced dimerization interface, but not for Cry11Aa, wherein this interface is 62 % larger. In line with this hypothesis, four regions are predicted to form short adhesive motifs of the Low Complexity, Amyloid-like Reversible Kinked Segments (LARKS) type (Supplementary Fig. 3).

Cry11 toxins depart from the canonical Cry δ-endotoxins architecture. The structures of Cry11Aa and Cry11Ba shed light on features that would not have been predicted based on sequence alignments (i.e., by homology modelling), and which largely deviate from the canonical organization observed in other Cry δ-endotoxins[5–7,13–15,28,45–47]. The most notable difference is the presence of a ~36 to 38 residue insertion between strands β4 and β5 in domain II, which results in an extra β-strand, coined β_{pin}. The β_{pin} not only participates in the formation of a modified β-prism, but contributes to a two-fold axis that supports tetramerization of Cry11 toxins through formation of two large β_h-β4-β3-β2-β_{pin}—β_{pin}-β2-β3-β4-β_h sheets between symmetry-related dimers into a tetramer. The observation of proteolytic cleavage sites at both the N- and C-termini of the β_{pin} suggests that it is removed upon activation by mosquito gut proteases, in line with the observation of ~32 and ~36 kDa fragments upon proteolytic activation of the Cry11 toxins[30–32]. If true, the unique role of the β_{pin} would be to support in vivo crystallization and its removal would entail the dissociation of tetramers into dimers and eventually monomers. While mutagenesis results indicate that this interface does not play a major role in crystal dissolution (see below), it seems likely that upon pH elevation and deprotonation of tyrosines and acidic groups, electrostatic repulsion will occur between Y349(OH) and E295(OE2) in Cry11Aa, and between Y350(OH) and P362(O) in Cry11Ba. Increased disorder of these regions could facilitate the access of proteases, and thus favor the activation of the Cry11Aa and Cry11Ba toxins. This hypothesis would rationalize the reluctance of the two toxins to recrystallize in vitro after pH induced dissolution, due to an impossibility to re-form tetramers – or at least, to re-match the exact positioning of the β_{pin}. The Cry11 toxins also differ from other Cry δ-endotoxins by the presence of a conserved N/D-DDLGITT insertion between β21 and β22, contributing a short helix, and by deletions of ~5–10 residues in the α3-α4 and β20-β21 loops, respectively. Compilation of these changes likely explains failures to phase the Cry11 structures by the molecular replacement method, even when *Btk* Cry2Aa, which also features a $\alpha_h\beta_h$-handle, was used as a starting model.

Mapping the interfaces involved in crystal dissolution. Our efforts to determine the structures of Cry11Aa and Cry11Ba at alkaline pH were unsuccessful, due to high sensitivity of crystals diffraction quality to pH increase. In the case of Cry11Aa we could not collect data, while in the case of Cry11Ba, we obtained a non-isomorphous structure which, while showing possible inter-domain rearrangements, did not inform on specific side chain rearrangements. Therefore, we resorted to site-specific mutagenesis to obtain information regarding the crystal dissolution pathway. We found that in Cry11Aa, the crystal interface most sensitive to pH elevation is the one enabling the honey-comb brick-wall packing of Cry11 tetramers, with the Cry11Aa-F17Y mutant displaying increased pH sensitivity (with an SP$_{50}$ of 10.6 ± 1.0 compared to 11.2 ± 1.0 for WT Cry11Aa crystals). In contrast, the dimerization (Y349F mutant) and tetramerization interfaces (Y449F mutant) appear to be less pH-sensitive. At the monomer level, we found that the three-domain interface to which α8 and the $\alpha_h\beta_h$-handle contribute is not very sensitive to pH increase (Y272Q and D507N-D514N mutants), possibly due to burying of mutated residues at the interface, preventing bulk solvent to access these sites and therefore retarding pH-sensing. Alternatively, interaction of Cry11 toxins with their membrane-bound receptors[40] could be a required step to expose α8, shown to play a major role in binding and toxicity[51]. Supporting this hypothesis is the observation that disruption of the H-bond between Y273(OH) and D518(O) in the Cry11Ba Y273F mutant increases pH sensitivity, demonstrating that Y273 plays a more important role in structuring the interface between α8 and the three domains than in triggering dissolution by electrostatic repulsion upon pH elevation.

The intramolecular domain I vs. domain III interface was found to be important for the pH-induced crystal dissolution, with the Cry11Aa E583Q and Cry11Ba Y241F mutants displaying reduced and increased sensitivity to pH (SP$_{50}$ of 12.6 ± 1.0 and 11.3 ± 1.0, respectively). Yet unlike the other tested interfaces, which are overall well conserved, the domain I vs. domain III interface differs in Cry11Aa and Cry11Ba, suggesting that caution is advised upon reflecting results obtained from Cry11Aa mutants onto Cry11Ba. Indeed, while Y241 and its H-bonding partner (D586 and D590 in Cry11Aa and Cry11Ba, respectively) are conserved in the two toxins (Supplementary Fig. 16), E583 is substituted for glycine in Cry11Ba (G587), suggesting that further mutagenesis experiments will be needed to further characterize the residues involved in the pH-induced separation of domain I and domain III. Amidst candidate residues to further tune the pH sensitivity is Cry11Ba E247, whose substitution for a glutamine would be expected to reduce electrostatic repulsion of V494 (β14) upon pH elevation. Likewise, E234 H-bonds to Q625(NE2; 2.6 Å) in Cry11Aa, and to K629(NZ; 2.8 Å) and R553(NH1; 2.8 Å) in Cry11Ba, suggesting that a E234Q mutation would reduce pH sensitivity in the two toxins whilst not affecting their folding. Inversely, the mutation into a glutamic acid of Q511/Q515, squeezed between a tryptophan (W584/W588), an arginine (R549/R553) and a glutamic acid (E234), would be expected to increase the pH sensitivity of the domain I vs. domain III intramolecular interface in both Cry11Aa and Cry11Ba – and by extension, that of their crystals.

The Cry11 structures afford rationalization of previous mutagenesis results. Prior to our work, the roles of select domains of Cry11 toxins (α3, α3-α4 loop, β1-α8 loop, α8, β3) have been investigated by mutagenesis. In light of the structures, the effect of single-point mutations can tentatively be clarified. Specific to Cry11Aa, one mutation was shown to fully abrogate crystal formation (V104E), whereas seven others were reported to reduce or suppress toxicity against *Aedes aegypti* mosquitoes (Supplementary Table 2)[31,51–54]. The α3 residue V104 is not present at a crystal interface, but rather exposed at the surface of domain I[31]. Thus, the V104E mutation is more likely to challenge crystal formation by affecting the folding of domain I. Given the immediate environment of this residue, we see two possible explanations; either electrostatic repulsion of E104 by α4 residues L129 and A132, resulting in destabilization of domain I; or the formation of a salt-bridge to R136, replacing the native salt-bridge to α3 E97 (Supplementary Fig. 17a). Supporting the latter hypothesis is the report that the E97A mutation also produces a non-toxic variant of Cry11Aa[31]. Further pushing forward the centrality of the R136-E97 salt bridge between α3 and α4 is the report that the R90E mutation also leads to a non-toxic Cry11Aa variant[31]. Indeed, replacement of R90 by a glutamic acid would force the Q139 side chain to flip, resulting in electrostatic repulsion of E97 and thereby disruption of the H-bonding pattern between α3 and α4 (Supplementary Fig. 17a). More difficult to understand is, however, how the mutation into an aspartic acid of V142, facing R90, results in a non-toxic Cry11Aa variant[54]. One would indeed expect that higher stabilization of the two helices by a direct salt bridge between R90 and D142 would result in a more potent toxin. Two other mutations introduced in α3 led to non-toxic Cry11Aa variants, namely Y98E and S105E[31]. The structure shows that Y98 fills a hydrophobic groove contributed by F63, P68, M71, L94, I101, I102, F140, L152 and M156 (Supplementary Fig. 17b); while S105(OH) H-bonds to main chain α3 atoms I101(O) and I102(O) and S105(Cβ) plugs another hydrophobic groove contributed by F55, L59, F108 and L129 (Supplementary Fig. 17c). It thus seems plausible that replacement of either residue by a glutamic acid would result in a destabilized domain I. The other Cry11Aa region that was targeted by mutations[51–53] is helix α8 and preceding residues in the β1-α8 loop, both of which are sandwiched between the β-handle in domain II and α6 and α7 in domain I. Briefly, full suppression of toxicity to *Aedes aegypti* mosquitoes was observed for V262E. Concerning the P261A, V262A and E266A mutations, results are contradictory in the literature, either mentioning a reduced toxicity or no difference with the wild-type Cry11Aa (Supplementary Table 2)[51–53]. If confirmed, the reduced activity of the P261A mutant would highlight the importance of the kink between α8 and the β1-α8 loop for the activity of Cry11Aa (Supplementary Fig. 17d). Interestingly, in Cry11Ba whose β1-α8 loop sequence differs, a proline residue is also found (P265) yet three residues downstream. The V262 residue fits in a hydrophobic groove contributed by C211, L215, I260, W267, V271, C432 and L438 (Supplementary Fig. 17d). Hence, it is plausible that introduction of a glutamic acid in this groove (V262E) would completely disrupt the interface between domain I and domain II, while replacement by the shorter side chain of an alanine (V262A) would loosen it. Last, the α8 residue E266 is exposed to the solvent and H-bonds to both the side chain and main chain nitrogen atoms of N263, in the β1-α8 loop. Mutation of E266 into an alanine will result in disruption of these H-bonds, and therefore in a destabilization of the interface between α8 and the β1-α8 loop (Supplementary Fig. 17e).

In Cry11Ba, most mutations have been made in triplet, making it difficult to pinpoint the contribution of each residue to the observed phenotypes[55]. Nonetheless, five single-point mutations have been reported, which led to Cry11Ba variants with reduced or suppressed toxicity against *Aedes aegypti*, *Anopheles stephensis* and *Culex quinquefasciatus* mosquitoes. Echoing the work described above, three of these have been introduced in α8 and the β1-α8 loop. Residue I263 is structurally equivalent to V262 in Cry11Aa and likewise, its side chain fits into a groove contributed by L211, L215, I260, W268, L272, C436 and L442 at the interface between domain I and domain II (Supplementary Fig. 18a). Hence, we propose that the I263A mutant exerts its detrimental effect on Cry11Ba toxicity in the same fashion than the V262A mutation in Cry11Aa. Likewise, the exposed side chains of S264 and K269 contribute to structuring α8 and the β1-α8 loop through interaction with D267(N) and P265(O), respectively (Supplementary Fig. 18b). Hence it is possible that the S264A and K269A mutations exert their effects in the same fashion as the E266A mutation in Cry11Aa, i.e., a destabilization of the region encompassing α8 and the β1-α8 loop. Since the effect of the three above mutations on the three mosquito species is not the same (Supplementary Table 2), it may be proposed that α8 and the β1-α8 loop are part of the binding epitope on the midgut microvillar receptor. The two other Cry11Ba residues that have been challenged by single point mutations are G257 and I306, both mutated into alanine (G257A, I306A). The former is involved in the conserved turn between β1 and the β1-α8 loop (the equivalent residue in Cry11Aa is also G257), hence the results highlight the necessity of a tight turn at the end of β1 to achieve full toxicity (Supplementary Fig. 18a). The latter is located in the middle of domain II β-prism, where its side chain fits into a hydrophobic grove contributed by F289, Y291, V309, W400, I402, L410 and I466 (Supplementary Fig. 18c). It is presumable that the I306 side chain replacement by the shorter side chain of an alanine effects in loosening the β-prism, which in turn would result in a reduced interaction with the receptor. Again, the effect of the two latter mutations varies depending on the considered mosquito species, suggesting an involvement in the binding to the receptors (Supplementary Table 2).

Implication for the future of nanocrystallography using SFX. In this study, de novo phasing was required—not because of the absence of homologous structures, but because none of those available were sufficiently close to serve as a search model for molecular replacement. Using Tb-Xo4, a caged terbium compound, we could phase the *Bti* Cry11Aa structure by SAD, from ~77,000 diffraction patterns collected on crystals consisting of ~75,000 unit cells on average – an achievement to compare to the determination of the *Ls* BinAB structure from >370,000 patterns (native and three derivatives) collected on crystals with ~100,000 unit cell[28]. Our success in phasing the Cry11Aa structure likely stemmed from a combination of the use of a dramatically powerful phasing agent[16,17] and advances in SFX data processing tools over the last five years, including the Xgandalf[56] indexing algorithm and improvements in data handling and integration in CrystFEL[57]. It should offer hope to investigators seeking to determine the structure of proteins of which no known structural homologue exists and that have to resort to SFX due to smallness of their crystals. It is foreseeable, however, that de novo structure determination will be helped by recent advances in comparative and ab initio modelling and the availability of programs such as RosettaFold[58] and AlphaFold2[59], capable of producing a decently-accurate structure for virtually all proteins and thus a good model for phasing of crystallographic data by molecular replacement. Latest releases of the two programs were published in the final stage of the writing of this manuscript, hence we asked whether or not the availability of these tools would have facilitated our journey towards the Cry11 toxins structures, and submitted the sequence of Cry11Aa to the two servers. For RosettaFold, the r.m.s.d. to the final refined structure of the five best models was over 4 Å, with discrepancies observed mostly in domain II. For AlphaFold2, however, the two first models displayed r.m.s.d. of 1.2 and 1.0 Å to the final structure, respectively. Using the worst of these two models, we could find a molecular replacement solution using Phaser, and a partial model featuring 95% of the residues in sequence was obtained after 20 cycles of automatic iterative model-building and refinement using Bucanneer[60] and Refmac5[61]. Thus, a problem which occupied five crystallographers for several years could have been solved in less than an hour using the new tools recently made available to the structural biology community. Based on our results, it is tantalizing to claim that the phase problem in crystallography has been solved, or that experimental structural biology will be abandoned, but such assertions would likely be shortsighted. Rather, we encourage investigators to challenge AlphaFold2 and RosettaFold as much as humanly possible, but to not forsake de novo phasing as it may remain the only route to success in difficult cases where molecular replacement based on such models does not work[62]. It must also be emphasized that in the case of Cry11 toxins and, more generally, naturally-crystalline proteins, the issue is not just phasing, but packing. For such proteins, crystal formation and dissolution serve function, hence characterization of packing interfaces is central to finely comprehend their bioactivation cascades. Without the naturally-occurring crystals and the atomic resolution experimental SFX data, it would not have been possible to make predictions on potential mutations affecting Cry11Aa crystal formation or dissolution.

Methods

Crystal production and purification. Crystals of Cry11Aa and Cry11Ba were produced by electrotransformation of the plasmids pWF53 and pPFT11S[63] into the acrystalliferous strain 4Q7 of *Bacillus thuringiensis* subsp. *israelensis* (*Bti*; The Bacillus Genetic Stock Center (BGSC), Columbus OH, USA), respectively[64]. Colonies were selected on LB agar medium supplemented with erythromycin (25 µg/mL) and used to inoculate precultures of LB liquid medium. For Cry11Aa production, precultures were spread on T3 sporulation medium. After incubation at 30 °C for 4 days, spores/crystals suspensions were collected using cell scrapers and resuspended in ultrapure water. After sonication-induced cell lysis and subsequent centrifugation at 4,000 g for 45 min to discard cell and medium debris, pellets were resuspended in water and crystals were purified by ultracentrifugation (23,000 × *g*, 4 °C, 16 h) on discontinuous sucrose gradient (67-72-79%). After ultracentrifugation, crystals were recovered and several rounds of centrifugation/resuspension in ultrapure water allowed discarding as much sucrose as possible for proper downstream application. Crystal purity was verified by SDS-PAGE on 12% gels. Purified crystals were conserved in ultrapure water at 4 °C until use. For Cry11Ba, a glycerol stock of the 4Q7/pPFT11S was streaked onto 25 µg/mL erythromycin Nutrient Agar plates. From here a single colony was selected and added to a Glucose-Yeast-Salts (GYS) media culture and allowed to grow continuously at 30 °C, 250 rpm for 5 days. This culture was then spun down, resuspended in ultrapure water, and the lysate was sonicated for 3 min at 50% duty. The sonicated lysate was added to the 30–65% discontinuous sucrose gradient (35-40-45-50-55-60-65 %) and spun down for 70 min at 20,000 rpm and 4 °C. The sucrose gradient was then hand fractionated with Cry11Ba crystals collected around 57–60% and dialyzed into ultrapure water. Crystal characterization and purity was completed by phase contrast light microscopy, X-ray powder diffraction, transmission electron microscopy, and 4–12% SDS-PAGE gels. The pure Cry11Ba crystals were stored at 4 °C in ultrapure water.

Cry11Aa mutagenesis. Based on the SFX structure of Cry11Aa, a total of 7 mutants of Cry11Aa were constructed to challenge the different crystal packing and intramolecular interfaces. The rationale behind these mutations is illustrated in Supplementary Fig. 10 and discussed in the main text. Point-mutations were inserted into *cry11aa* gene by Gibson assembly using pWF53 as a backbone[64]. Two different primer couples were used for each mutation to amplify two fragments that were complementary by their 15–18 bp overlapping 5′ and 3′ overhangs with a target Tm of 50 °C. Point mutations were inserted in the complementary part of the overhangs of the two fragments spanning the *cry11aa* region to be mutated. The double mutant D507N-D514N was successfully constructed in a single-step by respectively adding the D507N mutation on the non-overlapping overhang region of the forward primer, and the D514N on the non-overlapping overhang of the reverse one. The triple mutant Y272Q-D507N-D514N was constructed by using the primers containing the Y272Q mutation and the plasmid pWF53-D507N-D514N as a backbone. In addition to the point mutants, a Cry11Aa-Cry11Ba chimeric toxin—coined C11AB—was also constructed. For this, the sequence of the *cry11aa* gene was fused with the 234 bp extra 3′ extension of *cry11ba* gene, which is suggested to feature a low complexity region (LCR) based on sequence analysis using the LCR-eXXXplorer web platform (http://repeat.biol.ucy.ac.cy/fgb2/gbrowse/swissprot)[65], which implements the CAST[19] and SEG[20] computational methods to identify LCR. The C11AB chimera was constructed by Gibson assembly following a "1 vector, 2 fragments" approach. The plasmid pWF53 containing the *cry11aa* gene was used as a backbone and the *cry11ba* 3′ fragment was amplified from the extracted and purified plasmid of the WT strain of *Btj* containing the *cry11ba* gene. The list of primers used for plasmids construction is available in Supplementary Table 3. For each plasmid construction, the fragments with overlapping overhangs were assembled using the NEBuilder HiFi DNA Assembly (New England BioLabs) as previously described[12]. Briefly, after 90 min incubation at 50 °C, the constructed plasmids were transformed by heat shock into chemically competent Top10 *Escherichia coli* (New England BioLabs). Plasmids were extracted from colonies selected on LB agar medium containing ampicillin (100 µg/mL) using the NucleoSpin Plasmid extraction kit (Macherey-Nagel) following the manufacturer's instructions. The successful construction of each plasmid was assessed by double digestion (EcoRI and BamHI) followed by migration on 1% agarose gel stained with SYBR Safe (Invitrogen) and by Sanger sequencing of the region containing the mutation at the Eurofins Genomics sequencing platform. Of note, the *cry11aa* gene was also fully sequenced to validate its sequence for mutagenesis primer design and for comparing the expected toxin size to the observed one in mass spectrometry analyses. All mutants were produced as crystals in *Bti*, as described above. The presence of the mutated *cry11aa* gene sequence in the transformed *Bti* colony used for production was verified by colony PCR using specific primers and Sanger sequencing at the Eurofins Genomics sequencing platform. Crystals from all mutants were analyzed by SDS-PAGE on 12% gels. For C11AB, its proper size was confirmed by using the "gel analysis" module implemented in the software ImageJ v1.51k ($N = 7$)[66].

Cry11Ba mutagenesis. gBLOCK gene sequences composed of 2877 bp harboring the open reading frame of Cry11Ba, and targeted point mutations resulting in single amino acid replacements (Y241F, Y273F, Y350F, Y453F) in Cry11Ba, expressed under the cyt1Aa BtI, BtII, and BtIII promoters and featuring its transcription termination sequence[67], were synthesized at Integrated DNA Technologies (IDT, San Diego; Supplementary Table 4). These constructs were designed for cloning into the *E. coli - Bt* shuttle vector pHT3101[68], and contained homology sequences at the 5′ end (gaccatgattacgaatt) and 3′ end (gcatgcaagcttggc) for directional assembly in pHT3101 linearized with EcoR1 and Pst1 using the Choo-Choo Cloning Kit (Molecular Cloning Laboratories, MCLAB, San Francisco), according to manufacturer's protocol. Recombinant plasmids were propagated in E. coli DH5a, purified using the Wizard Plus Miniprep Kit (Promega), and point mutations were confirmed by sequencing (Genomic Core Facility, University of California Riverside) using internal forward and reverse primers that flanked the sites of interest (forward: DB11BseqF1 5′-GAATTTAGGAGGAAGCGATTGGGGA-3′ and DB11BseqF2 5′GTATTGTACTGAAAAGAATTTGGACGG-3′; reverse: DB11Seq11R 5′CTGGTGTATCTTCTAAGAATGATCTAT-3′). The acrystalliferous *B. thuringiensis* subsp. *israelensis* strain 4Q7 was transformed with the recombinant plasmids by electroporation as described previously[67]. Transformants were selected on Nutrient Agar supplemented with erythromycin (25 mg/ml) and the presence of crystals were initially monitored by phase contrast microscopy.

Crystal visualization by scanning electron microscopy (SEM). Purified crystals of Cry11Aa WT and of the 7 mutants were visualized using either a Zeiss LEO 1530 scanning electron microscope from the SEM facility of the European Synchrotron Radiation Facility (ESRF, Grenoble, France), a Thermo Fisher Quanta 650 FEG environmental SEM (ESEM) available for users at the European XFEL (EuXFEL, Hamburg, Germany) or a JEOL JSM-6700M FE-SEM (UCLA, Los Angeles, USA). For SEM at ESRF, samples were coated with a 2 nm thick gold layer with the Leica EM ACE600 sputter coater before imaging. For ESEM at the European XFEL, samples were diluted (1/1000) and mixed with 25 mM of ammonium acetate. Samples were then coated with a thin gold layer as described above using a Leica EM ACE600 sputter coater as well. Images were recorded at 15 kV acceleration voltage by collecting secondary electrons using an Everhart-Thornley-Detector (ETD detector) in high-vacuum mode. For SEM at UCLA, samples were diluted (1/5) in ultrapure H_2O. They were then added to 300 mesh Cu F/C grids that were positively glow discharged. These samples were then wicked away and washed with ultrapure water, wicked, and allowed to dry overnight to ensure all moisture had evaporated inside of a dessicator. These were then attached to a holder with carbon tape and coated with an Anatech Hummer VI sputter coater with approximately 2 nm of thick gold layer. Images were recorded at 5 kV acceleration voltage by collecting secondary electrons using a Lower secondary electron (LEI) or Upper secondary electron in-lens (SEI) detector.

Crystal visualization by transmission electron microscopy (TEM). Non-purified crystals of Cry11Aa WT were visualized using a Thermofisher TF20 electron microscope from the IBS electron microscopy platform. For negative staining TEM, samples were diluted 5 times in H_2O and 4 µL of the diluted sample was introduced to the interface of an amorphous carbon film evaporated on a mica sheet. The carbon film was then floated off the mica sheet in ~200 µL 2% sodium silicotungstate (SST) solution. The carbon film with the crystal sample was then recovered onto a Cu 300 mesh TEM grid after 30 s, let dry, and imaged at 200 keV. Images were recorded on a Gatan OneView CMOS detector. Non-purified crystals of Cry11Ba WT were visualized using a FEI Tecnai T12 electron microscope within the UCLA California Nanoscience Institute, EICN facility. For negative staining TEM, samples were prepared by adding 5 µL of pure crystal fractions in 10 µL ultrapure H_2O. In total, 2.5 µL of this sample was added to 300 mesh Cu F/C grids that were positively glow discharged. These samples were then wicked away using Whatman 1 filter paper; washed with 2.5 µL ultrapure H_2O, wicked; and negatively stained with 2.5 µL 2% uranyl acetate, wicked. These were allowed to dry overnight to ensure all moisture had evaporated and imaged at 120 keV. Images were recorded on a Gatan 2 K x 2 K CCD.

Crystal characterization by atomic force microscopy (AFM). Crystals of Cry11Aa were visualized by AFM as previously described[12]. Briefly, 5 µL of crystals suspended in ultrapure water were deposited on freshly cleaved mica. After 30 min in a desiccation cabinet (Superdry cabinet, 4% relative humidity), crystals were imaged on a Multimode 8, Nanoscope V (Bruker) controlled by the NanoScope software (Bruker, Santa Barbara, CA). Imaging was done in the tapping mode (TAP) with a target amplitude of 500 mV (about 12 nm oscillation) and a variable setpoint around 70% amplitude attenuation. TESPA-V2 cantilevers ($k = 42\ \mathrm{Nm}^{-1}$, $Fq = 320$ kHz, nominal tip radius = 7 nm, Bruker probes, Camarillo, CA, USA) were used and images were collected at ~1 Hz rate, with 512- or 1024-pixel sampling. Images were processed with Gwyddion[69], and if needed stripe noise was removed using DeStripe[70]. Measurements were performed on Cry11Aa WT and on mutants selected on the basis of their aspect in eSEM images (Y449F) or their solubilization pattern (F17Y and E583Q). Size measurements were performed on AFM images using Gwyddion[69] in a semi-automated protocol. A classical height threshold was applied to each image to select as many individual crystals as possible. Sometimes, partially overlapping crystals were individualized using the manual edition of the mask of selected crystals by adding a separation line. Finally, a filter was applied to remove very small selections (artefacts) or crystals touching the edge of the image. Measures were obtained using the 'distribution of grains' feature in Gwyddion where the crystal thickness (T) is the returned mean value, the volume (V) is the Laplacian background basis volume, and the length and width are the major and minor semi-axes of equivalent ellipses, respectively. The total number of crystals measured are: 45 for WT, 93 for F17Y, 60 for Y449F, and 94 for E583Q.

Data collection history. The Cry11Aa/Cry11Ba structure determination project was initiated in 2015. Data were collected at five different occasions, in two XFEL facilities, namely at the Linac Coherent Light Source (LCLS), Stanford (USA) and EuXFEL, Hamburg (Germany). During our first LCLS-SC3 beamtime (cxi04616), we collected data from native Cry11Ba (2.3 Å resolution), and in our second beamtime (LO91), we collected data from native Cry11Aa (2.8 Å resolution). Nanocrystals grown by recombinant expression in the modified acrystalliferous 4Q7 strain of *Bti* were injected by a microfluidic electrokinetic sample holder (MESH) device[22] in the microfocus chamber of LCLS-SC3[21]. After data reduction using cctbx.xfel and dials (hit-finding through merging)[71–74], we attempted phasing of both datasets by molecular replacement (MR), using sequence-alignment based multi-model approaches implemented in Mr Bump (based on MR by Molrep[75]) as well as custom-scripts testing models produced by Rosetta[24] (using the Robetta server; http://robetta.bakerlab.org/) and SwissProt[25] (https://www.ebi.ac.uk/uniprot/) servers (based on MR by Phaser[76]). Failure to find a homologue of a sufficiently-close structure led us to attempt de novo phasing of the Cry11 nanocrystalline proteins.

Initially, we aimed at obtaining experimental phases for Cry11Ba, considering that its larger crystals would produce a stronger diffraction signal which in turn would facilitate phasing. Hence, we collected derivative data on Cry11Ba, from crystals soaked with Gd, Pt and Au salts (P127 experiment) before injection using a MESH device[22]. Unfortunately, the data did not allow phase determination, as indicated by very weak and absent peaks in the isomorphous and anomalous difference maps, respectively (Supplementary Fig. 2), due to low occupancy of the soaked metal ions. Hence, we shifted focus to Cry11Aa crystals soaked with a recently introduced caged-terbium compound, Tb-Xo4[16,17] (P125 experiment). Crystals were injected using a GDVN[23] liquid microjet in the microfocus chamber of LCLS-SC3[21]. Online data processing was performed using NanoPeakCell[77] and CASS[78]. Offline data processing with NanoPeakCell[77] (hit finding) and CrystFEL[57] (indexing and merging) revealed a strong anomalous signal that enabled determination of the substructure and phasing of the SFX data, using Crank2[79] and its dependencies in the CCP4 suite[80] (see below for more details). The Cry11Aa structure was thereafter used to phase the Cry11Ba datasets by molecular replacement. *A posteriori*, we discovered that two of the heavy atom compounds that we used for soaking actually did bind Cry11Ba (Supplementary Fig. 2b-c). Difference Fourier maps revealed 7–8 σ peaks indicating Pt bound near Met 19 and 200, and Gd bound near Asp83 and Asp427 (Supplementary Fig. 2b). Surprisingly, however, there were no peaks in the anomalous difference Fourier maps. We speculate that if we had achieved higher heavy-atom occupancy and/or higher multiplicity in our measurements, the anomalous signal would have been strong enough to detect and perhaps used for phasing. We note that an alternative strategy could have been to first obtain experimental phases (either by seleno-methionylation or soaking with heavy metals) from in vitro-grown macro-crystals obtained by isolation, dissolution, recrystallization and derivatization, which could have allowed phasing by molecular replacement. However, as we could not exclude that Cry11 toxins would undergo large structural changes upon pH-induced activation, which would have complicated molecular replacement, we opted for the current strategy.

We last attempted data collection on Cry11Aa and Cry11Ba crystals soaked at elevated pH and injected by a MESH device (P141 experiment). Only Cry11Ba crystals could sustain the pH jump and yielded usable data. From the comparative analysis of the Cry11Aa and Cry11Ba structures, we nonetheless designed mutations aimed at increasing or decreasing the resilience of crystals; these were introduced in the *cry11aa* gene, and crystals were produced by recombinant expression in *Bti*. From these, SFX data were collected at the MHz pulse rate, during experiment P2545 at the SPB/SFX beam line of EuXFEL where a GDVN was used to inject crystals. The data were also processed with NanoPeakCell[77] (hit finding) and CrystFEL (indexing and merging).

It might be asked whether or not differences in data quality, arising from the use of different injection methods, could have played a role in the success in phasing Cry11Aa data, but not Cry11Ba data. Indeed, the use of a GDVN device, compatible with injection of a colloidal suspension of crystals in pure water, enables constant background in the diffraction patterns. This is less straightforward to achieve using a MESH device as the smaller size of the jet (reducing sample consumption by 5–10 fold) results in decreased stability (requiring to reposition the jet more often) and in the necessity to add highly viscous solvents to the crystal slurry (to avoid freezing in the vacuum chamber). To conclude on this point, a systematic study would be needed, whereby datasets assembled from the same number of images collected with either type of injector would be compared.

Data collection and processing, and structure refinement. During the P125 beamtime at LCLS, where the SAD data used for the phasing of the Cry11Aa structure were collected, the X-ray beam was tuned to an energy of 9800 eV (i.e., a wavelength of 1.27 Å), a pulse duration of 50 fs, a repetition rate of 120 Hz, and a focal size of 5 µm. SAD data were collected from nanocrystals soaked for 30 h with Tb-Xo4 at 10 mM in water, prior to GDVN injection[23]. Of 558747 images collected using the 5 µm beam available at LCLS-SCC, 76687, 292, 217 and 177 were indexed (i.e., a total of 77373 images) using Xgandalf[56], Dirax[81], taketwo[82] and Mosflm[83], respectively, in CrystFEL v.0.8.0[57]. Post-refinement was not attempted, but images were scaled one to another using the 'unity' model in CrystFEL *partialator*, yielding a derivative dataset extending to 2.55 Å resolution. *A posteriori*, we found that simple Monte Carlo averaging using the 'second-pass' option in CrystFEL *process_hkl* would have yielded data of similar quality, most probably because of the large number of indexed images. A native dataset was also collected and processed in the same fashion yielding, from 792,623 collected patterns of which 48,652 were indexed, a dataset extending to 2.60 Å resolution. The substructure of the derivative dataset was easily determined by ShelxD (figure of merit (FOM): 0.22), prompting us to try automatic methods for structure determination. Using Crank2[79] and its dependencies (ShelxC, ShelxD, Solomon, Bucanneer, Refmac5, Parrot) in CCP4 Online[84], the FOM was 0.52 after density modification, and rose to 0.88 upon building of 613 residues. This first model was characterized by R_{work}/R_{free} of 27.7/32.1% and was further improved by automatic and manual model building in phenix.autobuild[85] and Coot[86] until 630 residues were correctly built. This model was then used to phase the native data. Final manual rebuilding (using Coot[86]) and refinement (using phenix.refine[87] and Refmac5[61]) afforded a native model characterized by R_{work}/R_{free} of 17.2/24.1 % and consisting of most of the 643 residues. Only the first 12 N-terminal residues are missing (Table 1).

Cry11Ba data were collected during the cxi04616 and P141 beamtimes at LCLS-CXI. At both occasions, the photon energy was 9503 eV (i.e., a wavelength of 1.30 Å), a pulse duration of 50 fs, a repetition rate of 120 Hz, and a focal size of 1 µm—i.e., a similar standard configuration (pulse length, repetition rate) to that used for Cry11Aa, notwithstanding the beam size and wavelength. Data were collected from crystals at pH 6.5 (30% glycerol in pure water; cxi04616) and pH 10.4 (30% glycerol in 100 mM CAPS buffer; P141), presented to the X-ray beam using a MESH injector[22]. Of 813133 images collected for the pH 6.5 dataset, 19708 were indexed and scaled, post-refined, and merged using cctbx.xfel[71–74] and PRIME[88], yielding a dataset extending to 2.4 Å resolution. The Cry11Aa structure was used as a starting model to phase the Cry11Ba pH 6.5 dataset by molecular replacement using Phaser[76] with initial R_{work}/R_{free} being 34.4/40.4 %. Manual model building (using Coot[86]) and refinement (using Refmac5[61], phenix.refine[87] and Buster[89]) afforded a model characterized by R_{work}/R_{free} of 18.7/23.1 % (Table 1).

We used the refined Cry11Ba pH 6.5 structure as the starting model for the Cry11Ba pH 10.4 structure. We used Refmac5[61] to perform rigid body refinement and then Refmac5[61], phenix.refine[87], and Buster[89] to perform individual atomic refinement at a resolution of 2.65 Å. We performed manual model building with Coot[86]. The R_{work}/R_{free} of the final model was 23.7/24.7 %. The structural changes between the pH 6.5 and pH 10.4 models were difficult to interpret. No major peaks were observed in the difference Fourier difference map obtained by subtracting the pH 10.4 structure factors from the pH 6.5 structure factors. Consistent with this observation, there were no significant local structural changes, only a small contraction in the separation between subdomains. This contraction is consistent with a 1% shrinkage of the unit cell volume at pH 10.4. We ascribe this shrinkage to crystal dehydration caused by the use of a higher glycerol concentration during injection of the pH 10.4 sample. The conformational changes arising from elevated glycerol confound our interpretation of pH-related structural changes. Hence, we do not further discuss it in our manuscript.

Diffraction data on the Cry11Aa mutants at pH 7.0 was acquired on the SPB/SFX beamline at EuXFEL during our P002545 beamtime allocation, using a GDVN injector and X-ray energy and focal size of 9300 eV (1.33 Å) and 1.3 µm (FWHM), respectively. Technical problems allowed us to collect only a limited number of diffraction pattern of the Cry11Aa-Y349F mutant. 3150500; 5993679 and 3523741 images were collected for the F17Y, Y449F and E583Q mutant, respectively, of which 28227, 104359 and 21833 could be processed using CrysFEL0.8.0[57] and MonteCarlo based scaling and merging. The three structures were solved using MR with Phaser[76], using the Cry11Aa WT structure as input model. The structures were refined with only two B-factors per residue and secondary structure restraints in Phenix.refine[87] and Coot[86], with final R_{work}/R_{free} values of 21.2/25.1 % for Cry11Aa-F17Y, 22.4/25.1 % for Cry11Aa-Y449F and 21.5/25.4 % for Cry11Aa-E583Q (Table 2).

Structure analysis. Figures were prepared using PyMOL v. 2.5[90] (Figs. 2, 3 and Supplementary Fig. 4, 5, 10, 14, 15, 16, 17) and aline (Supplementary Fig. 3)[91]. Radii of gyration were predicted using the PyMOL script rgyrate (https://pymolwiki.org/index.php/Radius_of_gyration). Interfaces were analyzed with PISA[27] and root mean square deviations (r.m.s.d.) among structures were calculated using PyMOL using the 'super' algorithm. Sequence based alignment—performed using EBI laglign and ClustalW[92]—was challenged by the large gaps between *Bti* Cry11Aa, *Btj* Cry11Ba, *Btk* Cry2Aa and *Btt* Cry3Aa, while structure-based alignment—performed using SSM[93]—was blurred by the varying size of secondary structure elements in the three domains of the various toxins. Hence, for Supplementary Fig. 1, 3, the alignment of *Bti* Cry11Aa, *Btj* Cry11Ba, *Btk* Cry2Aa and *Btt* Cry3Aa was performed using strap[94] which takes into account both sequence and structural information. Specifically, the online version of the program was used (http://www.bioinformatics.org/strap/)[95]. To generate the phylogenetic tree in Fig. 1, We used the CCP4 program lsqkab to compute all pairwise superpositions of the 15 delta-endotoxins, and the r.m.s.d. of aligned alpha carbons. We uploaded a 15 × 15 matrix of r.m.s.d. values to the T-REX phylogenetic tree plot server www.trex.uqam.ca[96].

Structure prediction using AlphaFold2 and RosettaFold. RosettaFold[58] predictions were obtained by submitting the sequence to the Rosetta structure-prediction server (https://robetta.bakerlab.org). AlphaFold2[59] predictions were obtained by use of the Collaboratory service from Google Research (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/AlphaFold2_advanced.ipynb). The mmseq2 method[97,98] was employed for the multiple-sequence alignment instead of the slower jackhmmer method[99,100] used in Ref. [59]. Structural alignments were performed using the *align* tool in PyMOL[90]. Molecular replacements trials were carried out with Phaser[76]. Using the best five RosettaFold models, all characterized by an overall r.m.s.d. to the final structure superior to 4 Å, no molecular replacement solution could be found, due to inaccurate prediction of domain II β_{pin} and α_h-β_h regions, resulting in clashes. The best AlphaFold2 model was yet successful at predicting the domain II structure, which enabled successful phasing by molecular replacement, yielding a model characterized by R_{free} and R_{work} values of 0.322 and 0.292, respectively. This model was further used as a starting model for automatic model building and refinement using the Buccaneer pipeline in CCP4, resulting in a model characterized by R_{free} and R_{work} values of 0.245 and 0.215, respectively, after only five automatic cycles of iterative model-building, refinement and density modification using Buccaneer[60] and Refmac5[61] in the CCP4 suite[80].

Crystal solubilization assays. The solubility of crystals of Cry11Aa WT and of the mutants F17Y, Y272Q, Y349F, Y449F, D507N-D514N and E583Q was measured at different pH values as previously described[12]. Briefly, crystal suspensions were centrifuged at 11,000 × *g* for 2 min and resuspended in 18 different buffers with pH ranging from 8.6 to 14.2. After 1 h incubation in each buffer, crystals were centrifuged and the supernatant was collected. The concentration of soluble toxin in the supernatant was quantified using a Nanodrop 2000 (Thermo Fisher Scientist) by measuring the OD at 280 nm and by using the molar extinction coefficient and toxin size (102,930 M^{-1} cm^{-1} and 72.349 kDa, respectively, as calculated with the ProtParam tool of ExPASy (https://www.expasy.org) using the Cry11Aa protein sequence available under accession number "P21256"). Solubility was measured in triplicate for each toxin (Cry11Aa WT and mutants) and each pH. Data are normalized and represented as percentage of solubilization by dividing the concentration measured at a given pH by the maximum measured concentration.

For Cry11Ba and its mutants, the crystal suspensions were centrifuged at 13,300 × *g* for 3 min and ultrapure H_2O was removed from crystals. They were then resuspended in one of 18 buffers ranging from pH 7 to 14. These crystals were incubated for 1 h, afterwards the solution was centrifuged at 13,300 × *g* and the supernatant was separated from the crystal pellet. The concentration of the supernatant was then quantified by a ThermoFisher Nanodrop One (Thermo) by measuring the OD at 280 nm and using the molar extinction coefficient and toxin size (114600 $M^{-1}.cm^{-1}$ and 81.344 kDa respectively) that were calculated with Expasy ProtParam using the Cry11Ba sequence available at Uniprot.org under accession number Q45730. Solubility was measured in triplicate for each toxin at each pH measured. Data are normalized and represented as percentage of solubilization by dividing the concentration measured at a given pH by the maximum measured concentration. For Cry11Ba WT, this was then further tested by conducting a turbidity assay by resuspending the crystal pellet in 150 µL ultrapure H_2O and placed in a 96-well plate to be read on a NEPHELOstar Plus (BMG Labtech) nephelometer. These counts were normalized by subtracting the background signal and conducted in triplicate.

Solubility of Cry11Aa WT, Cry11Ba WT and their mutants was compared by calculating SP_{50} (pH leading to solubilization of 50% of crystals) as previously described[12], by globally fitting the data using a logistic regression model for binomial distribution using a script modified from Ref. [101]. Differences in SP_{50} between mutants were considered significant when 95% confidence intervals (CI), calculated using a Pearson's chi square goodness-of-fit test, did not overlap[102], and was confirmed by the individual fits of each of the triplicate measurement (Supplementary Fig. 19). All statistics were conducted using the software R 3.5.2[103].

Proteomic characterization. For SDS-PAGE experiments, samples heated to 95 °C were migrated on 12 % SDS-PAGE gels (1 h, 140 V) after addition of Laemmli buffer devoid of DTT. After staining by overnight incubation in Instant*Blue* (Sigma Aldrich, France), gels were washed twice in ultrapure water and migration results were digitalized using a ChemiDoc XRS + imaging system controlled by Image Lab software version 6.0.0 (BioRad, France).

MALDI TOF mass spectrometry. MALDI TOF mass spectra on Cry11Aa were acquired on an Autoflex mass spectrometer (Bruker Daltonics, Bremen, Germany) operated in linear positive ion mode. External mass calibration of the instrument, for the *m/z* range of interest, was carried out using as calibrants the monomeric (66.4 kDa) and dimeric (132.8 kDa) ions of bovine serum albumin (reference 7030, Sigma Aldrich). Just before analysis, crystals of Cry11Aa were firstly dissolved in acetonitrile/water mixture (70:30, *v/v*). For samples under reducing condition, DTT was added at a final concentration of 10 mM. The obtained solutions were therefore directly mixed in variable ratios (1:5, 1:10, 1:20, *v/v*) with sinapinic acid matrix (20 mg/mL solution in water/acetonitrile/trifluoroacetic acid, 70:30:0.1, *v/v/v*, Sigma Aldrich) to obtain the best signal-to-noise ratio for MALDI mass spectra. 1 to 2 µL of these mixtures were then deposited on the target and allowed to air dry (at room temperature and pressure). Mass spectra were acquired in the 10 to 160 kDa *m/z* range and data processed with Flexanalysis software (v.3.0, Bruker Daltonics).

MALDI TOF mass spectra on Cry11Ba were collected at the USC Mass Spectrometry Core Facility, Los Angeles, CA, USA. Purified Cry11Ba protein was dissolved in water (~5 mg/mL) and heated at 70 °C for 10 min to facilitate dissolution. One microliter of protein solution was spotted on a 384 Big Anchor MALDI target and let dry at room temperature. Crystallized protein was washed on-target twice with ultrapure water, on top of which 0.5 µL of 2,6 dihydroxyacetophenone (DHAP) solution (30 mg/mL in 50% acetonitrile:0.1% formic acid) was spotted and let dry at room temperature. Crystallized sample was then analyzed using Bruker Rapiflex® MALDI-TOF MS equipped with a Smartbeam 3D, 10 kHz, 355 nm Nd:YAG laser. The laser parameters were optimized as follows: scan range = 26 µm; number of shots per sample = 1000; laser frequency = 5000 Hz. The mass spectrometer was calibrated for high-mass range using Protein A and Trypsinogen standards under Linear Mode. Data were analyzed using FlexAnalysis software and plotted using Graphpad Prism.

In-gel digestion and peptide mass fingerprinting of Cry11Aa using MALDI. Selected bands were in-gel digested with trypsin as previously described[104]. MALDI mass spectra of the tryptic peptides were recorded on an Autoflex mass spectrometer (Bruker Daltonics, Bremen, Germany) in the reflectron positive ion mode. Before analysis samples were desalted and concentrated on RP-C18 tips (Millipore) and eluted directly with 2 µL of α-cyano-4-hydroxy cinnamic acid matrix (10 mg/mL in water/acetonitrile/trifluoroacetic acid: 50/50/0.1, *v/v/v*) on the target.

In-gel digestion and peptide mass fingerprinting of Cry11Ba using GeLC-MS/MS. Gel Liquid Chromatography tandem mass spectrometry spectra collected on Cry11Ba were acquired on a ThermoFisher Q-Exactive Plus (UCLA Molecular Instrumentation Center, Los Angeles, CA, USA). Before analysis, the Cry11Ba crystals were diluted at a 1:5 dilution with ultrapure H_2O and 4x SDS Loading Buffer Dye. These samples were then boiled for 3 min at 98 °C and were loaded on a 4–12% Bis-Tris SDS-PAGE gel. Protein embedded in gel bands were extracted and digested with 200 ng trypsin at 37 °C overnight. The digested products were extracted from the gel bands in 50% acetonitrile/49.9% H_2O/ 0.1% trifluoroacetic acid (TFA) and desalted with C18 StageTips prior to analysis by tandem mass spectrometry. Peptides were injected on an EASY-Spray HPLC column (25 cm × 75 µm ID, PepMap RSLC C18, 2 µm, ThermoScientific). Tandem mass spectra were acquired in a data-dependent manner with a quadrupole orbitrap mass spectrometer (Q-Exactive Plus Thermo Fisher Scientific) interfaced to a nanoelectrospray ionization source. The raw MS/MS data were converted into MGF format by Thermo Proteome Discoverer (VER. 1.4, Thermo Scientific). The MGF files were then analyzed by a MASCOT sequence database search.

Native mass spectrometry. Crystals of Cry11Aa were centrifuged for 5 min at 5000 × *g* during the buffer wash and washed twice with ammonium acetate buffer (pH adjusted to 6.4 with acetic acid). Pelleted crystals were then dissolved in ammonium acetate buffer (pH adjusted to 11.5 using ammonium hydroxide). Gold-coated capillary emitters were prepared as previously described and used to load the protein sample[105]. The sample was analyzed on a Synapt G1 mass spectrometer (Waters Corporation). The instrument was tuned to preserve non-covalent interactions. Briefly, the capillary voltage was set to 1.60 kV, the sampling cone voltage was 20 V, the extraction cone voltage was 5 V, the source temperature was 80 °C, the trap transfer collision energy was 10 V, and the trap collision energy (CE) was set at 30 V. For MS/MS characterization, a particular charge state was isolated in the quadrupole and the complex was dissociated by application of 200 V of CE. The data collected were deconvoluted and analyzed using UniDec[106].

Heat stability and aggregation propensity. The thermal unfolding of Cry11Aa WT and mutants was measured by following changes as a function of temperature (15–95 °C) in tryptophan fluorescence leading to an increase of the F350/F330 ratio. Scattering was also monitored to address aggregation propensity of Cry11Aa WT and of the mutants F17Y, Y272Q, Y349F, Y449F, D507N-D514N and E583Q (Supplementary Fig. 14). All the measurements were performed on a Prometheus NT.48 (Nanotemper) following manufacturer's instructions.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support this study are available from the corresponding author upon reasonable request. Structures and structure factor amplitudes are available in the PDB databank under accession codes 7QX4 (Cry11Aa WT, pH 7.0), 7QX5 (Cry11Aa Y449F, pH 7.0), 7QX6 (Cry11Aa E583Q, pH 7.0), 7QX7 (Cry11Aa F17Y, pH 7.0), 7QYD (Cry11Ba WT, pH 6.5), 7R1E (Cry11Ba WT, pH 10.4). Raw image files are deposited in the Coherent X-ray Imaging Data Bank (CXIDB) under accession number 190 [https://doi.org/10.11577/1873154]. The source data for Fig. 4 and for Supplementary Figs. 6, 7, 8, 9, 12 and 13, as well as uncropped blot scans for Supplementary Figs. 6, 9 and 11 are provided in a combined Source Data file. Source data are provided with this paper.

Received: 8 February 2022; Accepted: 30 June 2022;

References

1. Lacey, L. A. Bacillus thuringiensis serovariety israelensis and Bacillus sphaericus for mosquito control. *J. Am. Mosq. Control Assoc.* **23**, 133–163 (2007).
2. Vachon, V., Laprade, R. & Schwartz, J. L. Current models of the mode of action of Bacillus thuringiensis insecticidal crystal proteins: a critical review. *J. Invertebr. Pathol.* **111**, 1–12 (2012).
3. Delecluse, A., Rosso, M. L. & Ragni, A. Cloning and expression of a novel toxin gene from Bacillus thuringiensis subsp. jegathesan encoding a highly mosquitocidal protein. *Appl. Env. Microbiol* **61**, 4230–4235 (1995).
4. Federici, B. A., Park, H.-W. & Sakano, Y. Insecticidal Protein Crystals of Bacillus thuringiensis. in *Inclusions in Prokaryotes* (ed. Shively, J. M.) 195–236 (Springer, 2006).
5. Boonserm, P., Mo, M., Angsuthanasombat, C. & Lescar, J. Structure of the functional form of the mosquito larvicidal Cry4Aa toxin from Bacillus thuringiensis at a 2.8-Angstrom resolution. *J. Bacteriol.* **188**, 3391–3401 (2006).
6. Boonserm, P., Davis, P., Ellar, D. J. & Li, J. Crystal structure of the mosquito-larvicidal toxin Cry4Ba and its biological implications. *J. Mol. Biol.* **348**, 363–382 (2005).
7. Cohen, S. et al. Cyt1Aa toxin: crystal structure reveals implications for its membrane-perforating function. *J. Mol. Biol.* **413**, 804–814 (2011).
8. Gutierrez, P., Alzate, O. & Orduz, S. A theoretical model of the tridimensional structure of Bacillus thuringiensis subsp. medellin Cry 11Bb toxin deduced by homology modelling. *Mem. Inst. Oswaldo Cruz* **96**, 357–364 (2001).
9. Chapman, H. N. et al. Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73–77 (2011).
10. Schlichting, I. Serial femtosecond crystallography: the first five years. *IUCrJ* **2**, 246–255 (2015).
11. Boutet, S., Fromme, P. & Hunter, M. S. *X-ray Free Electron Lasers - A Revolution in Structural Biology*. vol. 1 (Springer International Publishing, 2018).
12. Tetreau, G. et al. Serial femtosecond crystallography on in vivo-grown crystals drives elucidation of mosquitocidal Cyt1Aa bioactivation cascade. *Nat. Commun.* **11**, 1153 (2020).
13. Li, J. D., Carroll, J. & Ellar, D. J. Crystal structure of insecticidal delta-endotoxin from Bacillus thuringiensis at 2.5 A resolution. *Nature* **353**, 815–821 (1991).
14. Morse, R. J., Yamamoto, T. & Stroud, R. M. Structure of Cry2Aa Suggests an Unexpected Receptor Binding Epitope. *Structure* **9**, 409–417 (2001).
15. Grochulski, P. et al. Bacillus thuringiensis CrylA(a) Insecticidal Toxin: Crystal Structure and Channel Formation. *J. Mol. Biol.* **254**, 447–464 (1995).

16. Engilberge, S. et al. Crystallophore: a versatile lanthanide complex for protein crystallography combining nucleating effects, phasing properties, and luminescence. *Chem. Sci.* **8**, 5909–5917 (2017).
17. Engilberge, S. et al. Protein crystal structure determination with the crystallophore, a nucleating and phasing agent. *J. Appl. Crystallogr.* **52**, 722–731 (2019).
18. Xu, C., Wang, B. C., Yu, Z. & Sun, M. Structural insights into Bacillus thuringiensis Cry, Cyt and parasporin toxins. *Toxins* **6**, 2732–2770 (2014).
19. Promponas, V. J. et al. CAST: an iterative algorithm for the complexity analysis of sequence tracts. *Bioinformatics* **16**, 915–922 (2000).
20. Wootton, J. C. & Federhen, S. Statistics of local complexity in amino acid sequences and sequence databases. *Comput. Chem.* **17**, 149–163 (1993).
21. Liang, M. et al. The Coherent X-ray Imaging instrument at the Linac Coherent Light Source. *J. Synchrotron Radiat.* **22**, 514–519 (2015).
22. Sierra, R. G. et al. Concentric-flow electrokinetic injector enables serial crystallography of ribosome and photosystem II. *Nat. Methods* **13**, 59–62 (2016).
23. DePonte, D. P. et al. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *J. Phys. -Appl. Phys.* **41**, 7 (2008).
24. Leaver-Fay, A. et al. Rosetta3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. *Methods Enzymol.* **487**, 545–574 (2011).
25. Bairoch, A. & Apweiler, R. The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1999. *Nucleic Acids Res.* **27**, 49–54 (1999).
26. Pardo-Lopez, L., Soberon, M. & Bravo, A. Bacillus thuringiensis insecticidal three-domain Cry toxins: mode of action, insect resistance and consequences for crop protection. *FEMS Microbiol. Rev.* **37**, 3–22 (2013).
27. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
28. Colletier, J.-P. et al. De novo phasing with X-ray laser reveals mosquito larvicide BinAB structure. *Nature* **539**, 43–47 (2016).
29. Oestergaard, J., Ehlers, R. U., Martinez-Ramirez, A. C. & Real, M. D. Binding of Cyt1Aa and Cry11Aa toxins of Bacillus thuringiensis serovar israelensis to brush border membrane vesicles of Tipula paludosa (Diptera: Nematocera) and subsequent pore formation. *Appl. Environ. Microbiol.* **73**, 3623–3629 (2007).
30. Yamagiwa, M., Sakagawa, K. & Sakai, H. Functional Analysis of Two Processed Fragments of Bacillus thuringiensis Cry11A Toxin. *Biosci. Biotechnol. Biochem.* **68**, 523–528 (2014).
31. Munoz-Garay, C. et al. Oligomerization of Cry11Aa from Bacillus thuringiensis Has an Important Role in Toxicity against Aedes aegypti. *Appl. Environ. Microbiol.* **75**, 7548–7550 (2009).
32. Fernandez, L. E. et al. Cry11Aa toxin from Bacillus thuringiensis binds its receptor in Aedes aegypti mosquito larvae through loop alpha-8 of domain II. *Febs Lett.* **579**, 3508–3514 (2005).
33. Carmona, D. et al. Dominant Negative Phenotype of Bacillus thuringiensis Cry1Ab, Cry11Aa and Cry4Ba Mutants Suggest Hetero-Oligomer Formation among Different Cry Toxins. *PLoS ONE* **6**, e19952 (2011).
34. Du, J. P., Knowles, B. H., Li, J. & Ellar, D. J. Biochemical characterization of Bacillus thuringiensis cytolytic toxins in association with a phospholipid bilayer. *Biochem. J.* **338**, 185–193 (1999).
35. Ntountoumi, C. et al. Low complexity regions in the proteins of prokaryotes perform important functional roles and are highly conserved. *Nucleic Acids Res.* **47**, 9998–10009 (2019).
36. Lanzarotti, E., Defelipe, L. A., Marti, M. A. & Turjanski, A. G. Aromatic clusters in protein–protein and protein–drug complexes. *J. Cheminformatics* **12**, 30 (2020).
37. Wiedorn, M. O. et al. Megahertz serial crystallography. *Nat. Commun.* **9**, 4025 (2018).
38. Grünbein, M. L. et al. Megahertz data collection from protein microcrystals at an X-ray free-electron laser. *Nat. Commun.* **9**, 3487 (2018).
39. Henrich, B. et al. The adaptive gain integrating pixel detector AGIPD a detector for the European XFEL. *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.* **633**, S11–S14 (2011).
40. Likitvivatanavong, S. et al. Multiple Receptors as Targets of Cry Toxins in Mosquitoes. *J. Agric. Food Chem.* **59**, 2829–2838 (2011).
41. Fernandez, L. E., Aimanova, K. G., Gill, S. S., Bravo, A. & Soberon, M. A GPI-anchored alkaline phosphatase is a functional midgut receptor of Cry11Aa toxin in Aedes aegypti larvae. *Biochem. J.* **394**, 77–84 (2006).
42. Bourgouin, C., Delecluse, A. & Rapoport, G. Specificity and Synergism of Bacillus thuringiensis israelensis Toxins on Mosquito Larvae. in *Bacterial Protein Toxins* (ed. Fehrenbach) 225–226 (Gustav Fisher, 1988).
43. Lopez-Diaz, J. A., Emiliano Canton, P., Gill, S. S., Soberon, M. & Bravo, A. Oligomerization is a key step in Cyt1Aa membrane insertion and toxicity but not necessary to synergize Cry11Aa toxicity in Aedes aegypti larvae. *Environ. Microbiol.* **15**, 3030–3039 (2013).
44. Pérez, C. et al. Bacillus thuringiensis subsp. israelensis Cyt1Aa synergizes Cry11Aa toxin by functioning as a membrane-bound receptor. *Proc. Natl. Acad. Sci. USA* **102**, 18303–18308 (2005).
45. Sawaya, M. R. et al. Protein crystal structure obtained at 2.9 Å resolution from injecting bacterial cells into an X-ray free-electron laser beam. *Proc. Natl. Acad. Sci.* **111**, 12769–12774 (2014).
46. Galitsky, N. et al. Structure of the insecticidal bacterial delta-endotoxin Cry3Bb1 of Bacillus thuringiensis. *Acta Crystallogr. D. Biol. Crystallogr.* **57**, 1101–1109 (2001).
47. Adalat, R., Saleem, F., Crickmore, N., Naz, S. & Shakoori, A. R. In vivo crystallization of three-domain Cry Toxins. *Toxins Basel* **9**, (2017).
48. Hughes, M. P. et al. Atomic structures of low-complexity protein segments reveal kinked β sheets that assemble networks. *Science* **359**, 698–701 (2018).
49. Salinas, N., Colletier, J. P., Moshe, A. & Landau, M. Extreme amyloid polymorphism in Staphylococcus aureus virulent PSMα peptides. *Nat. Commun.* **9**, 3512 (2018).
50. Sun, Y. et al. Construction and characterization of the interdomain chimeras using Cry11Aa and Cry11Ba from Bacillus thuringiensis and identification of a possible novel toxic chimera. *Biotechnol. Lett.* **36**, 105–111 (2013).
51. Fernandez, L. E. et al. Cry11Aa toxin from Bacillus thuringiensis binds its receptor in Aedes aegypti mosquito larvae through loop alpha-8 of domain II. *Febs Lett.* **579**, 3508–3514 (2005).
52. Perez, C. et al. Bacillus thuringiensis subsp israelensis Cyt1Aa synergizes Cry11Aa toxin by functioning as a membrane-bound receptor. *Proc. Natl. Acad. Sci. USA* **102**, 18303–18308 (2005).
53. Perez, C. et al. Bacillus thuringiensis ssp israelensis Cyt1Aa enhances activity of Cry11Aa toxin by facilitating the formation of a pre-pore oligomeric structure. *Cell. Microbiol.* **9**, 2931–2937 (2007).
54. Carmona, D. et al. Dominant negative phenotype of bacillus thuringiensis Cry1Ab, Cry11Aa and Cry4Ba mutants suggest hetero-oligomer formation among different cry toxins. *PLoS ONE* **6**, e19952 (2011).
55. Likitvivatanavong, S., Aimanova, K. G. & Gill, S. S. Loop residues of the receptor binding domain of Bacillus thuringiensis Cry11Ba toxin are important for mosquitocidal activity. *FEBS Lett.* **583**, 2021–2030 (2009).
56. Gevorkov, Y. et al. XGANDALF - extended gradient descent algorithm for lattice finding. *Acta Crystallogr. Sect. Found. Adv.* **75**, 694–704 (2019).
57. White, T. A. Processing serial crystallography data with CrystFEL: a step-by-step guide. *Acta Crystallogr. Sect. Struct. Biol.* **75**, 219–233 (2019).
58. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
59. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
60. Cowtan, K. Completion of autobuilt protein models using a database of protein fragments. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 328–335 (2012).
61. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 355–367 (2011).
62. Banneville, A.-S. et al. Structural and functional characterization of DdrC, a novel DNA damage-induced nucleoid associated protein involved in DNA compaction, *Nucleic Acids Res.* (2022) https://doi.org/10.1093/nar/gkac563.
63. Park, H. W., Delecluse, A. & Federici, B. A. Construction and characterization of a recombinant Bacillus thuringiensis subsp israelensis strain that produces Cry11B. *J. Invertebr. Pathol.* **78**, 37–44 (2001).
64. Wu, D. & Federici, B. A. Improved production of the insecticidal CryIVD protein in Bacillus-thuringiensis using CryLA(c) promoters to express the gene for an associated 20-KDa protein. *Appl. Microbiol. Biotechnol.* **42**, 697–702 (1995).
65. Kirmitzoglou, I. & Promponas, V. J. LCR-eXXXplorer: a web platform to search, visualize and share data for low complexity regions in protein sequences. *Bioinformatics* **31**, 2208–2210 (2015).
66. Girish, V. & Vijayalakshmi, A. Affordable image analysis using NIH Image/ImageJ. *Indian J. Cancer* **41**, 47 (2004).
67. Sakano, Y., Park, H. W., Bideshi, D. K., Ge, B. & Federici, B. A. Contributions of 5'-UTR and 3'-UTR cis elements to Cyt1Aa synthesis in Bacillus thuringiensis subsp. israelensis. *J. Invertebr. Pathol.* **149**, 66–75 (2017).
68. Lereclus, D., Arantès, O., Chaufaux, J. & Lecadet, M. Transformation and expression of a cloned delta-endotoxin gene in Bacillus thuringiensis. *FEMS Microbiol Lett.* **51**, 211–217 (1989).
69. Nečas, D. & Klapetek, P. Gwyddion: an open-source software for SPM data analysis. *Cent. Eur. J. Phys.* **10**, 181–188 (2012).
70. Chen, S. W. & Pellequer, J.-L. DeStripe: frequency-based algorithm for removing stripe noises from AFM images. *BMC Struct. Biol.* **11**, 7 (2011).
71. Waterman, D. G. et al. Diffraction-geometry refinement in the DIALS framework. *Acta Crystallogr. Sect. Struct. Biol.* **72**, 558–575 (2016).
72. Sauter, N. K., Hattne, J., Grosse-Kunstleve, R. W. & Echols, N. New Python-based methods for data processing. *Acta Crystallogr. D. Biol. Crystallogr.* **69**, 1274–1282 (2013).

73. Hattne, J. et al. Accurate macromolecular structures using minimal measurements from X-ray free-electron lasers. *Nat. Methods* **11**, 545–548 (2014).
74. Winter, G. et al. DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 85–97 (2018).
75. Vagin, A. & Teplyakov, A. MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **30**, 1022–1025 (1997).
76. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
77. Coquelle, N. et al. Raster-scanning serial protein crystallography using micro- and nano-focused synchrotron beams. *Acta Crystallogr. D. Biol. Crystallogr.* **71**, 1184–1196 (2015).
78. Foucar, L. et al. CASS—CFEL-ASG software suite. *Comput. Phys. Commun.* **183**, 2207–2213 (2012).
79. Skubák, P. & Pannu, N. S. Automatic protein structure solution from weak X-ray data. *Nat. Commun.* **4**, 1–6 (2013).
80. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 235–242 (2011).
81. Duisenberg, A. J. M. Indexing in single-crystal diffractometry with an obstinate list of reflections. *J. Appl. Crystallogr.* **25**, 92–96 (1992).
82. Ginn, H. M. et al. ıt TakeTwo: an indexing algorithm suited to still images with known crystal parameters. *Acta Crystallogr. Sect. D.* **72**, 956–965 (2016).
83. Leslie, A. G. W. The integration of macromolecular diffraction data. *Acta Crystallogr. D. Biol. Crystallogr.* **62**, 48–57 (2006).
84. Krissinel, E., Uski, V., Lebedev, A., Winn, M. & Ballard, C. Distributed computing for macromolecular crystallography. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 143–151 (2018).
85. Terwilliger, T. C. et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D. Biol. Crystallogr.* **64**, 61–69 (2008).
86. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr Biol. Crystallogr* **60**, 2126–2132 (2004).
87. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 352–367 (2012).
88. Uervirojnangkoorn, M. et al. Enabling X-ray free electron laser crystallography for challenging biological systems from a limited number of crystals. *eLife* **4**, e05421 (2015).
89. Blanc, E. et al. Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2210–2221 (2004).
90. *The PyMOL Molecular Graphics System.* (Schrödinger, LLC.).
91. Bond, C. S. & Schüttelkopf, A. W. ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr Biol. Crystallogr* **65**, 510–512 (2009).
92. Madeira, F. et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).
93. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr Biol. Crystallogr* **60**, 2256–2268 (2004).
94. Gille, C. & Frömmel, C. STRAP: editor for STRuctural Alignments of Proteins. *Bioinformatics* **17**, 377–378 (2001).
95. Gille, C., Fähling, M., Weyand, B., Wieland, T. & Gille, A. Alignment-Annotator web server: rendering and annotating sequence alignments. *Nucleic Acids Res.* **42**, W3–W6 (2014).
96. Boc, A., Diallo, A. B. & Makarenkov, V. T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucleic Acids Res.* **40**, W573–W579 (2012).
97. Mirdita, M., Steinegger, M., Breitwieser, F., Söding, J. & Levy Karin, E. Fast and sensitive taxonomic assignment to metagenomic contigs. *Bioinformatics* **37**, 3029–3031 (2021).
98. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
99. Prakash, A., Jeffryes, M., Bateman, A. & Finn, R. D. The HMMER Web Server for Protein Sequence Similarity Search. *Curr. Protoc. Bioinforma.* **60**, 3.15.1–3.15.23 (2017).
100. Potter, S. C. et al. HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204 (2018).
101. Savi, M. K., Mangamana, E. T., Deguenon, J. M., Hounmenou, C. G. & Kakaï, R. G. Determination of Lethal Concentrations Using an R Software Function Integrating the Abbott Correction. *J. Agric. Sci. Technol. A* **7**, 25–30 (2017).
102. Wheeler, M. W., Park, R. M. & Bailer, A. J. Comparing median lethal concentration values using confidence interval overlap or ratio tests. *Env. Toxicol. Chem.* **25**, 1441–1444 (2006).
103. R. *R: a language and environment for statistical computing.* (R Foundation for Statistical Computing, 2011).
104. Rosenfeld, J., Capdevielle, J., Guillemot, J. C. & Ferrara, P. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal. Biochem.* **203**, 173–179 (1992).
105. Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nat. Protoc.* **8**, 639–651 (2013).
106. Marty, M. T. et al. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).

Acknowledgements

We thank Aline LeRoy for help during experiments with the Nanotemper apparatus and Dr. Guy Schoehn for his support. We thank Dr. Filipe Maia for deposition of the raw diffraction data on CXIDB. We thank Sarah Rudd for assistance with the Cry11Ba mutagenesis. IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). We thank the LCLS for beamtime allocation under proposals cxi0416, L091, P127, P125 and P141, and the EuXFEL for beamtime allocation under proposals P2156 and P2545. This work was supported by the Agence Nationale de la Recherche (grants ANR-17-CE11-0018-01 and ANR-2018-CE11-0005-02 to J.-P.C.), the CNRS (PEPS SASLELX grant to M.W.) and used the atomic force microscopy (AFM) platform at the IBS and the electron microscopy (EM) platform of the Grenoble Instruct-ERIC center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB). Platform access was supported by FRISBI (ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). The EM facility is supported by the Auvergne-Rhône-Alpes Region, the Fondation Recherche Medicale (FRM), the fonds FEDER and the GIS-Infrastructures en Biologie Sante et Agronomie (IBiSA). Use of the LCLS at the Stanford Linear Accelerator Center (SLAC) National Laboratory, is supported by the US Department of Energy, Office of Science, and Office of Basic Energy Sciences under contract no. DE-AC02-76SF00515. The CXI instrument was funded by the Linac Coherent Light Source Ultrafast Science Instruments project, itself funded by the DOE Office of Basic Energy Sciences. Parts of the sample injector used at LCLS for this research were funded by the National Institute of Health, P41GM103393, formerly P41RR001209. Data processing was supported by National Institutes of Health grant GM117126 to N.K.S. N.A.S.'s contributions reported in this publication were supported by the Ruth L. Kirschstein National Research Service Award of the National Institutes of Health under award number GM007185. B.F. acknowledges funding support from the Pacific Southwest Regional Center of Excellence for Vector-Borne Diseases funded by the U.S. Centers for Disease Control and Prevention (Cooperative Agreement 1U01CK000516).

Author contributions

G.T., A.-So.B., D.B., E.L., F.L., L.D., H.-W.P., B.F. designed and constructed and transformed WT and mutants plasmids; G.T., E.A.A., A.-So.B., N.A.S., D.B., N.Z., H.-W.P. and B.F. produced crystals in vivo; L.S., E.J.C. and M.M.A. performed MALDI-TOF mass spectrometry experiments; P.Q. and A.L. performed native MS/MS mass spectrometry experiments; G.T., E.A.A. and N.A.S. performed solubilization assays; E.A.A. performed heat stability assays; N.A.S., M. B., W.L.L. and I.G. conducted transmission electron microscopy imaging; G.T., E.A.A., A.-So.B., N.A.S., R.S. and I.S. performed crystal visualization by SEM; G.T., J.-M.T., D.F. and J.-L.P. performed crystal visualization and size measurements by AFM; G.T., J.-L.P. and E.D.Z. performed the statistical analysis of solubilization data and G.T. and J.-L.P. performed the statistical analyses of AFM data; A.-So.B., M.B., M.W. and J.-P.C. secured beamtime at the ESRF for crystal screening; M.R.S., N.S., J.R., B.F., and D.C. secured beamtime at the APS for crystal screening; M.R.S., A.-So.B., M.S.H., J.R., M.W., N.K.S., B.F., D.C., I.S., J.-P.C. secured beamtime at the LCLS for data collection; I.S. and J.P.C. secured beamtime at the EuXFEL for data collection; S.E., E.G., A.R., C.C., F.R. and O.M. synthesized Tb-Xo4; G.T. derivatized Cry11Aa crystals for injection at LCLS; G.T., M.R.S., E.A.A., A.-So.B. and N.A.S. prepared crystals for data collection at XFEL and synchrotrons; R.G.S. developed and operated the MESH-on-a-stick injector; R.L.S. and R.B.D. developed the GDVN injector; E.A.A., G.S., M.G., G.N.-K., M.K., G.S., M.S., R.L.S. and R.B.D. operated the GDVN injector; G.T., M.R.S., E.A.A., A.-So.B., N.A.S., A.S.B., M.G., G.N.-K., M.S.H., M.K., R.G.S., G.S., M.S., I.D.Y., A.G., A.B., S.B., T.M.R.S., J.R., R.L.S., R.B.D., M.W., N.K.S., D.C., I.S. and J.-P.C. performed serial data collection at the LCLS; G.T., E.A.A., A.-So.B., N.C., M.G., G.N.-K., M.S.H., M.K., R.G.S., G.S., A.G., M.H., L.F., J.B., R.B., R.W., T.S., H.K., R.L., A.M., T.R.M.B., R.L.S., R.B.D., I.S. and J.-P.C. performed serial data collection at the EuXFEL; E.D.Z., N.C., A.S.B., I.D.Y. and N.K.S. produced new processing tools or devices; E.D.Z., N.C., A.S.B., A.G., I.D.Y., N.K.S. and J.-P.C. performed serial data processing; E.D.Z. and J.P.C. phased the structural data; M.R.S., E.D.Z., N.A.S. and J.-P.C. performed atomic model building, refinement and structure interpretation; G.T., M.R.S., E.D.Z. and J.-P.C. prepared figures and tables and wrote the manuscript with input from E.A.A., A.-So.B., N.A.S., A.S.B., M.L.G., D.B., S.E., L.S., R.S., W.L.L., J.-L.P., A.L., R.L.S., D.C. and I.S; J.-P.C. designed and coordinated the project; E. A. A., A.-S. B. and N. A. S. contributed equally.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-31746-x.

Correspondence and requests for materials should be addressed to Jacques-Philippe Colletier.

Peer review information *Nature Communications* thanks Mike Hough, Azadeh Shahsavar and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

[1]Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale, 71 Avenue des martyrs, F-38000 Grenoble, France. [2]UCLA-DOE Institute for Genomics and Proteomics, Department of Biological Chemistry, University of California, Los Angeles, CA 90095-1570, USA. [3]Max-Planck-Institut für medizinische Forschung, Jahnstrasse 29, 69120 Heidelberg, Germany. [4]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA. [5]Large-Scale Structures Group, Institut Laue-Langevin, F-38000 Grenoble, France. [6]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [7]Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA. [8]European XFEL GmbH, Holzkoppel 4, 22869 Schenefeld, Germany. [9]Department of Chemistry, Texas A&M University, College Station, TX 77845, USA. [10]Department of Entomology and Institute for Integrative Genome Biology, University of California, Riverside, CA 92521, USA. [11]Department of Biological Sciences, California Baptist University, Riverside, CA 92504, USA. [12]European Synchrotron Radiation Facility (ESRF), BP 220, 38043 Grenoble, France. [13]Mass Spectrometry Core Facility, School of Pharmacy, University of Southern California, Los Angeles, CA 90089, USA. [14]Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, CA 90089, USA. [15]Univ. Grenoble Alpes, CNRS, LECA, F-38000 Grenoble, France. [16]Univ. Lyon, ENS de Lyon, CNRS UMR 5182, Université Claude Bernard Lyon 1, Laboratoire de Chimie, F-69342 Lyon, France. [17]Polyvalan SARL, 15 parvis René Descartes, 69342 Lyon, France. [18]These authors contributed equally: Guillaume Tetreau, Michael R. Sawaya, Elke De Zitter. ✉email: colletier@ibs.fr

Chapter 3

Mutational and Biochemical Analysis of Cry11Ba

3.1 Background and Significance

From previous experiments, the Cry11Ba structure at pH 6.5 was successfully solved *de novo* from *in vivo* produced crystalline inclusions by X-ray Free Electron Laser (XFEL) crystallography (Chapter 2).[1] Unlike its fellow crystalline (Cry) and cytolytic (Cyt) parasporins, the Cry11Aa and Cry11Ba have a unique crystal packing, especially with the longer loops within their second domains which mainly consist of beta-sheets, and an unstructured C-terminus. Even with these structural similarities between these two isoforms, Cry11Ba exhibits greater toxicity levels against the major mosquito vectors. With the structure, a variety of programs provided structural insights about Cry11Ba. This provided the opportunity to select sites for mutations where they would target the largest interfaces or intrafaces, since these are typically responsible for strong crystal packing. To better understand the crystal stability, toxicity, and their relationship to one another, Cry11Ba was probed with select point mutations at tyrosine to elucidate whether these factors are independent or dependent of one another. These were also compared and mapped to various mutations from different δ-endotoxins[2] to see which regions are responsible for increased/decreased toxicity and determine sequence conservation (**Video 3.1**).

3.2 Introduction

With the de novo structures of Cry11Aa and Cry11Ba solved, analyzing the structures for potential mutations became the focus to better characterize the poorly understood δ-endotoxins function. There are a variety of programs that can aid in determining the best sites for mutations. For this study, the target was to disrupt the stability of crystals since they dissolved in the alkaline gut of their target host. While there are large swathes of inter- and intra- faces, there are programs that can calculate what the surface area of each interface is and what residues contribute most to these areas. For this study, Cry11Ba sites of stability were determined utilizing the Protein Data Bank in Europe’s Proteins, Interfaces, Structures, and Assemblies (PDBePISA) tool, where a protein structure’s interfaces in their crystal environment are analyzed, calculated, and quaternary structure can be predicted for a protein. This allows for a plethora of parameters, not only the size of crystal interfaces, to be determined, including energetics maintaining quaternary structure, and what the crystal contacts are for the structure. This was especially useful for the Cry11Ba structures since these were *in vivo* produced. Since these are a natural product of sporulation, they are unaltered or modified to force crystallization unlike most proteins for canonical crystallization. From these programs, Cry11Ba WT’s conserved globular 3-domain structural homology was found within these *Bt* protoxins, but are affected by various crystal packing secondary structural moieties. A variety of programs were utilized to compare and visualize these mutations, including Dali, STACCATO, and ConSurf. Dali is a protein structure comparison server, which compares the proteins’ 3D structures within the PDB.[3] This was ideal utilized for this family of proteins, since they do not appear to have any sequence homology, but have exceptionally high structural homology. The Dali program can undergo three types of database searches: “heuristic PDB,” which compares one query structure against all of that in the PDB; “exhaustive PDB25,” comparing one query structure against a representative PDB subset; and “hierarchical AF-DB,” this compares one query structure against a species subset of AlphaFold. These structures can then be compared “pairwise,” comparing one query structure against those specified, or “all against all,” returning a structural similarity for a set of specified structures. Next, the STructural sequence Alignment, Correspondence and Conservation Analysis Tool, STOCCATO, that took the PDB sequences provided by the Dali analysis and is a multiple sequence alignment (MSA) tool that combines the use of three-dimensional structure alignment probabilities and standard amino acid substitution probabilities.[4] This improves typical alignments by computing more accurate multiple-sequence alignments, analyzing protein conformational changes, and computation of amino acid structure sequence conservation. After obtaining an accurate multiple sequence alignment from STOCCATO, the final step was to visualize this alignment to show conservation of sequence, since the δ-endotoxins do have little to no observable sequence conservation. For this process, ConSurf was utilized, which is a bioinformatics tool for estimating evolutionary conservation of amino/nucleic acid positions in a protein/DNA/RNA molecule.[5–7] By doing so, the conservation of the amino acid sequence will be colored for more (maroon) or less (deep teal) conservation[7] mapped to the surface of the protein of interest's structure. By pairing each of these techniques, the δ-endotoxins high sequence variation can begin to be understood and find small stretches amongst all of the structurally homologous toxins that may be responsible for toxicity and/or elucidate sequence toxicity or specificity.

3.3 Results and Discussion

From PDBePISA[8,9], I discovered that native (WT) Cry11Ba did conserve the more globular 3-domain structural homology found within the majority of the *Bt* protoxins.[10] I observed Cry11Ba's largest interface is between the ɑ-helical bundle domains (named Domain 1) of two different Cry11Ba chains, the next was Cry11Ba's β-sheet domain (named Domain 2), that had the strongest interactions due to the antiparallel secondary structure between two different Cry11Ba chains' β-strands (**Fig. 3.1**). From this analysis, there were additional locations interior to the Cry11Ba domains' largest interfaces (interface #1 and interface #4), where some residues were key to interacting with the surrounding residues to stabilize and reduce their flexibility (**Appendix B**). While there were multiple residues observed in both of the aforementioned categories for mutation candidates (**Table 3.1**), I finally selected Tyr residues (Y241, Y273, Y350, Y453) at the most "load-bearing" locations, ie. areas with the highest Buried Surface Area (BSA) (**Fig. 3.2**). Since Tyr has a pK_a = 10.1-10.8, these were the top candidates to be deprotonated at the pH within the target hosts, ie. larval mosquitoes in the *Aedes* genus, of the *Bti* pesticidal toxins, which have had luminal pHs measured between 10.5-11.5.[11] Due to this, my mutation selections and designs focused on Tyr to Phe, since this should replicate the deprotonation that would occur within the luminal midgut of the larval mosquitoes. By losing this hydrogen-bonding site on the phenyl ring, the stabilizing interactions at key inter- and intra-faces should disperse and result in a reduction in stability. Furthermore, this would allow the pH trigger to shift due to diminished buffering capacity and increase solubilization at lower alkaline solutions. Though this would simulate the effects of destabilizing Cry11Ba, albeit the missing proteases and other chaperone proteins also theorized to be involved in this process,[12,13] this would provide insightful information of which sites are the most potent in increasing solubility, destabilizing Cry11Ba, and observe its effects on toxicity.

The initial constructs were developed by Brian Federici, Dennis Bideshi, Hyun-Woo Park, and Sarah-Jane Rudd at University of California, Riverside and California Baptist University by utilizing primer-induced mutations at the selected point mutation sites and then changed the DNA to encode for that new residue for both the forward and reverse primers with overhangs to help improve incorporation into the plasmid. This is then grown in *E. coli* and from there the plasmid undergoes a transfer to *B. thuringiensis* with a shuttle vector, where the initial vector is designed to propagate in two different host species by having origins of replication for the two different organisms.[8] When this was first attempted, Dennis Bideshi, Hyun-Woo Park, and Sarah-Jane Rudd were able to successfully express all of the mutant constructs after manipulation in *E. coli* and then insertion back into *B. thuringiensis* via (*E.coli-Bt*) shuttle vector pHT3101.[14] Each of the mutant constructs were cultured in a similar manner as the WT Cry11Ba (Chapter 2).[1] The purification of the crystalline inclusions followed the same protocol as the WT Cry11Ba crystals (Chapter 2)[1], where a sucrose gradient of varied percentages, 30-65%, were layered by hand within thin-walled UltraClear centrifugal tubes and linearized overnight. Cry11Ba lysate was then layered on top of the linearized sucrose gradient to completely fill the tubes, since the centrifugal tubes rely upon pressure displacement within the tubes to prevent the tubes from collapsing inward during the high-speed centrifugation. These are carefully balanced, within +/-0.005 of one another, to prevent any imbalance that would increase the pressure resistance within the ultracentrifugal chamber. After centrifuging, each mutation had distinct bands at varying sucrose percentages. These include the three main components to be separated within the lysate: cell membranes, spores, and the Cry11Ba crystals. With the sucrose gradients, bands were consistent amongst their mutants, but greatly differed from one another and the native gradients (**Fig. 3.3**). These were then further characterized by different analysis techniques including: SDS-PAGE (**Fig. 3.4**), powder diffraction, confocal microscopy, phase-contrast microscopy, and electron microscopy to determine purity and crystal morphology. While the initial morphological differences were less observable due to all the mutants retaining the native cubic bipyramidal shape, they did display a broad range of smaller sizes compared to native Cry11Ba crystals (**Fig. 3.3**). With these smaller crystals, 0.7 µm, observed more often than larger, 1.5 µm, which is a result of the mutants being selected to reduce stability and disrupt the crystal lattice contacts/interactions.

Utilizing pre-calculated XFEL data, the Y350F Cry11Ba at pH 6.5 structure was solved. For the mutant structure, the data quality was reduced when compared to the WT Cry11Ba at pH

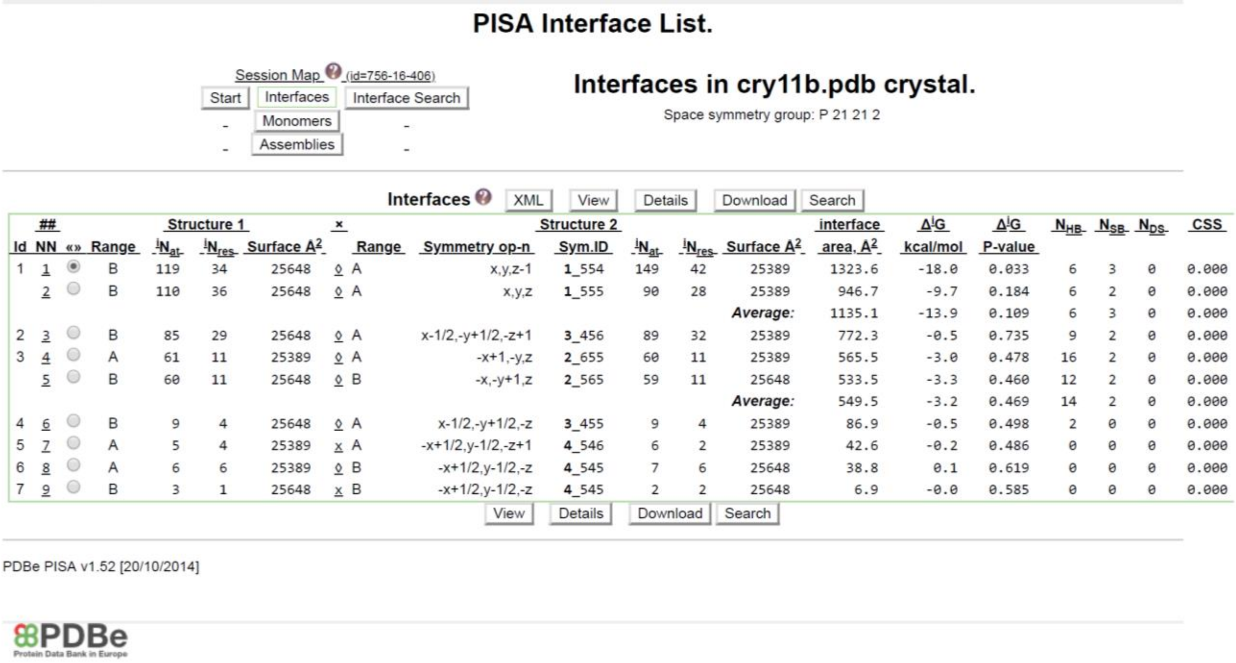
6.5 (**Chapter 2, Table 1&2**).[1] Upon solving the structure and utilizing WT Cry11Ba as a molecular replacement model, I went to mutate the Tyr350 in the 3D model structure to Phe and noticed positive density when the Phe was mutated (**Fig. 3.5a,b**). To remove any bias, the structure was reprocessed and solved with a Phe at position 350 from the beginning of refinement; however, the positive density around the para-position of the Phe aromatic ring was still present and was only satisfied once the Phe was mutated to Tyr with the hydroxyl group occupying the density. Upon discovering this during the Y350F Cry11Ba structure refinement, we further probed what the status of this mutant was, since the PCR results had confirmed that the mutations were up taken by the plasmid and shuttle vector. One possibility was that the mixing injector on the XFEL for serial femtosecond crystallography had residual WT Cry11Ba crystals present and this then contributed to its signal collected in the data. This was avoided though by the XFEL facilitators thoroughly rinsing the mixing injector before introducing new samples, which despite the preventative efforts may still been present. To confirm that the *Bti* cells crystalline inclusions were in fact producing solely mutant crystalline inclusions, I utilized Gel Extraction Liquid Chromatography Tandem Mass Spectrometry (GeLC-MS/MS) to determine the final purified Cry11Ba crystal mutation state from both the stock and diluted sample for the XFEL data collection. From the preliminary results, I observed within the Y350F Cry11Ba purified fraction, both WT and Y350F peptide fragments within the spectra (**Fig. 3.6**). To further probe and confirm this, I utilized vaporization iodination labeling (VIL), which will selectively iodinate the meta-position on the aromatic ring when $^{-}$OH is present and confirmed the status of the other mutants as well, where each mutant showed the +142 amu shift within that peptide fragment (**Fig. 3.7**), consistent with an iodination event on the alleged Phe or confirmed Tyr aromatic ring (**Fig. 3.5c,d**). From these studies, I was able to confirm that each of the mutants were contaminated with the WT Cry11Ba species. Due to how mass spectrometry functions, obtaining an accurate ratio of WT:Y350F was not possible to understand the magnitude of the mutant signal causing density within the structure or being masked during other biochemical assays. This was thus similar for all the mutants. From preliminary trials of the solubility assays, I observed an effect of uptake into solution at a lower pH than the WT for each mutant, but these were not as extreme as I believed the mutants should be. Due to the GeLC-MS/MS, we discovered that each of the purified mutant crystals were contaminated with WT Cry11Ba from different preparations utilizing the initial mutant colonies provided. The previous solubility data was still promising as shifts toward less alkaline pHs were observed when conducting solubility assays and the purifications displayed different sucrose gradient fraction fingerprints than the WT. The expected results for Y241F, Y273F, Y350F, and Y453F were to be further enhanced without the WT crystal's contributing signal masking/competing for the mutant.

After ruling out the possibility of the contamination of WT Cry11Ba crystals from the injector, we further investigated how to mitigate and eliminate these recurring in the next expression. While researching different mutation/expression approaches to prevent wildtype crystals from being expressed, Dennis Bideshi, Hyun-Woo Park, Sarah-Jane Rudd, and I decided to slightly alter the mutagenesis step in the construct's creation. The expression or shuttle vector used as these vary in how they are expressed and their efficiency. For example, an integrated shuttle vector is more effective for maintaining low spontaneous mutation frequency within the system, but takes time to establish a workable cell line and low DNA copy numbers. While potentially caused by the shuttle vector system being utilized to mutate the Cry11B expression vector, it is not common to have a "leaky" vector of the WT or never been reported previously to be a common occurrence. From this discovery, new mutations were made in GeneBlocks (gBlocks), which are gene fragments of double-stranded DNA between 125-3000 bps), and a unique gBlock was designed for each mutant. These were then shuttle-vectored back into *Bacillus thuringiensis* by Dennis Bideshi, Hyun-Woo Park, and Sarah-Jane Rudd and sequenced via PCR at the DNA level and confirmed the mutants' sequences and were then cultured, expressed, and purified for protein sequencing via GeLC-MS/MS. When I reconducted the GeLC-MS/MS, I was able to confirm the mutant constructs contained no WT contamination utilizing the same protocol as before displaying the contamination (**Fig. 3.8**). Upon confirming these with peptide sequence coverage containing each of the mutants respectively and an average of 75% sequence coverage for each mutant construct. Characterization profiles for each mutant were collected utilizing powder diffractions, OD600, phase contrast microscopy, electron microscopy, and purification extraction percentages. Additionally, these purified mutant crystals were used in solubility assays with pHs ranging 7 - 13.5. The concentration of the supernatant was measured by absorbance and turbidity was measured of the remaining pellet. I proceeded to reconduct the solubility assays and observed a more drastic shift, ~0.7 pH units, to a lower pH for each of the point mutations of Cry11Ba. Since these Cry and Cyt paratoxins display low sequence homology, yet are highly effective toxins, this was further analyzed. While working on these new expressions, I conducted further analysis utilizing Dali, STOCCATO, and ConSurf to study and visualize the conservation across 15 Cry and Cyt proteins that displayed high structural homology (**Figs. 3.9-11 & Video 3.2**). Amongst these only 1 residue, Tyr195, was completely conserved and will be added as a mutation to the panel of constructs with the expectation of higher effects on toxicity than stability, since it is not at a key inter- or intra- face and may be contained within a regulatory region for toxicity due to the complete/high conservation (**Fig. 3.11**).

From the solubility studies, a distinct shift of pH sensitivity from the previously conducted assays with pH 11.5 - 11.6 for the initial contaminated mutants to pH ~11.3 for the new mutants. This distinct shift was statistically significant from Cry11Ba wt (**Table 2a**) and is indicative that the mutations selected at those key intra- and inter- faces to target these large BSA sites to increase flexibility and/or reduce stability became more pH sensitive to their alkaline environment causing them to dissolve at a lower pH. From the assays, I can state that each of these sites is important for protein stability in the crystal state due to removing these hydrogen-bonds resulting in a decrease in stability and increase in pH sensitivity from pH 11.6 to 11.3 (**Fig. 3.12; Appendix A, Supp Fig. 13**). A future study designing a construct with multiple mutations to see if these effects are compounding and exponentially increase pH sensitivity and decrease stability. Another observation was at more neutral pHs, the solubility was increased for a few of the mutants and uptake was occurring at pHs that had minimal uptake for the WT Cry11Ba. This would be indicative that not only were these mutations more pH sensitivity, but that the crystal integrity/stability was affected in a significant enough way to allow for this uptake. According to a two-factor ANOVA, we were able to determine that these values for Y350F and Y241F were statistically significant with 90% and 95% confidence respectively (**Table 3.2b**) for the neutral or slightly alkaline pH (pH 7-9.5) values making multiple mutant constructs even more interesting for future studies. Given these results, making a new chimera combining interface or a mixture of intra- and inter-face mutations would be of interest to see if these effects are compounding and exponential in increasing pH sensitivity and decrease stability.

Lastly, the Cry11Ba mutants were evaluated for their biotoxicity to determine if the effects observed for pH sensitivity translated to toxin functionality. Each of the purified mutant crystals were lyophilized in nanopure H_2O and then evaluated for biotoxicity against *Culex quinquefasciatus*, the southern house mosquito. This was the first take instead of *Aedes aegypti*, the true target of Cry11Ba, as it is one of the most pervasive species worldwide, except the northern region of the temperate zone, and a host of various pathogens, including West Nile Virus. *C. quinquefasciatus* is similar to *A. aegypti* in their resting positions on water, but *A. aegypti* are carriers of fewer, albeit more deadly, pathogens, including Zika and Dengue viruses, and able to exist in every region of the world. Having a better understanding as to how these toxins function will aid in designing more effective synthetic toxins and help combat the growing resistance to pesticides within different vector species[2] (**Video 3.2**). The other important aspect is the age of the vectors being experimented on, as most of the pesticidal protein target larvae, since this is the life stage of most vectors that will be in the soil and will ingest these pesticidal proteins. During this larval stage, there are typically 4-6 levels of instars, which is the developmental stage of an arthropod between molts, before adulthood which varies depending on species. For *Culicidae*, where mosquitoes are categorized under *Diptera*, have 4 instars before entering the pupae stage and finally molt into their well-known adult mosquito form. The biotoxicity assays were conducted by Dennis Bideshi and Sarah-Jane Rudd, who exposed 4th instar *C. quinquefasciatus* larvae to different concentrations of the lyophilized crystals for 48 hrs where mortality was recorded. From this experiment's data, Hyun-Woo Park analyzed and calculated the LC_{50} and LC_{95} values (**Table 3.3**), except Y273F, were less toxic than the native Cry11Ba produced. The Y273F mutant is a targeted Tyr at an intraface between Domains 1 and 2 within the same chain indicates not only pH sensitivity, but also involvement in toxicity to the target. This mutation is buried within the protein and is a short helix with higher hydrophobicity than its surrounding hydrophilic residues, thus by removing the hydroxyl group from the Tyr, not only was the pH sensitivity increased, the change in hydropathy in this region could have allowed for more flexibility and switched to a more active state for the toxin. While the target vector species of Cry11Ba is not *C. quinquefasciatus*, this is indicative of the pervasiveness of these toxins in fine tuning their systems to their target vectors, despite being in the same Familiae Culicidae. While the other mutants displayed decreased toxicity, ie. higher LC_{95} values, this is still indicative that the Y241F, Y350F, and Y453F are each involved in the activity of the toxin, but not how I initially perceived they would. Displaying a change in pH sensitivity indicates that the stability effects do not necessarily correlate to higher toxicity but could still affect the mode of action of the toxin. Each of these less toxic mutations were affecting interface interactions and stabilizing the large BSAs observed via the PISA analysis, thus these residues would be more solvent exposed as the crystals interacted with the alkaline pH and able to interact with the receptors within the gut of the mosquitos. Without the hydroxyl group present, the Cry11Ba mutants' ability to shift from the crystalline toxin to protoxin may have been inhibited or lost binding effectiveness to the receptors to cause the pore-forming function of the δ-endotoxins. To further probe these effects, chimeras of the Cry11Ba mutants should be designed to see if the pH sensitivity and toxicity effects can be synergistic and mutate to more polar or charged residues to target the mode of action's biochemistry.

3.4 Figures



PISA Interface List.

Interfaces in cry11b.pdb crystal.

Space symmetry group: P 21 21 2

Interfaces

##			Structure 1				×			Structure 2				interface	$\Delta^{i}G$	$\Delta^{i}G$	N_{HB}	N_{SB}	N_{DS}	CSS
Id	NN	«»	Range	$^{i}N_{at}$	$^{i}N_{res}$	Surface Å^2		Range	Symmetry op-n	Sym.ID	$^{i}N_{at}$	$^{i}N_{res}$	Surface Å^2	area, Å^2	kcal/mol	P-value				
1	1	●	B	119	34	25648	◊	A	x,y,z-1	1_554	149	42	25389	1323.6	-18.0	0.033	6	3	0	0.000
	2	○	B	110	36	25648	◊	A	x,y,z	1_555	90	28	25389	946.7	-9.7	0.184	6	2	0	0.000
													Average:	1135.1	-13.9	0.109	6	3	0	0.000
2	3	○	B	85	29	25648	◊	A	x-1/2,-y+1/2,-z+1	3_456	89	32	25389	772.3	-0.5	0.735	9	2	0	0.000
3	4	○	A	61	11	25389	◊	A	-x+1,-y,z	2_655	60	11	25389	565.5	-3.0	0.478	16	2	0	0.000
	5	○	B	60	11	25648	◊	B	-x,-y+1,z	2_565	59	11	25648	533.5	-3.3	0.460	12	2	0	0.000
													Average:	549.5	-3.2	0.469	14	2	0	0.000
4	6	○	B	9	4	25648	◊	A	x-1/2,-y+1/2,-z	3_455	9	4	25389	86.9	-0.5	0.498	2	0	0	0.000
5	7	○	A	5	4	25389	x	A	-x+1/2,y-1/2,-z+1	4_546	6	2	25389	42.6	-0.2	0.486	0	0	0	0.000
6	8	○	A	6	6	25389	◊	B	-x+1/2,y-1/2,-z	4_545	7	6	25648	38.8	0.1	0.619	0	0	0	0.000
7	9	○	B	3	1	25648	x	B	-x+1/2,y-1/2,-z	4_545	2	2	25648	6.9	-0.0	0.585	0	0	0	0.000

PDBe PISA v1.52 [20/10/2014]

Figure 3.1 PDBePISA analysis of Cry11Ba interfaces. This displayed the various interfaces that were predicted via the partiality models that were initially obtained from multiple X-ray crystallography attempts with and without heavy metal soaking. From this analysis, the Cry11Ba sequence was threaded on the highly conserved δ-endotoxin 3 domain globular structure and matched the discovered residues with the PISA identified interfaces. This specified the different chains and calculated the surface area for each interface. The most stable residues were also highlighted when further expanding the interface data and were able to select Tyr residues at key interface points and whether they were responsible for stabilizing interactions. Each interface was further studied to identify other potential residues for different mutations that were not related to pH sensitivity.

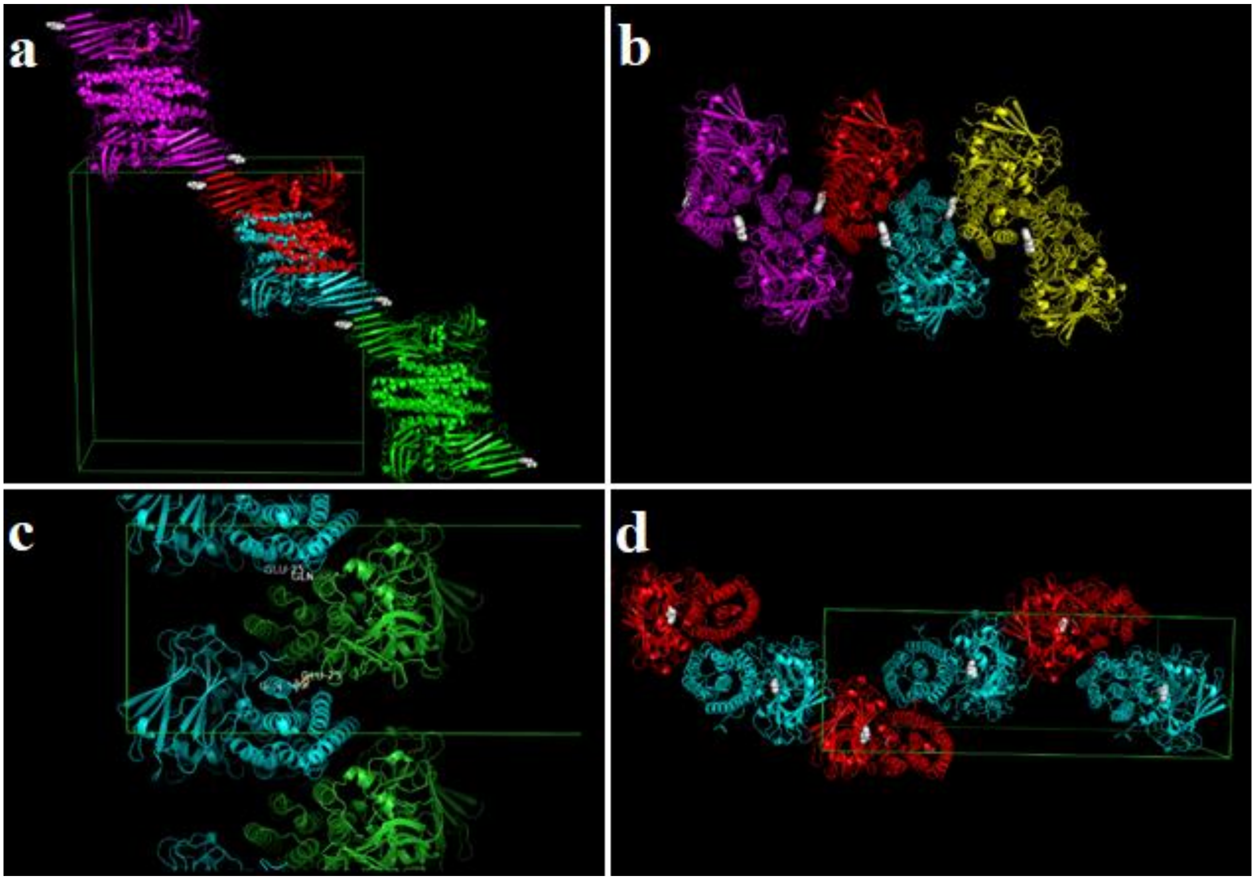


Figure 3.2 Cry11Ba selected mutations at PISA predicted interfaces and intrafaces. Each of the interfaces and intrafaces identified by PISA, contained Tyr residues near stabilizing residues. Each of these was selected for stability effects, along with possible toxicity effects. This included mutations at interfaces (**a**) Y350F and (**b**) Y453F and intrafaces (**c**) Y241F and (**d**) Y273F with the mutation sites marked as white spheres. The asymmetric unit containing a Cry11Ba dimer were colored as different chains (red and cyan/green) with different dimers being the another set colored together (magenta and yellow).

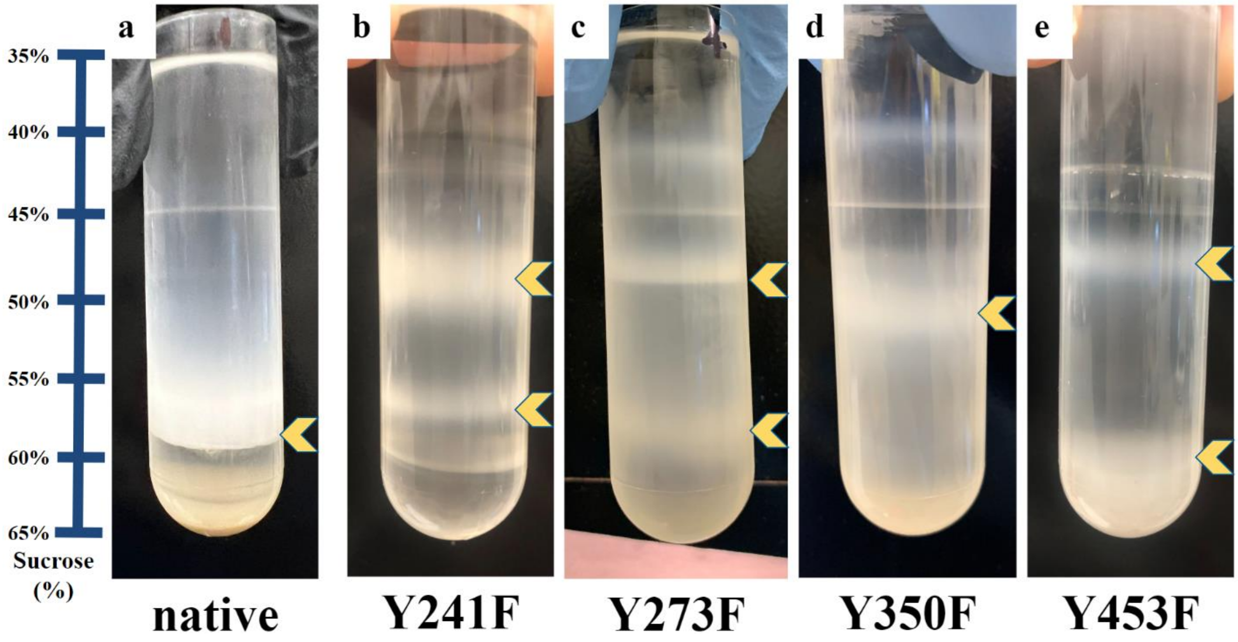


Figure 3.3 Comparison of Cry11Ba WT and mutants separated by sucrose gradient. Upon ultracentrifugation of the (**a**) WT and (**b-e**) mutant Cry11Ba, they are separated within the discontinuous 35-65% sucrose gradient. Each mutation displayed a "fingerprint" sucrose gradient, which differ from the native gradient. The separate bands were fractionated manually and contained either cell membranes, crystals (yellow arrows), endospores, or *Bti* cells. The crystal fractions were analyzed via phase contrast confocal and electron microscopy to confirm morphology with many smaller crystals and varying sizes produced by the mutations than the uniform native crystal fraction.

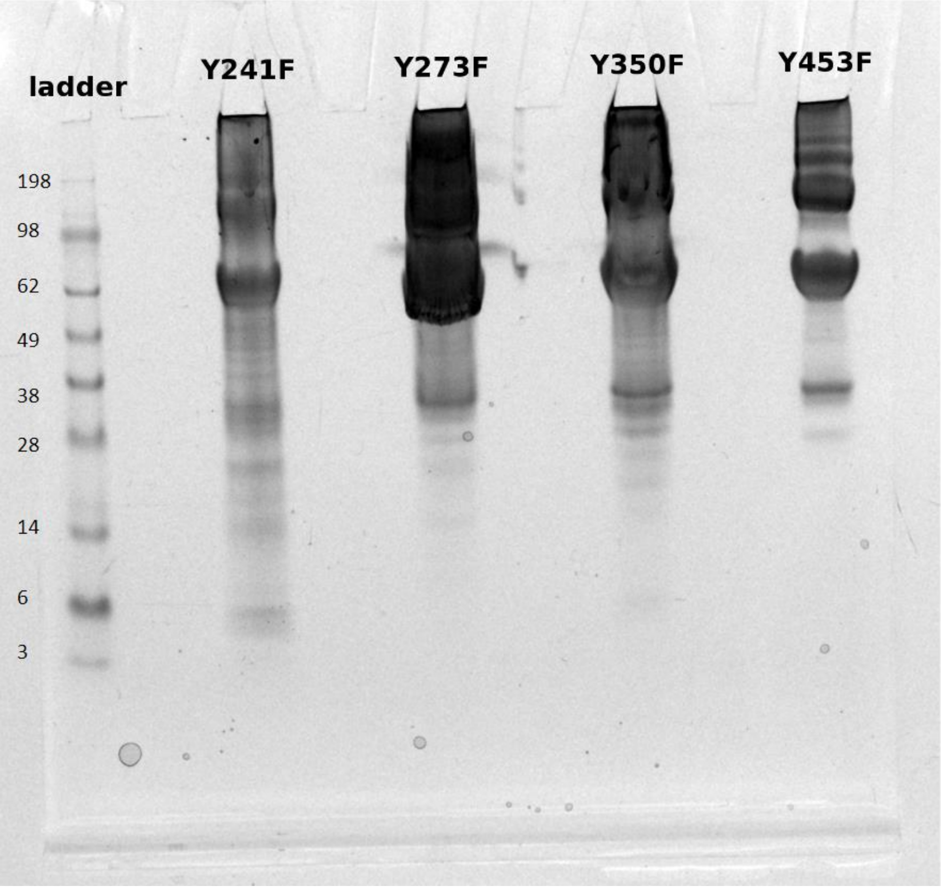


Figure 3.4 Purified Cry11Ba mutation constructs crystal fractions on SDS-PAGE. After separating the crystal fractions from the sucrose gradient, these were analyzed by SDS-PAGE to confirm the full-length proteins were present (~72 kDa). High-order multimers and possible degradation products were also observed above and below the monomer band, which would be expected as the crystalline form quaternary structure would contain multimers and the dissolution could induce degradation.

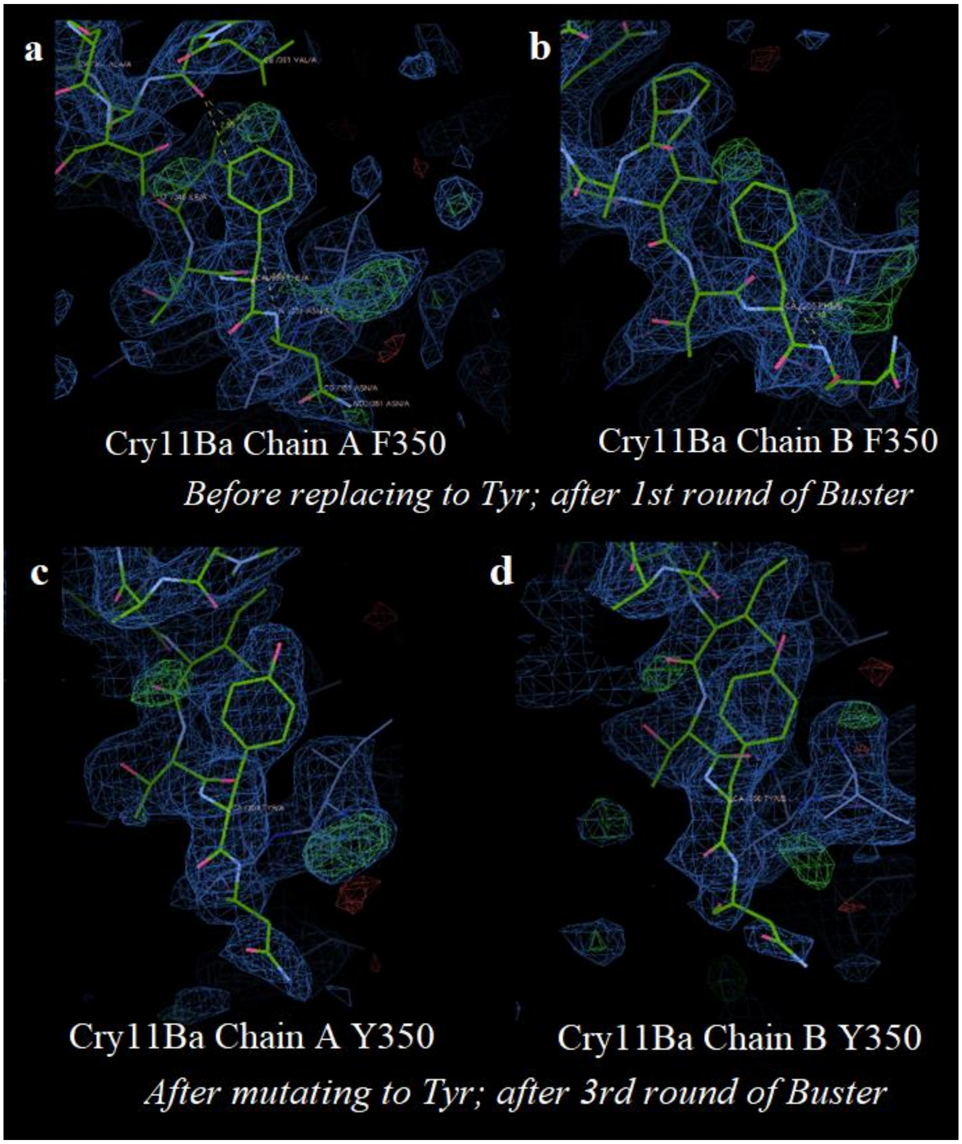
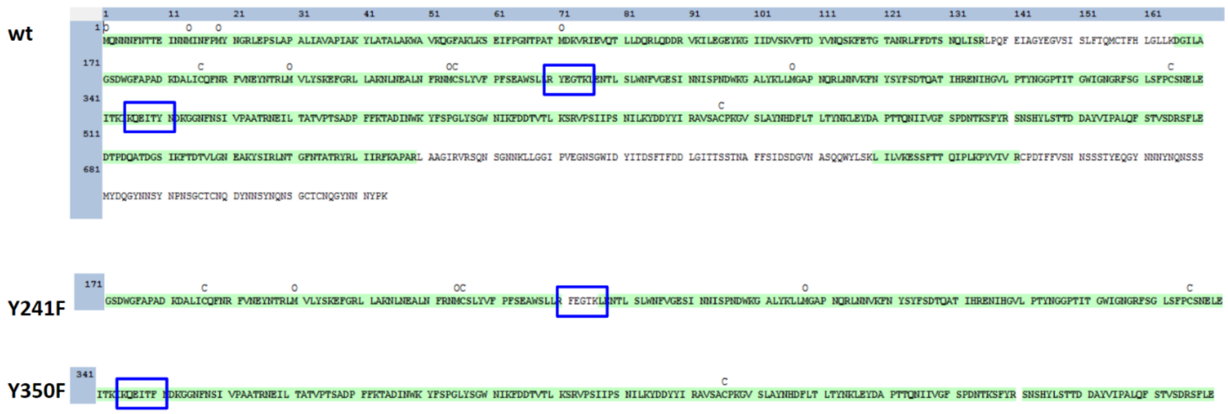


Figure 3.5 Structural refinement of Cry11Ba Y350F structure with Phe and Tyr. (**a,b**) The first Y350F Cry11Ba structure was refined using Buster to improve the structural statistics. The initial structures contained positive density above the F350, that had their initial sequnce mutated to remove any potential bias for the density above the phenyl ring in both Cry11Ba chains. (**c,d**) This was then mutated back to Y350 after 3 rounds of Buster refinement, which then satisfied the positive density with the addition of the hydroxyl group in that region. This was the first indication that the mutant constructs may contain WT Cry11Ba despite PCR sequencing of the primers and mutated vectors.



```
wt
1    MQNNNFNTTE INNMINFPMY NGRLEPSLAP ALIAVAPIAK YLATALAKWA VKQGFAKLKS EIFPGNTPAT MDKVRIEVQT LLDQRLQDDR VKILEGEYKG IIDVSKVFTD YVNQSKFETG TANRLFFDTS NQLISRLPQF EIAGYEGVSI SLFTQMCTFH LGLLKDGILA
171  GSDWGFAPAD KDALICQFNR FVNEYNTRLM VLYSKEFGRL LAKNLNEALN FRNMCSLYVF PFSEAWSLLR YEGTKLENTL SLWNFVGESI NNISPNDWKG ALYKLLMGAP NQRLNNVKFN YSYFSDTQAT IHRENIHGVL PTYNGGPTIT GWIGNGRFSG LSFPCSNELE
341  ITKIKQEITY NDKGGNFNSI VPAATRNEIL TATVPTSADP FFKTADINWK YFSPGLYSGW NIKFDDTVTL KSRVPSIIPS NILKYDDYYI RAVSACPKGV SLAYNHDFLT LTYNKLEYDA PTTQNIIVGF SPDNTKSFYR SNSHYLSTTD DAYVIPALQF STVSDRSFLE
511  DTPDQATDGS IKFTDTVLGN EAKYSIRLNT GFNTATRYRL IIRFKAPARL AAGIRVRSQN SGNNKLLGGI PVEGNSGWID YITDSFTFDD LGITTSSTNA FFSIDSDGVN ASQQWYLSKL ILVKESSFTT QIPLKPYVIV RCPDTFFVSN NSSSTYEQGY NNNYNQNSSS
681  MYDQGYNNSY NPNSGCTCNQ DYNNSYNQNS GCTCNQGYNN NYPK

Y241F
171  GSDWGFAPAD KDALICQFNR FVNEYNTRLM VLYSKEFGRL LAKNLNEALN FRNMCSLYVF PFSEAWSLLR FEGTKLENTL SLWNFVGESI NNISPNDWKG ALYKLLMGAP NQRLNNVKFN YSYFSDTQAT IHRENIHGVL PTYNGGPTIT GWIGNGRFSG LSFPCSNELE

Y350F
341  ITKIKQEITF NDKGGNFNSI VPAATRNEIL TATVPTSADP FFKTADINWK YFSPGLYSGW NIKFDDTVTL KSRVPSIIPS NILKYDDYYI RAVSACPKGV SLAYNHDFLT LTYNKLEYDA PTTQNIIVGF SPDNTKSFYR SNSHYLSTTD DAYVIPALQF STVSDRSFLE
```

Figure 3.6 GeLC-MS/MS of Y350F Cry11Ba purified crystals sequence identification. The mass spectroscopy sequence idenfication (green) for Y350F Cry11Ba of WT, Y350F, and Y241F peptide regions focused upon (blue box). The Y350F mutation was detected by mass spectroscopy (positve control). Y241F Cry11Ba (negative control) was not detected, which should not have been as they were not within the same vector. WT Cry11Ba was detected within the Y350F Cry11Ba confirming that the native sequence was also present within the mutation.

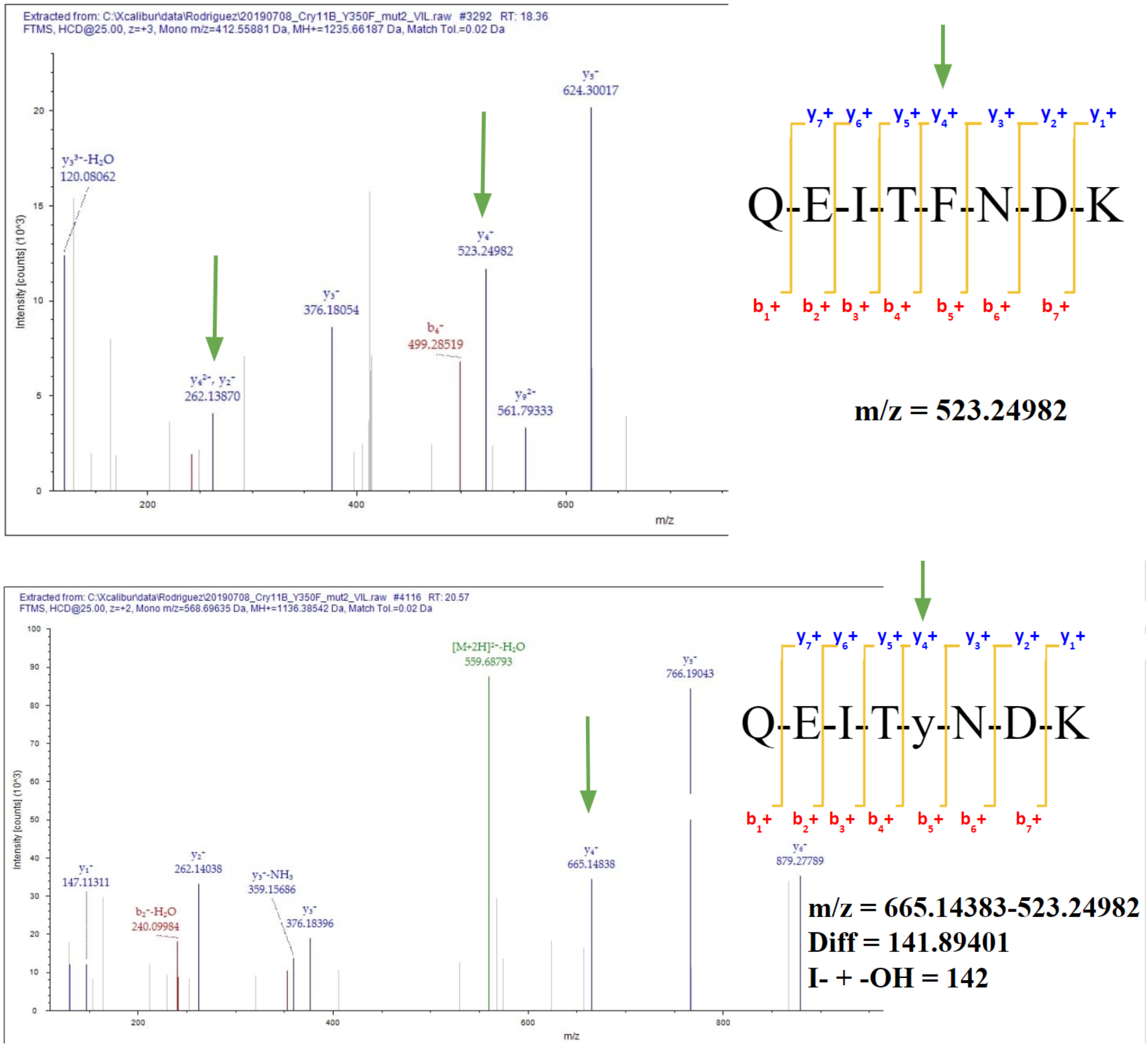
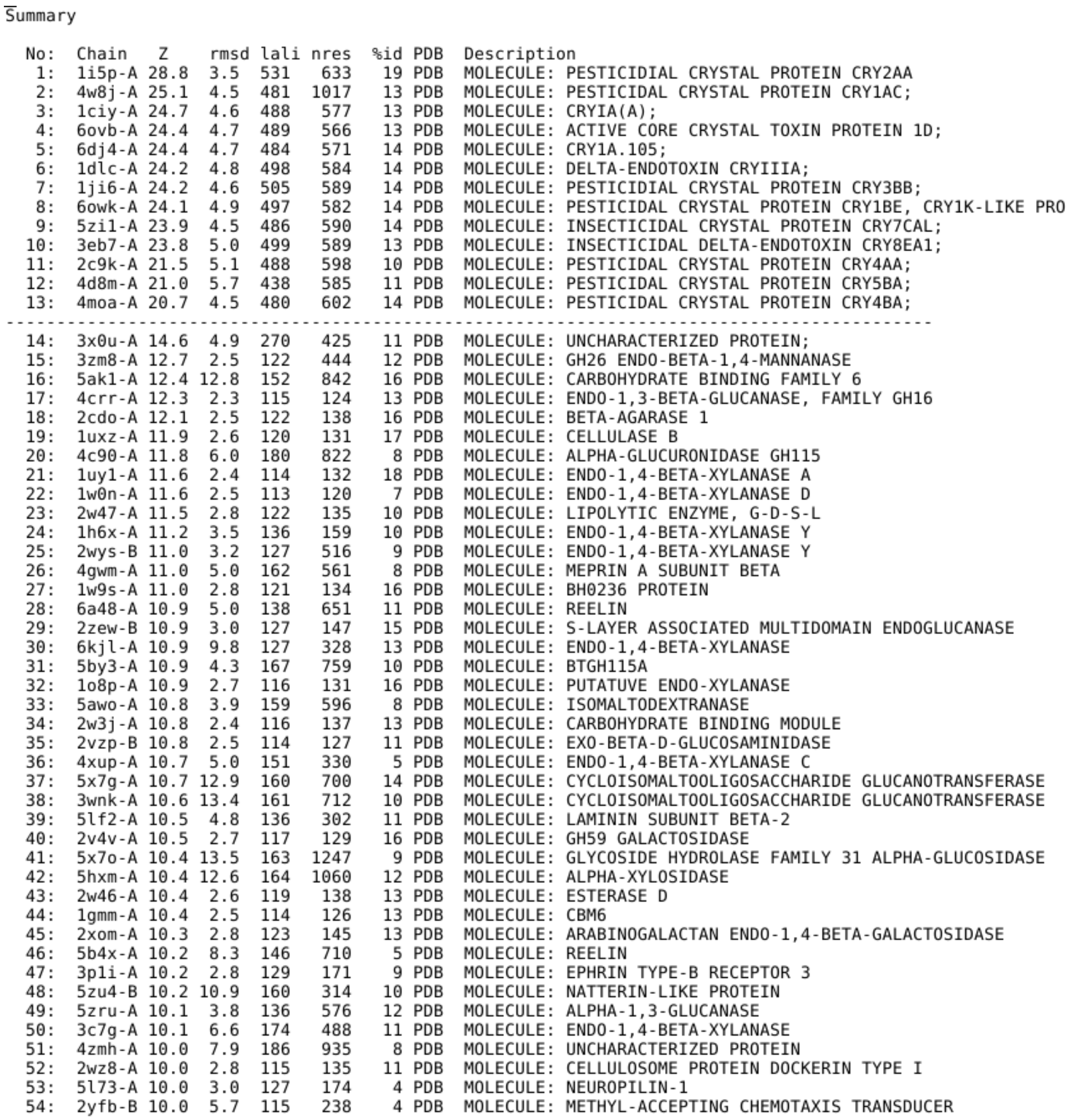


Figure 3.7 Mass spectroscopy MS2 spectras after vaporization iodonation labelling for Y350F Cry11Ba. After the vaporization iodonation labelling, the QEITFNDK peptide was identified by mass spectorscopy and displayed correct ionization peaks for the m/z= 523.25 amu. If WT Cry11Ba with Y350 was not present, then the iodonation labelling would have failed and no spectra or peaks would be associated with the QEITyNDK peptide sequence. The QEITyNDK with y being modified by -I atom, the shifted m/z=665.14 amu peak was identified along with the 523.25 amu peak. The difference of 141.89 amu is equivalent to the iodide and hydroxyl groups both being present on the Y350 phenyl ring.

Figure 3.8 GeLC-MS/MS of Y273F Cry11Ba purified crystals sequence identification. The mass spectroscopy sequence idenfication (green) for Y273F Cry11Ba of WT and Y273F peptide regions focused upon (blue box). The Y273F mutation was detected by mass spectroscopy. WT Cry11Ba was not detected within the Y273F Cry11Ba confirming that the native sequence was not present within the mutation. This was confirmed for each of the Cry11Ba mutations upon recreation with GeneBlocks to produce the new Cry11Ba mutations.



```
Summary

  No:  Chain   Z    rmsd lali nres  %id PDB  Description
   1:  1i5p-A 28.8  3.5  531   633   19 PDB  MOLECULE: PESTICIDIAL CRYSTAL PROTEIN CRY2AA
   2:  4w8j-A 25.1  4.5  481  1017   13 PDB  MOLECULE: PESTICIDAL CRYSTAL PROTEIN CRY1AC;
   3:  1ciy-A 24.7  4.6  488   577   13 PDB  MOLECULE: CRYIA(A);
   4:  6ovb-A 24.4  4.7  489   566   13 PDB  MOLECULE: ACTIVE CORE CRYSTAL TOXIN PROTEIN 1D;
   5:  6dj4-A 24.4  4.7  484   571   14 PDB  MOLECULE: CRY1A.105;
   6:  1dlc-A 24.2  4.8  498   584   14 PDB  MOLECULE: DELTA-ENDOTOXIN CRYIIIA;
   7:  1ji6-A 24.2  4.6  505   589   14 PDB  MOLECULE: PESTICIDIAL CRYSTAL PROTEIN CRY3BB;
   8:  6owk-A 24.1  4.9  497   582   14 PDB  MOLECULE: PESTICIDAL CRYSTAL PROTEIN CRY1BE, CRY1K-LIKE PRO
   9:  5zi1-A 23.9  4.5  486   590   14 PDB  MOLECULE: INSECTICIDAL CRYSTAL PROTEIN CRY7CAL;
  10:  3eb7-A 23.8  5.0  499   589   13 PDB  MOLECULE: INSECTICIDAL DELTA-ENDOTOXIN CRY8EA1;
  11:  2c9k-A 21.5  5.1  488   598   10 PDB  MOLECULE: PESTICIDAL CRYSTAL PROTEIN CRY4AA;
  12:  4d8m-A 21.0  5.7  438   585   11 PDB  MOLECULE: PESTICIDAL CRYSTAL PROTEIN CRY5BA;
  13:  4moa-A 20.7  4.5  480   602   14 PDB  MOLECULE: PESTICIDAL CRYSTAL PROTEIN CRY4BA;
-------------------------------------------------------------------------------------------
  14:  3x0u-A 14.6  4.9  270   425   11 PDB  MOLECULE: UNCHARACTERIZED PROTEIN;
  15:  3zm8-A 12.7  2.5  122   444   12 PDB  MOLECULE: GH26 ENDO-BETA-1,4-MANNANASE
  16:  5ak1-A 12.4 12.8  152   842   16 PDB  MOLECULE: CARBOHYDRATE BINDING FAMILY 6
  17:  4crr-A 12.3  2.3  115   124   13 PDB  MOLECULE: ENDO-1,3-BETA-GLUCANASE, FAMILY GH16
  18:  2cdo-A 12.1  2.5  122   138   16 PDB  MOLECULE: BETA-AGARASE 1
  19:  1uxz-A 11.9  2.6  120   131   17 PDB  MOLECULE: CELLULASE B
  20:  4c90-A 11.8  6.0  180   822    8 PDB  MOLECULE: ALPHA-GLUCURONIDASE GH115
  21:  1uy1-A 11.6  2.4  114   132   18 PDB  MOLECULE: ENDO-1,4-BETA-XYLANASE A
  22:  1w0n-A 11.6  2.5  113   120    7 PDB  MOLECULE: ENDO-1,4-BETA-XYLANASE D
  23:  2w47-A 11.5  2.8  122   135   10 PDB  MOLECULE: LIPOLYTIC ENZYME, G-D-S-L
  24:  1h6x-A 11.2  3.5  136   159   10 PDB  MOLECULE: ENDO-1,4-BETA-XYLANASE Y
  25:  2wys-B 11.0  3.2  127   516    9 PDB  MOLECULE: ENDO-1,4-BETA-XYLANASE Y
  26:  4gwm-A 11.0  5.0  162   561    8 PDB  MOLECULE: MEPRIN A SUBUNIT BETA
  27:  1w9s-A 11.0  2.8  121   134   16 PDB  MOLECULE: BH0236 PROTEIN
  28:  6a48-A 10.9  5.0  138   651   11 PDB  MOLECULE: REELIN
  29:  2zew-B 10.9  3.0  127   147   15 PDB  MOLECULE: S-LAYER ASSOCIATED MULTIDOMAIN ENDOGLUCANASE
  30:  6kjl-A 10.9  9.8  127   328   13 PDB  MOLECULE: ENDO-1,4-BETA-XYLANASE
  31:  5by3-A 10.9  4.3  167   759   10 PDB  MOLECULE: BTGH115A
  32:  1o8p-A 10.9  2.7  116   131   16 PDB  MOLECULE: PUTATUVE ENDO-XYLANASE
  33:  5awo-A 10.8  3.9  159   596    8 PDB  MOLECULE: ISOMALTODEXTRANASE
  34:  2w3j-A 10.8  2.4  116   137   13 PDB  MOLECULE: CARBOHYDRATE BINDING MODULE
  35:  2vzp-B 10.8  2.5  114   127   11 PDB  MOLECULE: EXO-BETA-D-GLUCOSAMINIDASE
  36:  4xup-A 10.7  5.0  151   330    5 PDB  MOLECULE: ENDO-1,4-BETA-XYLANASE C
  37:  5x7g-A 10.7 12.9  160   700   14 PDB  MOLECULE: CYCLOISOMALTOOLIGOSACCHARIDE GLUCANOTRANSFERASE
  38:  3wnk-A 10.6 13.4  161   712   10 PDB  MOLECULE: CYCLOISOMALTOOLIGOSACCHARIDE GLUCANOTRANSFERASE
  39:  5lf2-A 10.5  4.8  136   302   11 PDB  MOLECULE: LAMININ SUBUNIT BETA-2
  40:  2v4v-A 10.5  2.7  117   129   16 PDB  MOLECULE: GH59 GALACTOSIDASE
  41:  5x7o-A 10.4 13.5  163  1247    9 PDB  MOLECULE: GLYCOSIDE HYDROLASE FAMILY 31 ALPHA-GLUCOSIDASE
  42:  5hxm-A 10.4 12.6  164  1060   12 PDB  MOLECULE: ALPHA-XYLOSIDASE
  43:  2w46-A 10.4  2.6  119   138   13 PDB  MOLECULE: ESTERASE D
  44:  1gmm-A 10.4  2.5  114   126   13 PDB  MOLECULE: CBM6
  45:  2xom-A 10.3  2.8  123   145   13 PDB  MOLECULE: ARABINOGALACTAN ENDO-1,4-BETA-GALACTOSIDASE
  46:  5b4x-A 10.2  8.3  146   710    5 PDB  MOLECULE: REELIN
  47:  3p1i-A 10.2  2.8  129   171    9 PDB  MOLECULE: EPHRIN TYPE-B RECEPTOR 3
  48:  5zu4-B 10.2 10.9  160   314   10 PDB  MOLECULE: NATTERIN-LIKE PROTEIN
  49:  5zru-A 10.1  3.8  136   576   12 PDB  MOLECULE: ALPHA-1,3-GLUCANASE
  50:  3c7g-A 10.1  6.6  174   488   11 PDB  MOLECULE: ENDO-1,4-BETA-XYLANASE
  51:  4zmh-A 10.0  7.9  186   935    8 PDB  MOLECULE: UNCHARACTERIZED PROTEIN
  52:  2wz8-A 10.0  2.8  115   135   11 PDB  MOLECULE: CELLULOSOME PROTEIN DOCKERIN TYPE I
  53:  5l73-A 10.0  3.0  127   174    4 PDB  MOLECULE: NEUROPILIN-1
  54:  2yfb-B 10.0  5.7  115   238    4 PDB  MOLECULE: METHYL-ACCEPTING CHEMOTAXIS TRANSDUCER
```

Figure 3.9 Dali analysis of Cry11Ba against heuristic PDB search. Utilizing the Dali analysis server that samples from the PDB at various levels, where other proteins will be identified and ranked based on their similarity to the supplied (WT Cry11Ba) structure. The highest Z-score structures were other insecticidal δ-endotoxin proteins that share a highly similar 3 domain globular protein fold. Below these was a significant drop in Z-score, indicating lower structural similarity. These first 13 candidates were selected for further comparsion and analysis for functional and structural analysis.

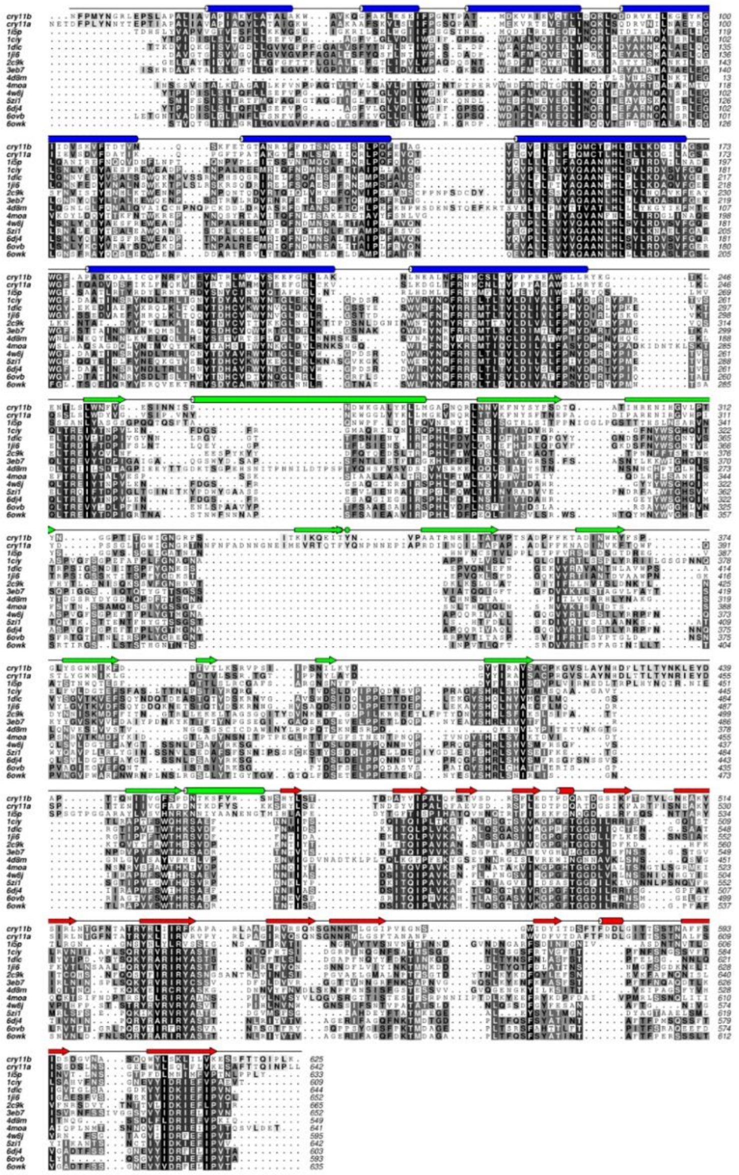


Figure 3.10 STOCCATO multiple sequence alignment of Cry11Ba and Cry11Aa against 13 δ-endotoxins. From the selected top Z-score candiates from Dali, these sequences were processed through STOCCATO for a multiple sequence alignment. This is aligning the sequences based upon their structure and accounts for any deletions or insertions into the individual protein's sequence. Domain 1 (blue), Domain 2 (green), and Domain 3 (red) have a significant amount of secondary structure with the most significant differences between WT Cry11Ba being the more central parts within Domain 2, with 2 large amino acid insertions.

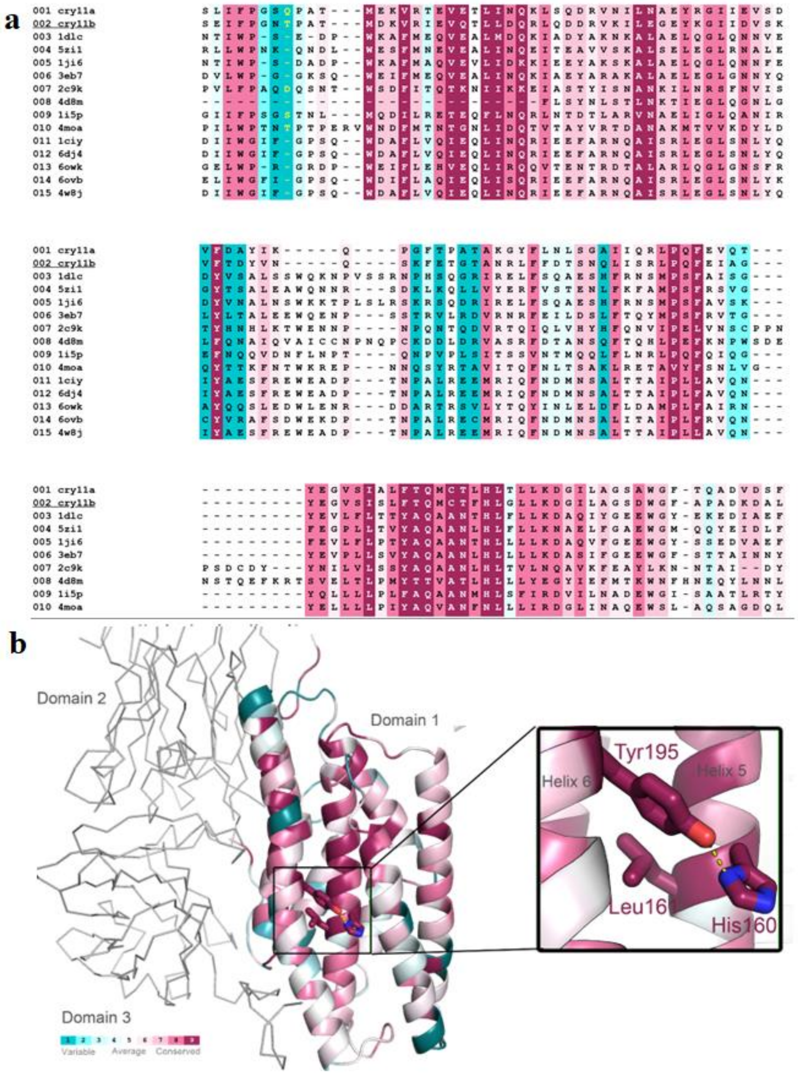


Figure 3.11 ConSurf multiple sequence alignment of Cry11Ba Domain 1 against 14 δ-endotoxins with conservation surface mapping. (a) The aligned sequences of Cry11Ba and the other selected candidates were then colored according the their conservation level with high (maroon) to low (dark teal) levels. These conservation levels also took into account the residue class as nonpolar, polar, and charged. **(b)** The same conservation mapping levels were colored on the Cry11Ba structure, which displayed more conservation towards the central part within the alpha-helical bundle in Domain 1. This also identified Tyr195 as the single completed conserved residue with His160 and Leu161 being highly conserved.

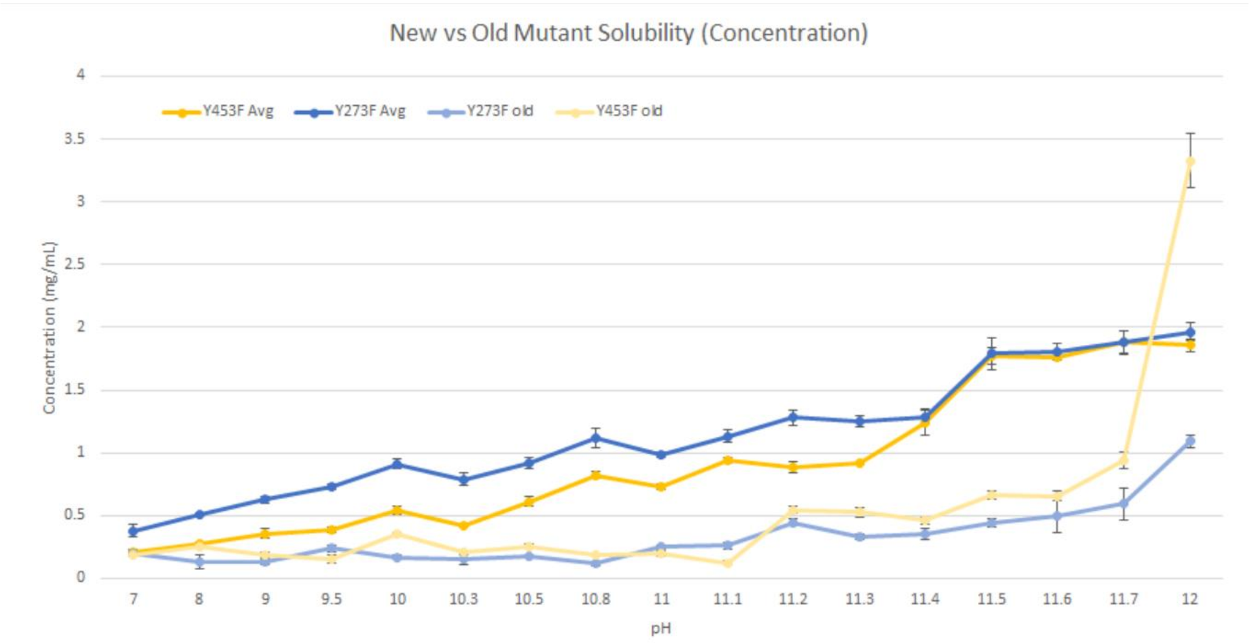
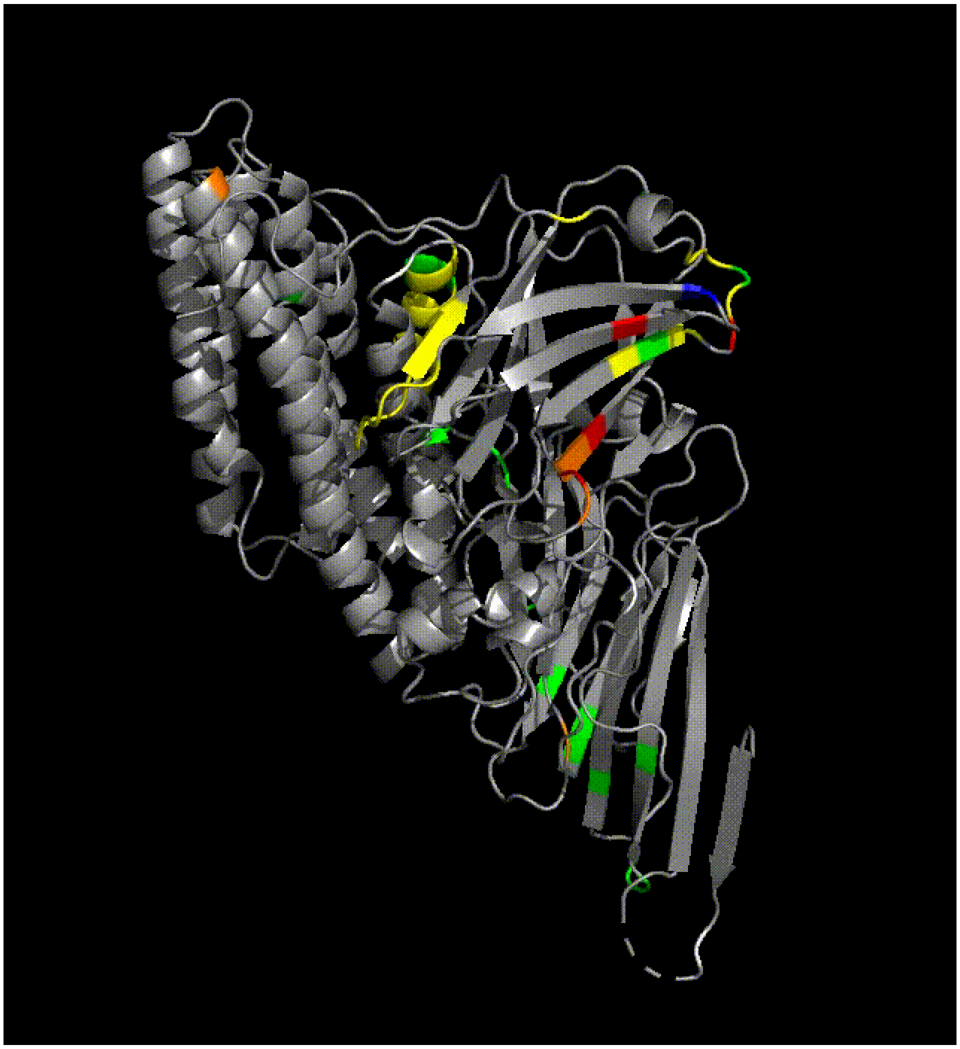
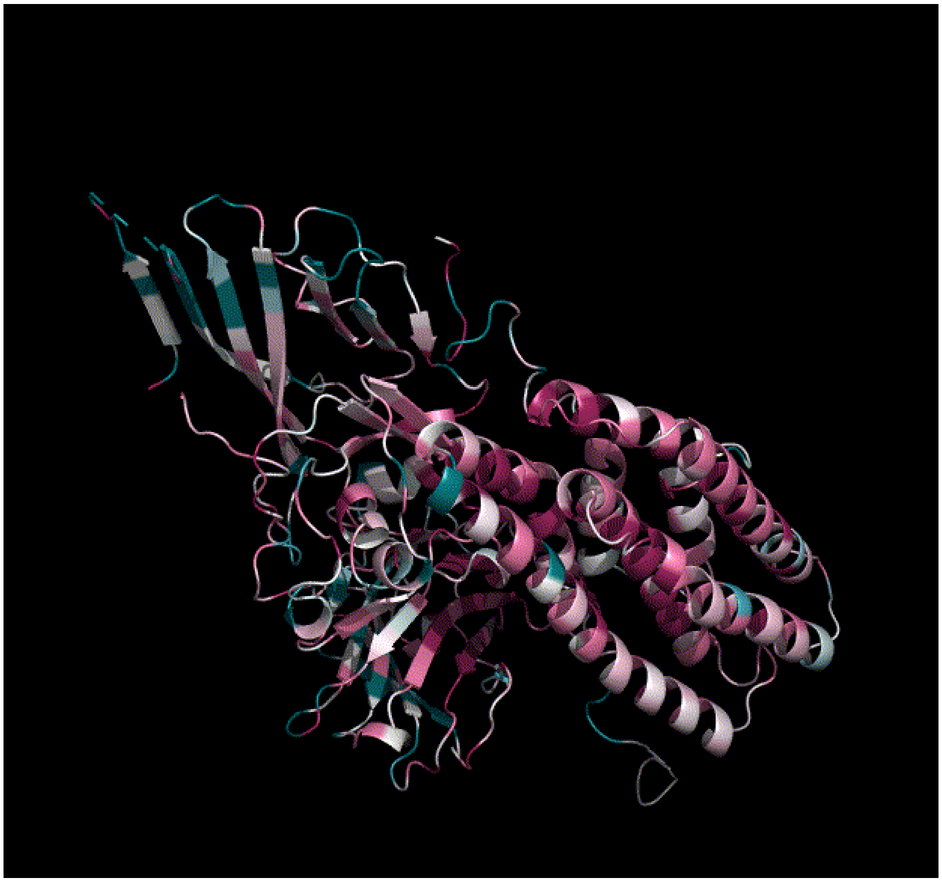


Figure 3.12 Solubility assay comparing old and new mutant constructs of Cry11Ba. The old and new mutatnt preparations displayed varying levels of dissolution (mg/mL) indicating a shift in pH sensitivity. The old Y273F (light blue) and Y453F (light yellow) indicated a notable uptake into solution at pH 11.7, which were contaminated with WT Cry11Ba. The new Y273F (dark yellow) and Y453F (dark blue) that were confirmed by mass spectroscopy, displayed a gradual increase in uptake with the largest increase between pH 11.3-11.5. The initial levels for Y273F were higher than the original and each of the mutants, indicative of a decrease in stability of the crystal.

3.5 Videos



Video 3.1 δ -endotoxin mutations mapped with different levels of toxicity. A variety of mutations have been conducted upon Cry and Cyt proteins to probe their toxicity levels. These levels consisted of higher toxicity (green), lower toxicity (red), equal toxicity (yellow), mixed toxicity (blue), and toxicity not reported (yellow). The majority of mutations that increased toxicity were in Domain 2 within the beta-sheet and towards the central part of the protein, were most conservation occurred. Domain 3 had the greatest variability in mutations and their effect on toxicity.



Video 3.2 ConSurf conservation surface map of Cry11Ba and Cry11Aa against 13 δ-endotoxins. The ConSurf conservation mapping levels colored on all 3 domains of the WT Cry11Ba structure, which displayed more conservation (maroon) towards the central part of the overal protein in Domains 1, 2, and 3. There is higher variablity (deep teal) at the edges of the B-sheets and strands of Domains 2 and 3.

3.6 Tables

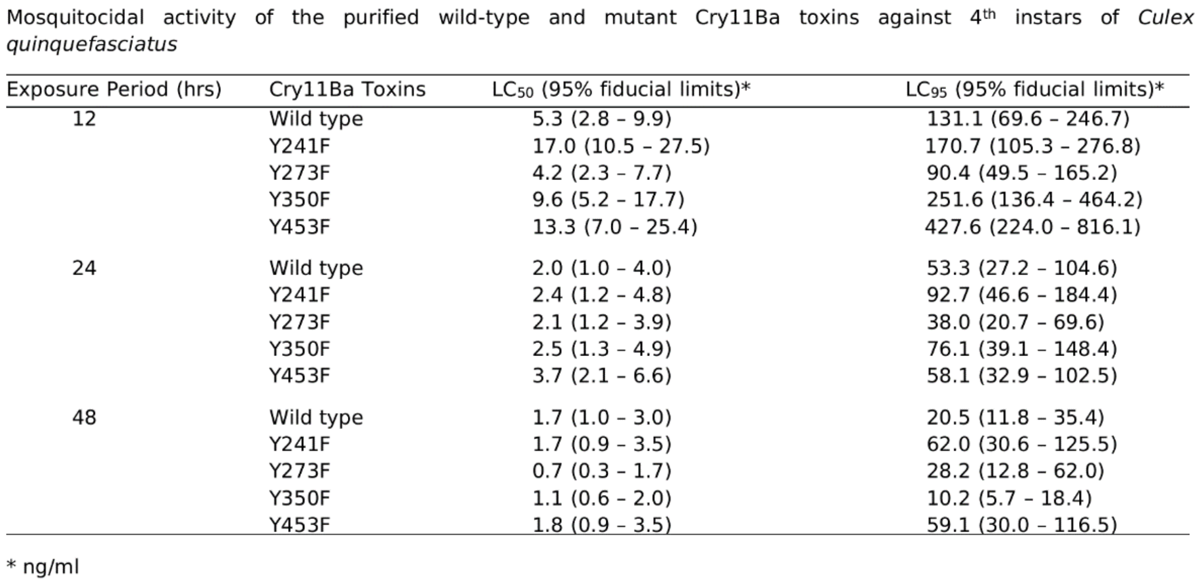
Mut.	Interaction	Reasoning	Cause	Effect/Impact	Location	Symmetry
Y241F	Intramolecular (at an interface)	H-bonded between domains, same chain, D590	Loss of H-bond	More flexible, higher entropy barrier to crystallization	Between Helical and C-term B-sheet domains	None
Y273F	Intramolecular	H-bonded between domains, same chain, R518	Loss of H-bond	More flexible, higher entropy barrier to crystallization	Between interrupted helix and loop	None
Y350F	Intramolecular (at an interface)	H-bond to neighboring strand, P352 carbonyl	Loss of H-bond	Change of backbone conformation with neighboring strand near interface (supporting conformation compatible with interface)	H-bond sheet-sheet	Two-fold symmetry
Y453F	Intermolecular	H-bond to neighboring strand, T318 Carbonyl	Loss of H-bond	Maintain the interface contact, but reduce of the strength of the interaction	H-bond loop-loop (450 - 460, Chain A) (315 - 325, Chain B)	Two-fold symmetry

Table 3.1 Cry11Ba mutation selection and rationale. The selection process for Cry11Ba mutations. These are the factors contributing to the mutations and the hypothesized result when inducing that mutation.

a	**Y241F**	**Y273F**	**Y350F**	**Y453F**
F	16.79195	10.21232	22.2864	10.57612
P-Value	0.000411	0.003881	8.44E-05	0.003385
F Table Value (0.01)	7.823	7.823	7.823	7.823

b	**Y241F**	**Y273F**	**Y350F**	**Y453F**
F	49.10665	4.790597	7.565861	0.031991
P-Value	0.005968	0.116391	0.070706	0.869445
F Table Value (0.1)	5.53832	5.53832	5.53832	5.53832
F Table Value (0.05)	10.128	10.128	10.128	10.128

Table 3.2 Cry11Ba mutants vs WT pH ANOVA. The statistically significant values and analysis for (**a**) the pH trigger/uptake shift and (**b**) initial stability/pH floor uptake shift. Each value had the mutations' shift was statistically significant with 99% confidence level for the pH trigger shift and 95% confidence level for the pH floor uptake shift.



Mosquitocidal activity of the purified wild-type and mutant Cry11Ba toxins against 4th instars of *Culex quinquefasciatus*

Exposure Period (hrs)	Cry11Ba Toxins	LC_{50} (95% fiducial limits)*	LC_{95} (95% fiducial limits)*
12	Wild type	5.3 (2.8 – 9.9)	131.1 (69.6 – 246.7)
	Y241F	17.0 (10.5 – 27.5)	170.7 (105.3 – 276.8)
	Y273F	4.2 (2.3 – 7.7)	90.4 (49.5 – 165.2)
	Y350F	9.6 (5.2 – 17.7)	251.6 (136.4 – 464.2)
	Y453F	13.3 (7.0 – 25.4)	427.6 (224.0 – 816.1)
24	Wild type	2.0 (1.0 – 4.0)	53.3 (27.2 – 104.6)
	Y241F	2.4 (1.2 – 4.8)	92.7 (46.6 – 184.4)
	Y273F	2.1 (1.2 – 3.9)	38.0 (20.7 – 69.6)
	Y350F	2.5 (1.3 – 4.9)	76.1 (39.1 – 148.4)
	Y453F	3.7 (2.1 – 6.6)	58.1 (32.9 – 102.5)
48	Wild type	1.7 (1.0 – 3.0)	20.5 (11.8 – 35.4)
	Y241F	1.7 (0.9 – 3.5)	62.0 (30.6 – 125.5)
	Y273F	0.7 (0.3 – 1.7)	28.2 (12.8 – 62.0)
	Y350F	1.1 (0.6 – 2.0)	10.2 (5.7 – 18.4)
	Y453F	1.8 (0.9 – 3.5)	59.1 (30.0 – 116.5)

* ng/ml

Table 3.3 Cry11Ba WT and mutation biotoxicity assay against *Culex quinquefasciatus*. Each Cry11Ba mutation and WT was tested for toxicity against *C. quinquefasciatus*. They were fed to 4th-instars and displayed various levels of LC50 and LC95 levels. The majority of mortality is observed within 24 hrs, with initial death observed at 2-3 hrs, thus a 12 hr time point was collected. This displayed increased toxicity of Y273F initially, but upon reaching 48 hrs, both Y273F and Y350F displayed high toxicity and Y241F and Y453F maintaining toxicity levels equalt to WT.

3.7 References

(1) Tetreau, G.; Sawaya, M. R.; De Zitter, E.; Andreeva, E. A.; Banneville, A.-S.; Schibrowsky, N. A.; Coquelle, N.; Brewster, A. S.; Grünbein, M. L.; Kovacs, G. N.; Hunter, M. S.; Kloos, M.; Sierra, R. G.; Schiro, G.; Qiao, P.; Stricker, M.; Bideshi, D.; Young, I. D.; Zala, N.; Engilberge, S.; Gorel, A.; Signor, L.; Teulon, J.-M.; Hilpert, M.; Foucar, L.; Bielecki, J.; Bean, R.; de Wijn, R.; Sato, T.; Kirkwood, H.; Letrun, R.; Batyuk, A.; Snigireva, I.; Fenel, D.; Schubert, R.; Canfield, E. J.; Alba, M. M.; Laporte, F.; Després, L.; Bacia, M.; Roux, A.; Chapelle, C.; Riobé, F.; Maury, O.; Ling, W. L.; Boutet, S.; Mancuso, A.; Gutsche, I.; Girard, E.; Barends, T. R. M.; Pellequer, J.-L.; Park, H.-W.; Laganowsky, A. D.; Rodriguez, J.; Burghammer, M.; Shoeman, R. L.; Doak, R. B.; Weik, M.; Sauter, N. K.; Federici, B.; Cascio, D.; Schlichting, I.; Colletier, J.-P. De Novo Determination of Mosquitocidal Cry11Aa and Cry11Ba Structures from Naturally-Occurring Nanocrystals. *Nat Commun* **2022**, *13* (1), 4376. https://doi.org/10.1038/s41467-022-31746-x.

(2) Labbé, P.; David, J.-P.; Alout, H.; Milesi, P.; Djogbénou, L.; Pasteur, N.; Weill, M. Evolution of Resistance to Insecticide in Disease Vectors. In *Genetics and Evolution of Infectious Diseases*; Elsevier, 2017; pp 313–339. https://doi.org/10.1016/B978-0-12-799942-5.00014-7.

(3) Holm, L. DALI and the Persistence of Protein Shape. *Protein Science* **2020**, *29* (1), 128–140. https://doi.org/10.1002/pro.3749.

(4) Shatsky, M.; Nussinov, R.; Wolfson, H. J. Optimization of Multiple-Sequence Alignment Based on Multiple-Structure Alignment. *Proteins: Structure, Function, and Bioinformatics* **2005**, *62* (1), 209–217. https://doi.org/10.1002/prot.20665.

(5) Ashkenazy, H.; Abadi, S.; Martz, E.; Chay, O.; Mayrose, I.; Pupko, T.; Ben-Tal, N. ConSurf 2016: An Improved Methodology to Estimate and Visualize Evolutionary Conservation in Macromolecules. *Nucleic Acids Res* **2016**, *44* (W1), W344–W350. https://doi.org/10.1093/nar/gkw408.

(6) Ashkenazy, H.; Erez, E.; Martz, E.; Pupko, T.; Ben-Tal, N. ConSurf 2010: Calculating Evolutionary Conservation in Sequence and Structure of Proteins and Nucleic Acids. *Nucleic Acids Res* **2010**, *38* (Web Server), W529–W533. https://doi.org/10.1093/nar/gkq399.

(7) Glaser, F.; Pupko, T.; Paz, I.; Bell, R. E.; Bechor-Shental, D.; Martz, E.; Ben-Tal, N. ConSurf: Identification of Functional Regions in Proteins by Surface-Mapping of Phylogenetic Information. *Bioinformatics* **2003**, *19* (1), 163–164. https://doi.org/10.1093/bioinformatics/19.1.163.

(8) Krissinel, E. Crystal Contacts as Nature's Docking Solutions. *J Comput Chem* **2010**, *31* (1), 133–143. https://doi.org/https://doi.org/10.1002/jcc.21303.

(9) Krissinel, E.; Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J Mol Biol* **2007**, *372* (3), 774–797. https://doi.org/https://doi.org/10.1016/j.jmb.2007.05.022.

(10) Akiba, T.; Okumura, S. Parasporins 1 and 2: Their Structure and Activity. *J Invertebr Pathol* **2017**, *142*, 44–49. https://doi.org/https://doi.org/10.1016/j.jip.2016.10.005.

(11) Boudko, D. Y.; Moroz, L. L.; Harvey, W. R.; Linser, P. J. Alkalinization by Chloride/Bicarbonate Pathway in Larval Mosquito Midgut. *Proceedings of the National Academy of Sciences* **2001**, *98* (26), 15354–15359. https://doi.org/10.1073/pnas.261253998.

(12) Abdullah, M. A. F.; Valaitis, A. P.; Dean, D. H. Identification of a Bacillus Thuringiensis Cry11Ba Toxin-Binding Aminopeptidase from the Mosquito, Anopheles Quadrimaculatus. *BMC Biochem* **2006**, *7* (1), 16. https://doi.org/10.1186/1471-2091-7-16.

(13) Xu, C.; Wang, B.-C.; Yu, Z.; Sun, M. Structural Insights into Bacillus Thuringiensis Cry, Cyt and Parasporin Toxins. *Toxins (Basel)* **2014**, *6* (9), 2732–2770.

(14) Lereclus, D. Transformation and Expression of a Cloned δ-Endotoxin Gene in Bacillus Thuringiensis. *FEMS Microbiol Lett* **1989**, *60* (2), 211–217. https://doi.org/10.1016/0378-1097(89)90511-9.

Chapter 4

Electron microscopy of Cry11Ba crystalline inclusions at biologically relevant alkaline pH

4.1 Results and Discussion

The crystalline (Cry) and cytolytic (Cyt) paratoxins are unique in their ability to crystallize and dissolve into pro- and active toxin, which we are interested in understanding the mode of action of these Cry11Ba crystalline inclusions further by means of cryo-tomography to detect the changes occurring in the protein states. The first approach was inducing the alkaline state upon the Cry11Ba native crystals and injecting them utilizing the micro-injector on the XFEL to capture these changes over time. The data collected at the alkaline conditions was sufficient for another structure determination of Cry11Ba at pH 10.4, but did have reduced refinement statistics compared to the native structure at neutral pH (**Chapter 2, Table 1**).[1] By utilizing an F_{obs}-F_{obs} difference density maps, which subtract the electron densities from one another in order to determine change in one direction or another between two structures, differences were expected to have correlation in locations with either positive or negative densities to indicate changes within the structure that were affected by the pH change. These changes were predicted to be between domains to indicate a change from the crystal to protoxin or active toxin state. Upon conducting this calculation, the differences were detected within the structure; however, these differences were not as extensive as we originally predicted, and thus further analysis was necessary utilizing a different technique.

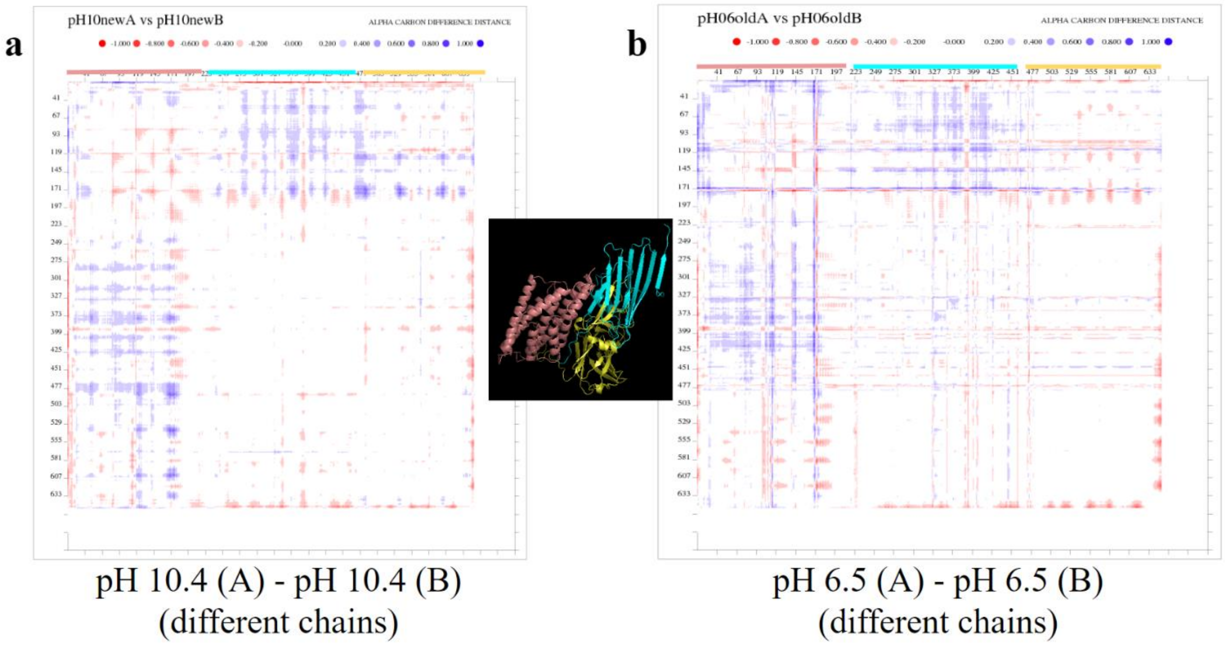
Utilizing pre-calculated XFEL data, an additional structure was solved alongside the wildtype (WT) Cry11Ba at neutral pH (6.5), an alkaline pH (10.4) Cry11Ba structure. The higher pH structure was to simulate the alkaline pH levels that the crystals would be subjected to within the *Aedes aegypti* posterior gut upon ingestion and binding. Differences were observed between the WT Cry11Ba pH 6.5 and pH 10.4, notably that the crystal appeared to be condensing at the alkaline pH compared to the neutral pH structure, which was observed by utilizing a distance difference plot (ddPlot) (**Figs. 4.1,4.2**). This would indicate that upon being introduced to the alkaline environment in the *A. aegypti*'s gut, the protein domains would retract and begin to increase the distance from one another and thus lose their interaction points with other chains, thus reducing stability and causing the crystal to fall apart and dissolve. During XFEL data collection, we noticed a significant reduction in the data quality and in order to combat this, we increased the glycerol percentage to help hold the crystals within the suspension and improve the data of the smaller crystals that were resulting from the WT Cry11Ba being exposed to the alkaline pH over the long data collection time period (ie. a few hours).[1] This could have conflated the observed separating distances between the domains and chains for the Cry11Ba crystal contacts as the increased glycerol, 30% to 40%, dehydrates the system and thus would result in constricting the domains from one another.

With these structures leading to more questions than answers, another structural technique was selected to try and study the changes occurring within Cry11Ba at alkaline pH. The biggest challenge to overcome with studying the Cry11Ba conformational changes is that this protein is packaged into a crystal. As previously discussed, these toxins are unique in that their packaging mechanism is crystalline. These crystals, while typically well-ordered in their crystal lattice, do experience the barrier that they are much too small for typical X-ray crystallography structure determination. However, they can be imaged at high magnification by electron microscopy to obtain high-resolution information. By negatively staining the Cry11Ba crystals with uranyl acetate, the lattice of the crystal was observed, and weak peaks were observable by taking the Fourier transform (FT) of the transmission electron micrographs (TEM). The conformational changes that were being induced at the alkaline pH could be changing the crystal lattice as well, resulting in different diffraction patterns. This could be more easily observed by cryo-electron tomography (cryo-ET), which takes images at different tilt angles of the sample stage and from here can be reconstructed into a stack of images resulting in a 3D structure.[2–6] cryo-ET's imaging power fills the imaging gap within the aforementioned range and would allow the possible elucidation of the pH-induced crystal lattice changes. In order to capture the changes, a time course on the crystals was conducted to see how quickly the crystals were dissolving into the solution and by how much. After conducting the typical solubility assay and shortening it to different time frames, it was found that most of the crystal pellet dissolution occurred within 10 minutes. With the majority of crystal dissolution occurring within the initial 10 minutes, time points of 1, 5, and 10 minutes were selected to capture the change that was expected to be occurring within the crystal lattice. After observing that crystals were more dissolved/damaged at 10 minutes than 1 minute via TEM images, optimization in terms of concentration, distribution, and freezing conditions were determined. These datasets would produce tilt series that, when the Fourier transform was performed, display a diffraction pattern movie (**Fig. 4.3a,b**). The procedure of grid preparation was modified for incorporate vitrification, aka freezing, of the grids, by utilizing 2:1 Cu Quantifoil grids to decrease background, removing negatively stain to preserve biological structure, and fiducials were added to aid in the tracking of the grid as data was collected for the cryo-ET tilt series. A total of 65 tilt series were collected on a 300 kEV Titan-KRIOS equipped with a Volta phase plate. This phase plate aids in enhancing image contrast and can collect in-focus data, which means an increase in the signal-to-noise ratio. Tilt series were collected and processed using the previously designed workflows through iMOD and etomo.[7] There were little to no changes observed in the reconstructed stacks of the crystal lattice to indicate a change from the alkaline pH environment (**Video 4.1**) and the FTs of these tilt series showed diffraction peaks out to 3.0 Ǻ, which was below the resolution (2.3 - 2.9 Å) of our previously solved structures of Cry11Ba by XFEL studies (Chapter 2).[1]

Even though the expected changes were not observed within the cryo-ET datasets, other interesting characteristics of these samples were observed. Small particles were observed to be coming off of the crystal. This was believed to be the crystal's reaction to the alkaline solution environment, as the Cry11Ba crystals would slowly dissolve and transition from Cry11Ba solid to soluble state. To further investigate this, Cry11Ba crystals underwent the same optimized conditions for cryo-ET, but were not vitrified and were analyzed via negative stain TEM. In order to amplify the Cry11Ba nano-structures, heavier staining than previous screening conditions was conducted and the Cry11Ba particles were observed (**Fig. 4.4a-d**). Upon initial imaging of the supernatants on F/C grids that had been glow discharged and negatively stained, small particulates were observed more clearly, with a few species were recurring throughout the analysis, including spheres (**Fig. 4.4e**), a trefoil (**Fig. 4.4f**), and a barrel (**Fig. 4.4g**). By further probing, the early points of protein dissolution did display the greatest or equal amounts of protein via normalized samples on an SDS-PAGE gel. These samples also displayed high molecular weights above the 192 kDa standard band, which is indicative of a multimer (ie. trimer or tetramer) above the monomer band at 82 kDa (**Fig. 4.5**). These higher molecular weight bands were not crystals within the samples, as the supernatant was separated from the pellet and was then filtered to remove any residual crystals remaining. The mode of action for the Cry and Cyt proteins has been known as pore-forming within the target vector's gut membrane; however, the structure of this active toxin state has not been observed or determined due to the difficulty of obtaining active toxins in their theorized quaternary structure.[8] Single particle EM would allow the natively produced toxins, in this case Cry11Ba, to be studied in a similar alkaline environment to the host organisms, only absent of cofactors and membrane receptors that may bind these protoxins. After observing these recurring particles, negatively stained EM samples were optimized to obtain the right concentration to allow for these particle species to be prominent on the grid without overcrowding the background, which would reduce the contrast and decrease the signal-to-noise ratio. After data collection of single TEMs, data processing was first conducted with REgularized LIkelihood OptimizatioN (RELION), which allows for refinement of macromolecular structures by single-particle analysis of EM samples. RELION, unlike other programs that require user-expertise of a program's paraments, utilizes a Bayesian approach to infer parameters of a statistical model from the data and recategorizes similarly shaped particles into unique groups/classes.[9] From here, these classes are selected for analysis to reconstruct 3D models of the data which, in turn, is dependent upon high resolution EM data and the number of particles at high magnification. These particles also need to be in random orientations in order to sample all angles of the protein and provide a full image of the particle of interest to produce a 3D volume model. After finishing the preprocessing of the TEMs, the particles of interest were selected "by-hand" (manually) and with strong contrast. This type of manual selection does ensure a more stringent selection of particles, but does limit the highest possible resolution as there are a limited number of particles that are selected for analysis. After combining data collections, the 2D class analysis consistently produced similarly shaped classes, including the initially described trefoil and barrels with 987 particles (**Fig. 4.6**). The trefoil 2D classes were more heavily populated and included multiple views of these particles of interest, with estimated resolutions ranging from 15.43 - 20.25 Å. By selecting and combining these 2D classes, an initial 3D model was produced at 8.1 Å. While from the 2D classes, the Cry11Ba monomers appear to interact at a central point, the 3D density volume shows these to have one continuous tube of density in the shape of the trefoil (**Fig. 4.7a,b**). After trying a few poses of the Cry11Ba monomers, there was not high enough resolution in the density to determine the correct pose of a single Cry11Ba chain(**Fig. 4.8a-c**). This is most likely due to the lack of particles in the pool, which is limiting the resolution from going higher than 8.1 Å. To overcome this, cryo-EM Single Particle Abi-Initio Reconstruction and Classification (cryoSPARC) was selected as the other single particle analysis software that was designed to refine multiple high-resolution 3D structures with little to no user input from single-particle images.[10] One of the features of cryoSPARC is the ability to select from a TEM and use this as a training set for auto (template) particle picking by the program, thus significantly reducing the work of the user with a trade-off of a less stringent particle selection process.[10] More classes were added as a results and binning the 81,320 particles led to 30 total classes. Some of the 2D classes were excluded as either "junk" classes, containing no or partial particles, or a sphere morphology, leaving the remaining 21,689 particles for analysis and an estimated resolution of 12 Å (**Fig. 4.9**).

The initial 3D model produced by cryoSPARC is ~5.6 Å and upon further refinement did obtain a 3D density map, which took shape of an hourglass shape upon further class averaging and refinement (**Fig. 4.10**). While this is different from the tube of density observed with the RELION single particle analysis, it is possible that due to the larger number of particles being auto-selected, there could be enough to average out the two lobes into this hourglass shape that is observed (**Fig. 4.11**). Despite the results, it is reassuring that RELION, cryoSPARC, and etomo, all displayed similar 2D classes (**Figs. 4.7,4.10,4.12**), especially the trefoil and the barrel. With each of these single-particle data analyses, a beginning of understanding Cry11Ba multimers visually can lead to how the mode of action for delta-endotoxins is occurring in these multimer states when solubilized at alkaline pH experienced within the target hosts' guts and begin their pore-forming toxin action.

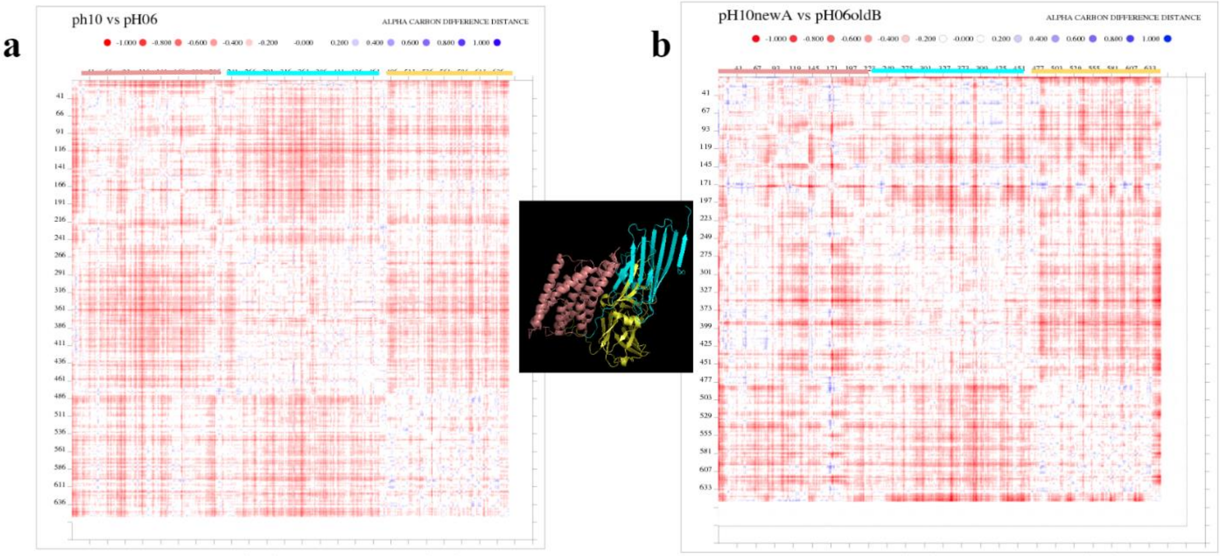
4.2 Figures



pH 10.4 (A) - pH 10.4 (B)
(different chains)

pH 6.5 (A) - pH 6.5 (B)
(different chains)

Figure 4.1 Difference distance plot (ddPlot) of WT Cry11Ba between two chains. These are the differences between the distances of two chains within the asymmetric unit at only pH (**a**) 10.4 or (**b**) 6.5. The closer two residues are the distance will be marked with blue, while two residues that are further apart in distance will be marked by red. Domain 1 for both pH displayed some slight distance differences (closer) with Domain 2 residues specifically. The domain bounds are indicated by the middle structure with Domain 1 (pink), Domain 2 (cyan), and Domain 3 (yellow.)



pH 10.4 (A) - pH 6.5 (A)
(same chains)

pH 10.4 (A) - pH 6.5 (B)
(different chains)

Figure 4.2 Difference distance plot (ddPlot) of WT Cry11Ba between two pH. These are the differences between the distances of two chains from different pH (**a**) 10.4 or (**b**) 6.5 with their corresponding chain (A or B). The closer two residues are the distance will be marked with blue, while two residues that are further apart in distance will be marked by red. Each of the domains appear to be shrinking/condensing from one another and moving further away. This global difference was found to be caused by the change in glycerol from 30% to 40% for pH 6.5 and 10.4, respectively. The domain bounds are indicated by the middle structure with Domain 1 (pink), Domain 2 (cyan), and Domain 3 (yellow.)

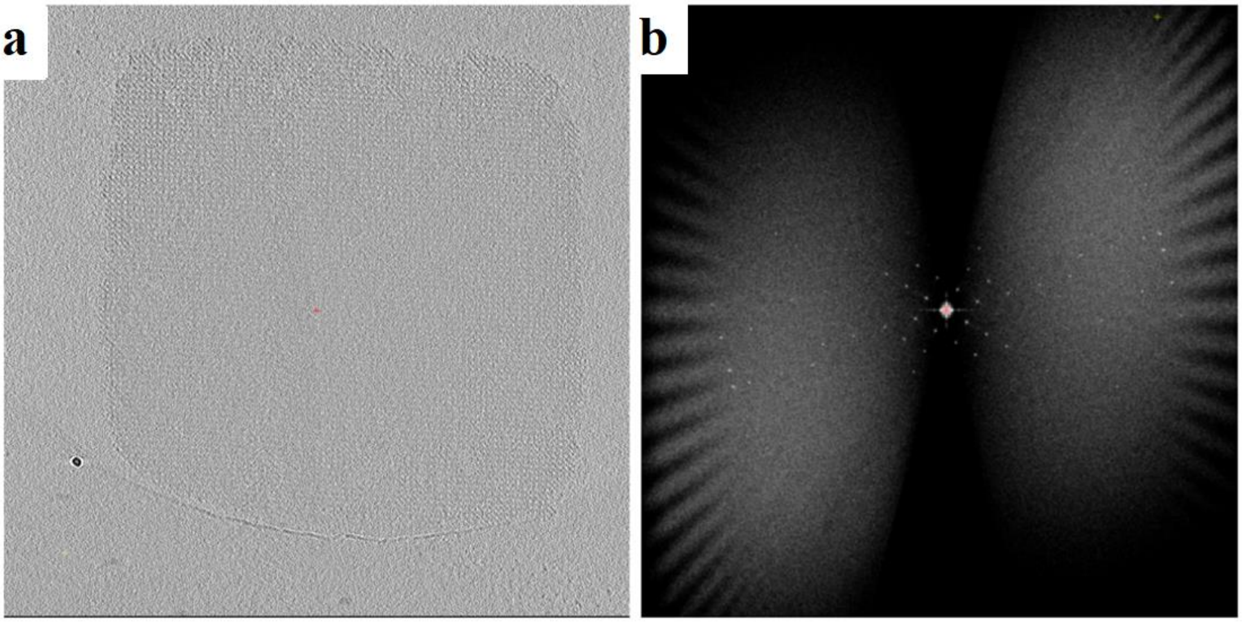


Figure 4.3 Cryo-electron tomography of WT Cry11Ba crystals. These tomograms were collected within an alkaline pH (11.3) environment with (**a**) the tomogram slice displaying the crystal at that tilt-value and its corresponding (**b**) Fourier transform pattern. This pattern displayed diffraction peaks out to approximately 3.0 Ǻ with regularity in the Bragg peaks indicative of the crystal lattice it was sampling.

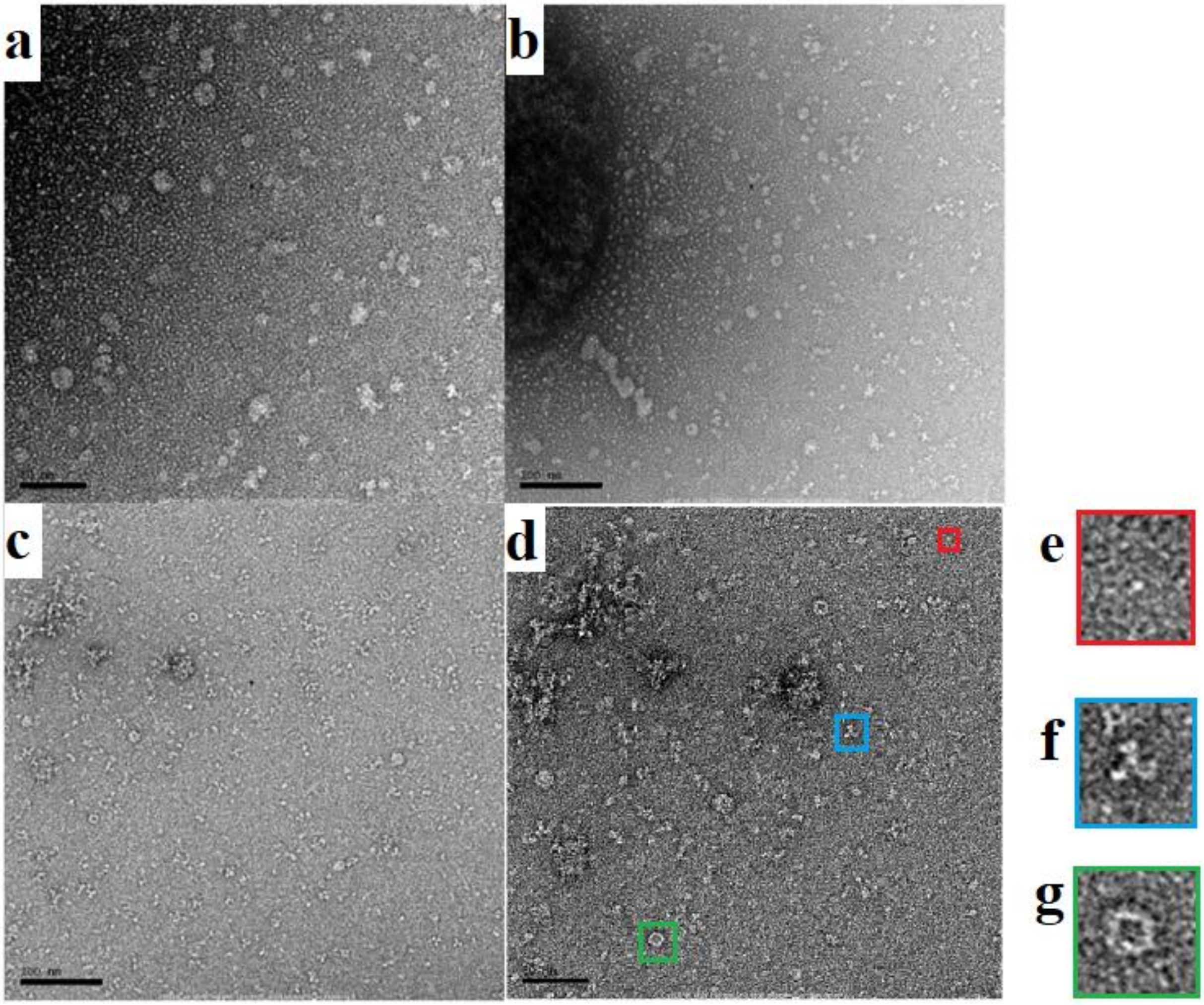


Figure 4.4 Cry11Ba crystals solubilized supernatant within pH 11.3 alkaline environment. These are from placing the supernatant after the solubility assay of WT Cry11Ba onto a grid and negatively staining with 2% uranyl acetate to improve contrast around the small particles (**a-d**) for EM. The particle shapes that were regularly observed were a (**e**) sphere, (**f**) trefoil, and (**g**) barrel highlighted in red, blue, and green boxes, respectively.

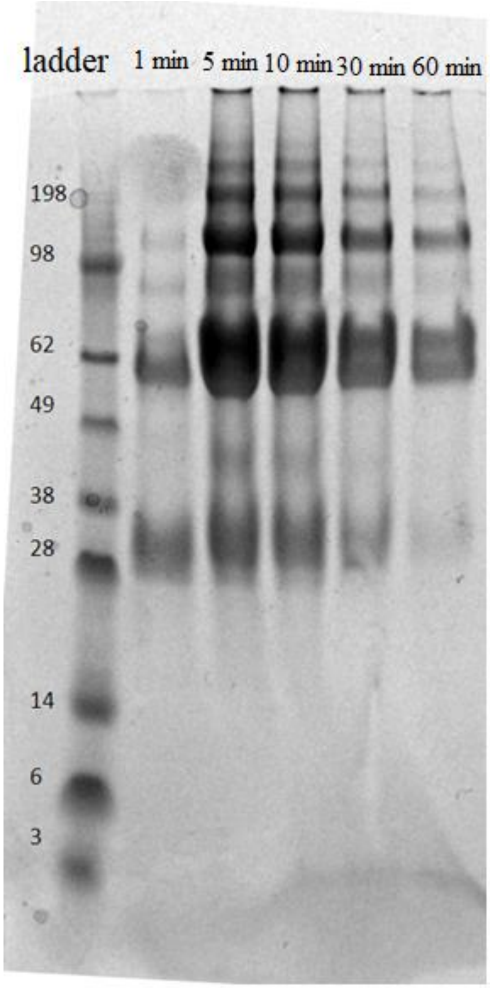


Figure 4.5 SDS-PAGE of WT Cry11Ba solubilized crystals. These are the WT Cry11Ba supernatants within alkaline environment (pH 11.3) across a time course (1-60 min). This displayed that uptake was immediate and caused by introduction to the alkaline environment with little increase in concentration after 10 min. The most drastic concentration difference was observed between 1 min to 5 min and was an ideal target for electron microscopy screening.

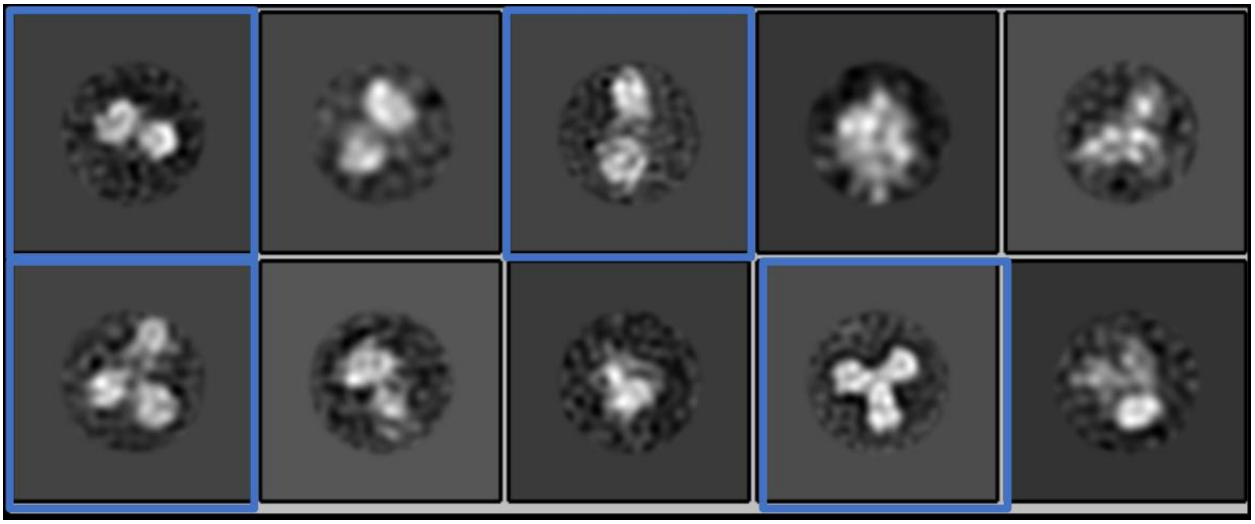


Figure 4.6 RELION WT Cry11Ba 2D classes. After selecting a total of 987 particles of solubilized Cry11Ba supernatant manually, they were classified via single-particle analysis into 10 representative 2D classes of varying resolutions. The highlighted classes 1, 3, 6, and 9 (blue box) were selected for further refinement and sampling as they had the highest amount of particles and best resolution levels.

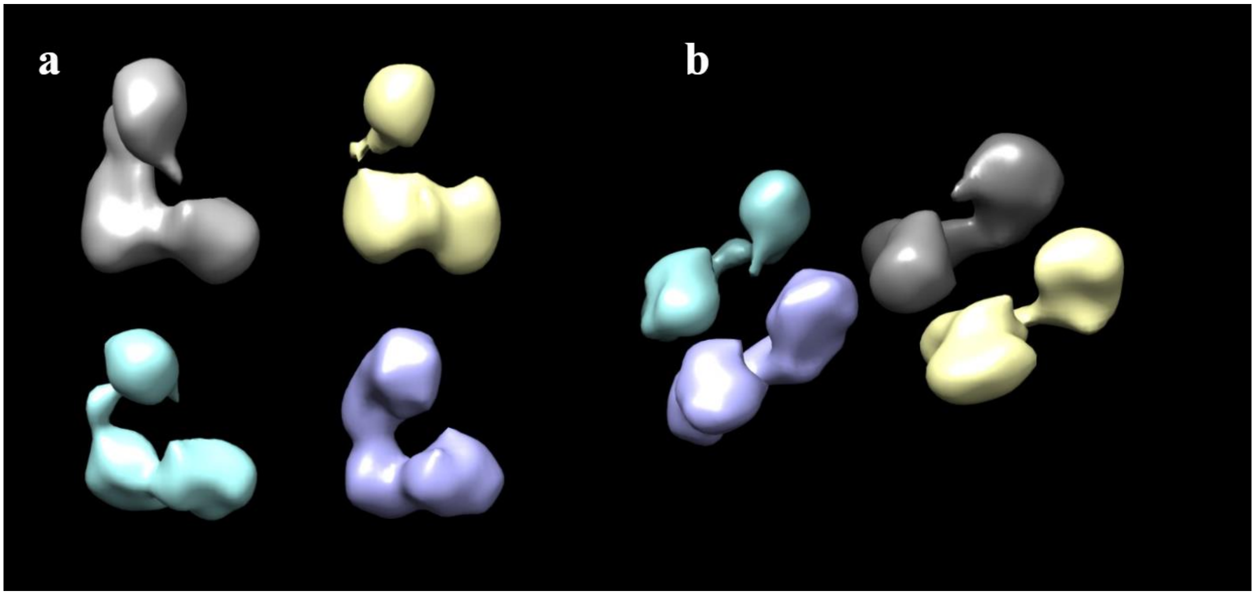


Figure 4.7 RELION WT Cry11Ba 3D *ab initio* models. Upon sampling the selected 2D classes, the WT Cry11Ba solubilized particles were processed to design these (**a**) 3D models of various levels of contour and (**b**) rotated 90°. These recapitulated the “trefoil” particles shape the best and were more stringently selected for further refinement.

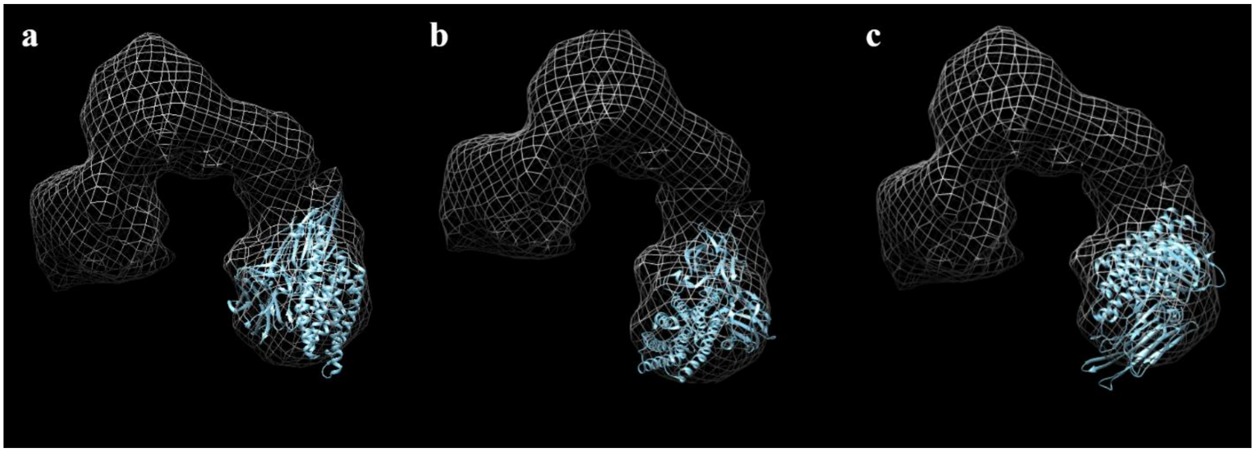


Figure 4.8 RELION *ab initio* 3D model with various poses of WT Cry11Ba. From the ab initio 3D model from RELION and the WT Cry11Ba structure, multiple poses of the Cry11Ba monomer was placed within the density and placed in various poses (**a-c**) to see which had the best fit. After multiple orientations were attempted within the density, none appeared to have a preferred pose within the 3D model's density, but 3 monomers were able to fit into the density easily with no perturbations made to the Cry11Ba monomers.

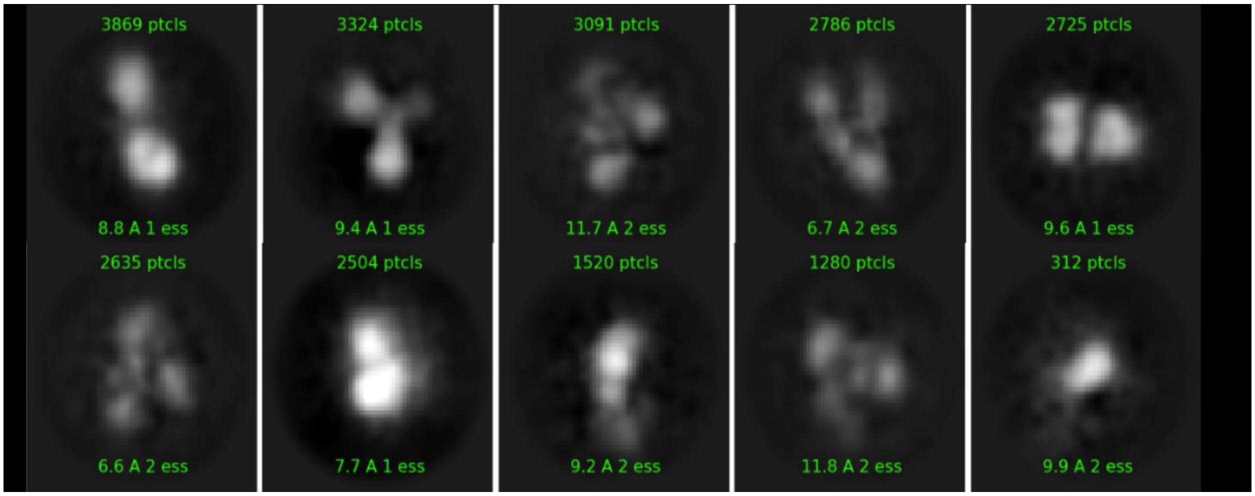


Figure 4.9 cryoSPARC WT Cry11Ba 2D classes. After selecting a total of 81,320 particles of solubilized Cry11Ba supernatant by auto-picking, they were classified via single-particle analysis into 10 representative 2D classes of varying resolutions. From these classes 21,689 particles were selected for further refinement and sampling as they had the highest amount of particles and best resolution levels.

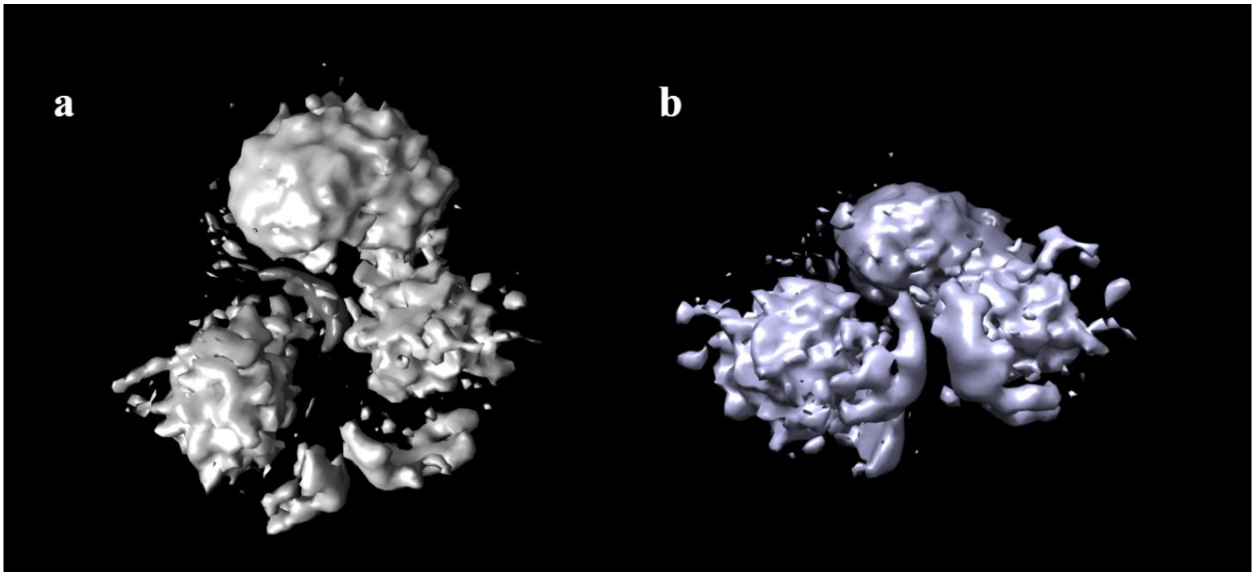


Figure 4.10 cryoSPARC WT Cry11Ba 3D *ab initio* model. Upon sampling the selected 2D classes, the WT Cry11Ba solubilized particles were processed to design a 3D model. These recapitulated the "trefoil" particles shape the best and were more stringently selected for further refinement. The density for the bottom two lobes/spheres observed (**a**) appears to be averaging together and converging, more evident when (**b)** rotated 90°. This could be a result of bias or oversampling or another particle interacting between these two lobes. To mitigate this, selecting more particles or more stringent selection of 2D classes/particles could improve the density.

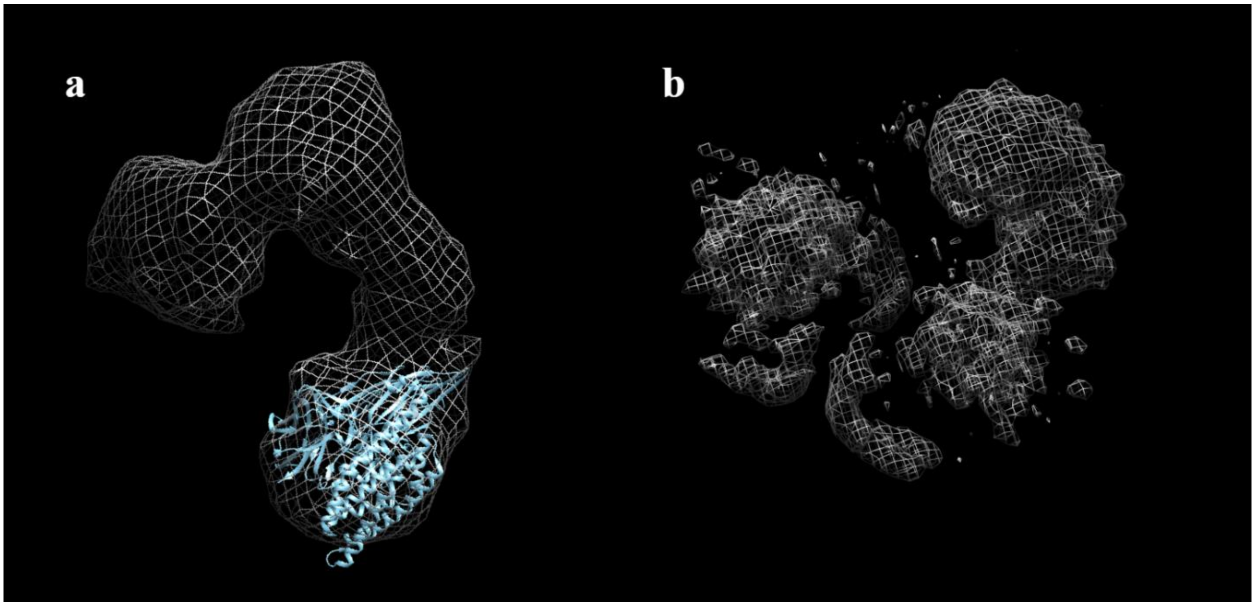


Figure 4.11 WT Cry11Ba 3D *ab initio* models by RELION and cryoSPARC. Comparing the two 3D ab initio models from (**a**) RELION and (**b**) cryoSPARC, the overall shape of the densities is quite similar, with clear trimer symmetry from the "trefoil" particles that were selected. The differences are in the amount of connecting density between the different lobes and the condensed density of each lobe from the cryoSPARC model. Despite the better estimated resolution of the cryoSPARC model, no specific pose was favored by Cry11Ba monomers, but this would insinuate the interactions between these lobes would be smaller area interfaces.

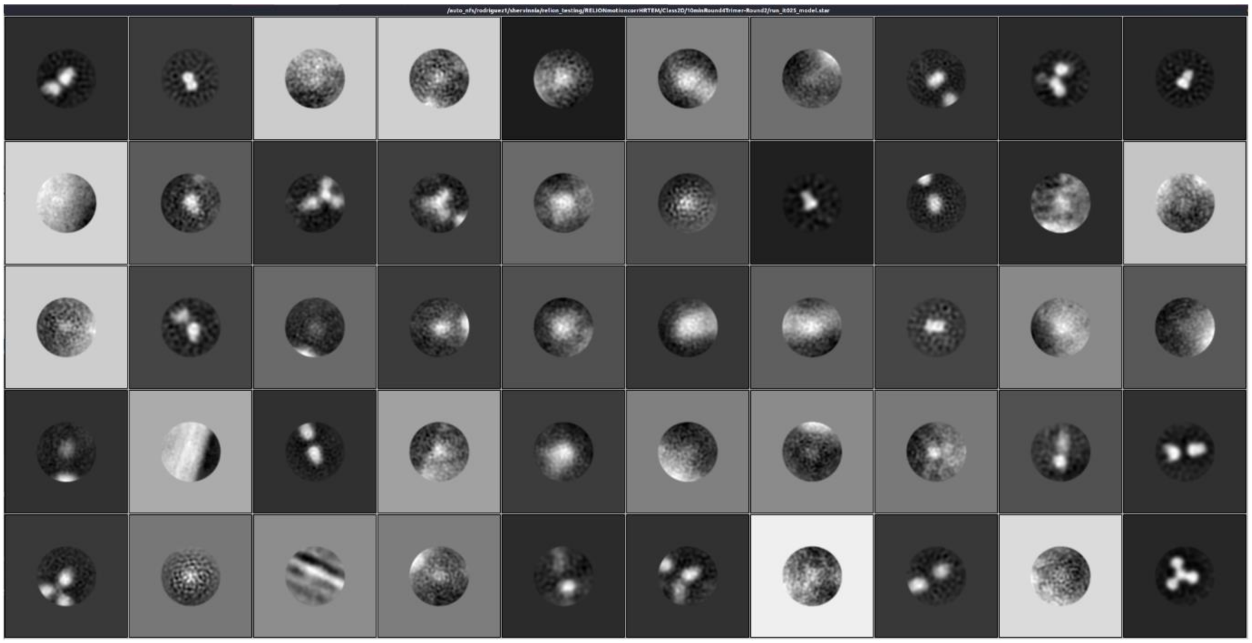
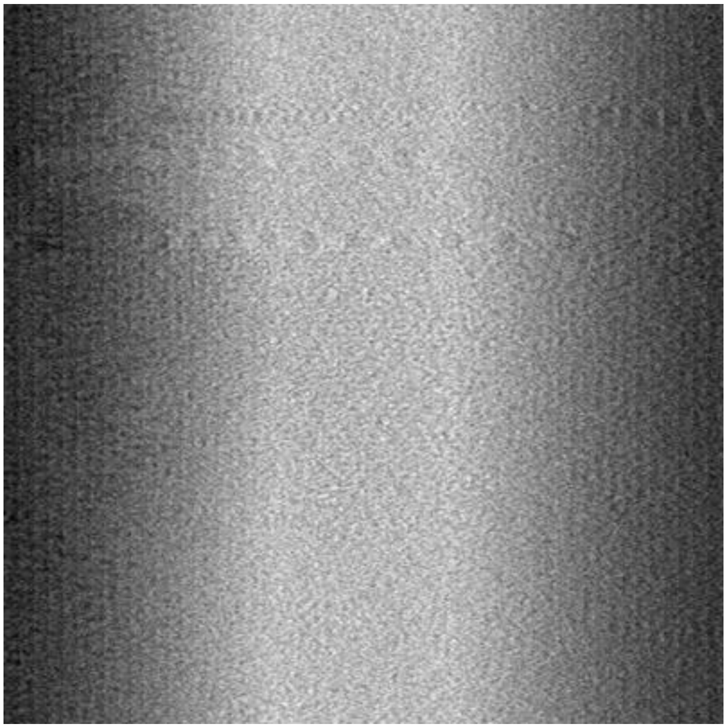


Figure 4.12 etomo and iMOD WT Cry11Ba 2D classes. Similar to RELION and cryoSPARC, the tomograms were able to be further analyzed and produce 2D classes from etomo and iMOD. While the tomograms were not able to distinguish difference within the crystal lattice, the tomograms were able to produce similar 2D classes that were produced from the separated WT Cry11Ba supernatant. The particles observed from cryoSPARC and RELION were preserved despite being separated from the dissolving crystals.

4.3 Videos



Video 4.1 Cryo-electron tomography of WT Cry11Ba crystals within alkaline pH (11.3) environment tomogram. This is the complete tilt-series for a WT Cry11Ba crystal that was reconstructed to make this tomogram. The Cry11Ba crystal can be seen throughout the tomogram and the crystal lattice of the cubic bipyramidal crystal as it samples the different planes of the crystal. The dark black dots are gold fiducials utilized to help with tracking the drift of the tomogram throughout the tilt-series.

References

(1) Tetreau, G.; Sawaya, M. R.; De Zitter, E.; Andreeva, E. A.; Banneville, A.-S.; Schibrowsky, N. A.; Coquelle, N.; Brewster, A. S.; Grünbein, M. L.; Kovacs, G. N.; Hunter, M. S.; Kloos, M.; Sierra, R. G.; Schiro, G.; Qiao, P.; Stricker, M.; Bideshi, D.; Young, I. D.; Zala, N.; Engilberge, S.; Gorel, A.; Signor, L.; Teulon, J.-M.; Hilpert, M.; Foucar, L.; Bielecki, J.; Bean, R.; de Wijn, R.; Sato, T.; Kirkwood, H.; Letrun, R.; Batyuk, A.; Snigireva, I.; Fenel, D.; Schubert, R.; Canfield, E. J.; Alba, M. M.; Laporte, F.; Després, L.; Bacia, M.; Roux, A.; Chapelle, C.; Riobé, F.; Maury, O.; Ling, W. L.; Boutet, S.; Mancuso, A.; Gutsche, I.; Girard, E.; Barends, T. R. M.; Pellequer, J.-L.; Park, H.-W.; Laganowsky, A. D.; Rodriguez, J.; Burghammer, M.; Shoeman, R. L.; Doak, R. B.; Weik, M.; Sauter, N. K.; Federici, B.; Cascio, D.; Schlichting, I.; Colletier, J.-P. De Novo Determination of Mosquitocidal Cry11Aa and Cry11Ba Structures from Naturally-Occurring Nanocrystals. *Nat Commun* **2022**, *13* (1), 4376. https://doi.org/10.1038/s41467-022-31746-x.

(2) Kochovski, Z.; Chen, G.; Yuan, J.; Lu, Y. Cryo-Electron Microscopy for the Study of Self-Assembled Poly(Ionic Liquid) Nanoparticles and Protein Supramolecular Structures. *Colloid Polym Sci* **2020**, *298* (7), 707–717. https://doi.org/10.1007/s00396-020-04657-w.

(3) Yakushevska, A. E.; Lebbink, M. N.; Geerts, W. J. C.; Spek, L.; Van Donselaar, E. G.; Jansen, K. A.; Humbel, B. M.; Post, J. A.; Verkleij, A. J.; Koster, A. J. STEM Tomography in Cell Biology. *J Struct Biol* **2007**, *159* (3), 381–391.

(4) Aoyama, K.; Takagi, T.; Hirase, A.; Miyazawa, A. STEM Tomography for Thick Biological Specimens. *Ultramicroscopy* **2008**, *109* (1), 70–80.

(5) Quentin, D.; Raunser, S. Electron Cryomicroscopy as a Powerful Tool in Biomedical Research. *J Mol Med* **2018**, *96* (6), 483–493. https://doi.org/10.1007/s00109-018-1640-y.

(6) Kirkwood, J.; Hargreaves, D.; O'Keefe, S.; Wilson, J. Analysis of Crystallization Data in the Protein Data Bank. *Acta Crystallogr F Struct Biol Commun* **2015**, *71* (10), 1228–1234.

(7) Kremer, J. R.; Mastronarde, D. N.; McIntosh, J. R. Computer Visualization of Three-Dimensional Image Data Using IMOD. *J Struct Biol* **1996**, *116* (1), 71–76. https://doi.org/https://doi.org/10.1006/jsbi.1996.0013.

(8) Xu, C.; Wang, B.-C.; Yu, Z.; Sun, M. Structural Insights into Bacillus Thuringiensis Cry, Cyt and Parasporin Toxins. *Toxins (Basel)* **2014**, *6* (9), 2732–2770.

(9) Scheres, S. H. W. RELION: Implementation of a Bayesian Approach to Cryo-EM Structure Determination. *J Struct Biol* **2012**, *180* (3), 519–530. https://doi.org/10.1016/j.jsb.2012.09.006.

(10) Punjani, A.; Rubinstein, J. L.; Fleet, D. J.; Brubaker, M. A. CryoSPARC: Algorithms for Rapid Unsupervised Cryo-EM Structure Determination. *Nat Methods* **2017**, *14* (3), 290–296. https://doi.org/10.1038/nmeth.4169.

(11) Bechtel, D. B.; Bulla, L. A. Electron Microscope Study of Sporulation and Parasporal Crystal Formation in *Bacillus Thuringiensis*. *J Bacteriol* **1976**, *127* (3), 1472–1481. https://doi.org/10.1128/jb.127.3.1472-1481.1976.

(12) Milne, J. L. S.; Subramaniam, S. Cryo-Electron Tomography of Bacteria: Progress, Challenges and Future Prospects. *Nat Rev Microbiol* **2009**, *7* (9), 666–675.

Chapter 5

Conclusion and Future Directions

5.1 Concluding Remarks

From these studies, a *de novo* structure of native Cry11Ba from *in vivo* produced crystalline inclusions at pH 6.5 and pH 10.4 were solved and led to the addition of the Cry11Ba Y350F at pH 6.5 structure. With the in-depth structural analysis, point mutations of Y241F, Y273F, Y350F, and Y453F were all selected as prime candidates for pH sensitivity, stability, and potential activity sites. These experiments have shown that each of the single point mutations were enough to cause significantly reduced stability within the crystal packing by the production of smaller, thinner crystals. In addition to the smaller sized crystals, each displayed a shift in pH sensitivity to less alkaline pHs that triggered dissolution when compared to the native Cry11Ba. The biotoxicity assay against *Culex quinquefasciatus* showed various results with Y273F displaying more toxic levels by LC_{50} values, Y350F displaying slightly more toxic levels, and Y241F & Y453F displaying the equal toxicity compared to native Cry11Ba. The Y273F and Y350F are intraface interactions compared to the Y241F and Y453F mutations that were closer to interface interactions. With Y273F being more centered at the core of the Cry11Ba chains, this could indicate the possible residue sites involved in regulating toxicity. The Cry11Ba native crystalline inclusions have been probed for the effects of introduction to an alkaline environment. These trials on the *in vivo* produced crystals displayed dissolution and Cry11Ba multimers were consistently produced and maintained within the alkaline pH. These led to single particle electron microscopy (EM) analyses to aid in structure elucidation and determination. Cry11Ba's "trefoil-trimer" class had the most success in producing structure density for orientation of each monomer with this being a potential intermediate state in the delta-endotoxin mode of action pore-forming process. Through these structural analyses, the mode of action and uptake of the visualized Cry11Ba multimers can be further investigated by other structural and biochemical techniques as these toxins' mode of action need to be more thoroughly understood within the target vector. Lastly, the *Bacillus thuringiensis* sporulation process has been recapitulated to previous schemes to display both endospore and crystalline inclusion co-maturation, along with additional details that were not previously observed, including microcrystal formation/nucleation early in sporulation, potential membrane formation throughout sporulation, and endospores being potentially bound to the cell membranes upon natural lysis. The investigation of one such protein that may be causing the membrane formation, Bt152, did show coalescence to the Cry11Ba crystalline inclusion by fluorescence microscopy and a possible function for this otherwise uncharacterized paratoxin-interacting protein. The collection of these findings is simply the beginning to understanding Cry11Ba crystalline inclusions as a paratoxin at the macromolecular level, mode of activity, and *in vivo* self-assembly. Each of these can be used as possible models for other crystalline and cytolytic toxins that play vital roles as pesticides and in further understanding how the crystallization is occurring mechanically by using the suite of EM techniques established and currently being developed. The possibilities are endless if this *Bacillus thuringiensis* system is fully investigated and can unlock the optimization for future exploitations for crystallography research and/or medical applications.

Appendix A

Supplementary Information for Chapter 2: Serial femtosecond structural determination and biochemical analysis of *in vivo* Cry11Ba crystalline inclusions

The work described in this section has been reproduced from:

Tetreau, G.‡; Sawaya, M. R.‡; De Zitter, E.‡; Andreeva, E. A.§; Banneville, A.-S.§; Schibrowsky, N. A. §; Coquelle, N.; Brewster, A. S.; Grünbein, M. L.; Kovacs, G. N.; Hunter, M. S.; Kloos, M.; Sierra, R. G.; Schiro, G.; Qiao, P.; Stricker, M.; Bideshi, D.; Young, I. D.; Zala, N.; Engilberge, S.; Gorel, A.; Signor, L.; Teulon, J.-M.; Hilpert, M.; Foucar, L.; Bielecki, J.; Bean, R.; de Wijn, R.; Sato, T.; Kirkwood, H.; Letrun, R.; Batyuk, A.; Snigireva, I.; Fenel, D.; Schubert, R.; Canfield, E. J.; Alba, M. M.; Laporte, F.; Després, L.; Bacia, M.; Roux, A.; Chapelle, C.; Riobé, F.; Maury, O.; Ling, W. L.; Boutet, S.; Mancuso, A.; Gutsche, I.; Girard, E.; Barends, T. R. M.; Pellequer, J.-L.; Park, H.-W.; Laganowsky, A. D.; Rodriguez, J.; Burghammer, M.; Shoeman, R. L.; Doak, R. B.; Weik, M.; Sauter, N. K.; Federici, B.; Cascio, D.; Schlichting, I.; Colletier, J.-P. De Novo Determination of Mosquitocidal Cry11Aa and Cry11Ba Structures from Naturally-Occurring Nanocrystals. *Nat Commun* **2022**, *13* (1), 4376. https://doi.org/10.1038/s41467-022-31746-x.

‡ and § Denotes equal contributions

Guillaume Tetreau, Michael R. Sawaya, Elke De Zitter, Elena A. Andreeva, Anne-Sophie Banneville, Natalie A. Schibrowsky, Nicolas Coquelle, Aaron S. Brewster, Marie Luise Grünbein, Gabriela Nass Kovacs, Mark S. Hunter, Marco Kloos, Raymond G. Sierra, Giorgio Schiro, Pei Qiao, Myriam Stricker, Dennis Bideshi, Iris D. Young, Ninon Zala, Sylvain Engilberge, Alexander Gorel, Luca Signor, Jean-Marie Teulon, Mario Hilpert, Lutz Foucar, Johan Bielecki, Richard Bean, Raphael de Wijn, Tokushi Sato, Henry Kirkwood, Romain Letrun, Alexander Batyuk, Irina Snigireva, Daphna Fenel, Robin Schubert, Ethan J. Canfield, Mario M. Alba, Frédéric Laporte, Laurence Després, Maria Bacia, Amandine Roux, Christian Chapelle, François Riobé, Olivier Maury, Wai Li Ling, Sébastien Boutet, Adrian Mancuso, Irina Gutsche, Eric Girard, Thomas R. M. Barends, Jean-Luc Pellequer, Hyun-Woo Park, Arthur D. Laganowsky, Jose Rodriguez, Manfred Burghammer, Robert L. Shoeman, R. Bruce Doak, Martin Weik, Nicholas K. Sauter, Brian Federici, Duilio Cascio, Ilme Schlichting, Jacques-Philippe Colletier

Supplementary Information for:

De novo determination of mosquitocidal Cry11Aa and Cry11Ba structures from naturally-occurring nanocrystals

by

Guillaume Tetreau, Michael R. Sawaya, Elke De Zitter, Elena A. Andreeva, Anne-Sophie Banneville, Natalie A. Schibrowsky, Nicolas Coquelle, Aaron S. Brewster, Marie Luise Grünbein, Gabriela Nass Kovacs, Mark S. Hunter, Marco Kloos, Raymond G. Sierra, Giorgio Schiro, Pei Qiao, Myriam Stricker, Dennis Bideshi, Iris D. Young, Ninon Zala, Sylvain Engilberge, Alexander Gorel, Luca Signor, Jean-Marie Teulon, Mario Hilpert, Lutz Foucar, Johan Bielecki, Richard Bean, Raphael de Wijn, Tokushi Sato, Henry Kirkwood, Romain Letrun, Alexander Batyuk, Irina Snigireva, Daphna Fenel, Robin Schubert, Ethan J. Canfield, Mario M. Alba, Frédéric Laporte, Laurence Després, Maria Bacia, Amandine Roux, Christian Chapelle, François Riobé, Olivier Maury, Wai Li Ling, Sébastien Boutet, Adrian Mancuso, Irina Gutsche, Eric Girard, Thomas R. M. Barends, Jean-Luc Pellequer, Hyun-Woo Park, Arthur D. Laganowsky, Jose Rodriguez, Manfred Burghammer, Robert L. Shoeman, R. Bruce Doak, Martin Weik, Nicholas K. Sauter, Brian Federici, Duilio Cascio, Ilme Schlichting, Jacques-Philippe Colletier.

This PDF file contains:

- Supplementary Figures S1-S19
- Supplementary Tables S1-S4

Supplementary Fig. 1. Cry11Aa and Cry11Ba exhibit high similarity but limited sequence identity. Cry11Aa and Cry11Ba were aligned using strap [97]. This shows that 54.0% (392 residues) of Cry11Aa and Cry11Ba residues are identical. Domain I is shown in blue; domain II is shown in orange except for the $\alpha_h\beta_h$-handle and β_{pin} which are shown in purple and red, respectively; domain III is shown in pink.

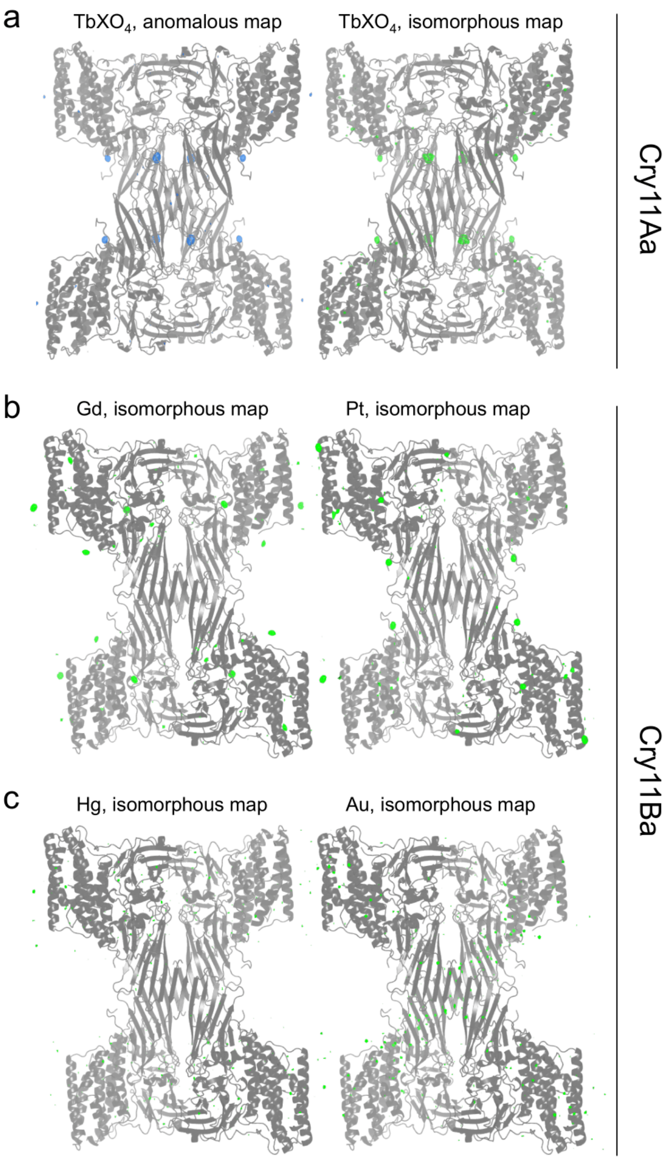
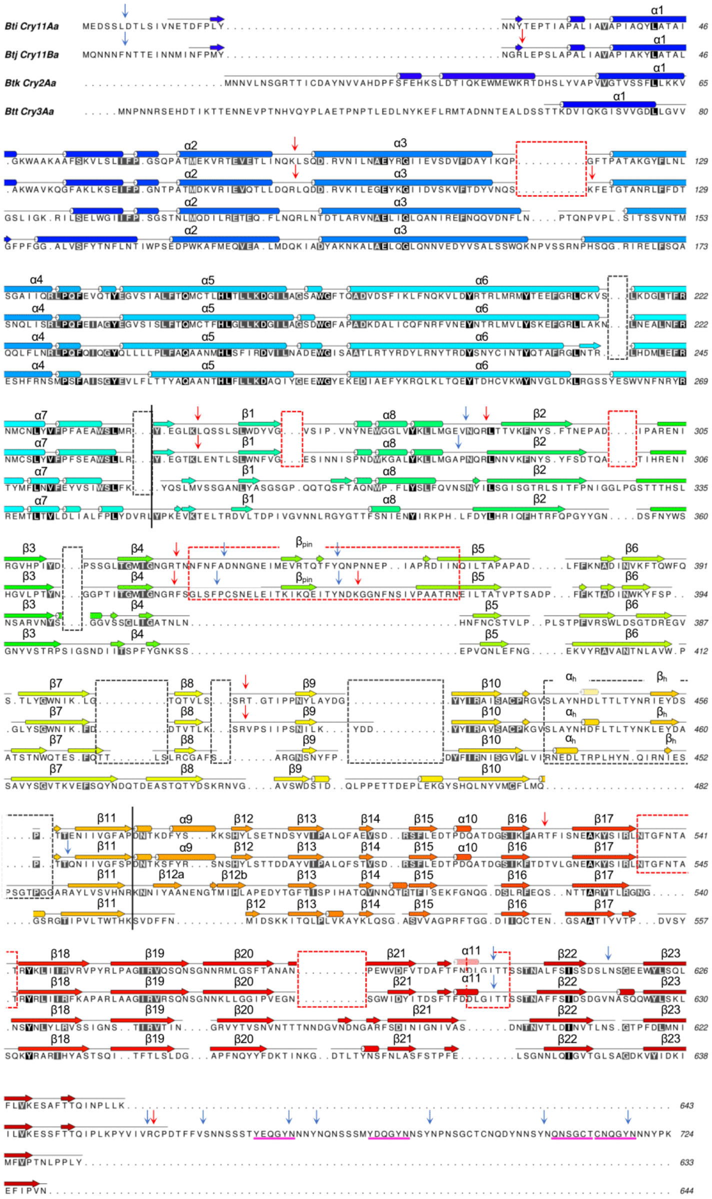
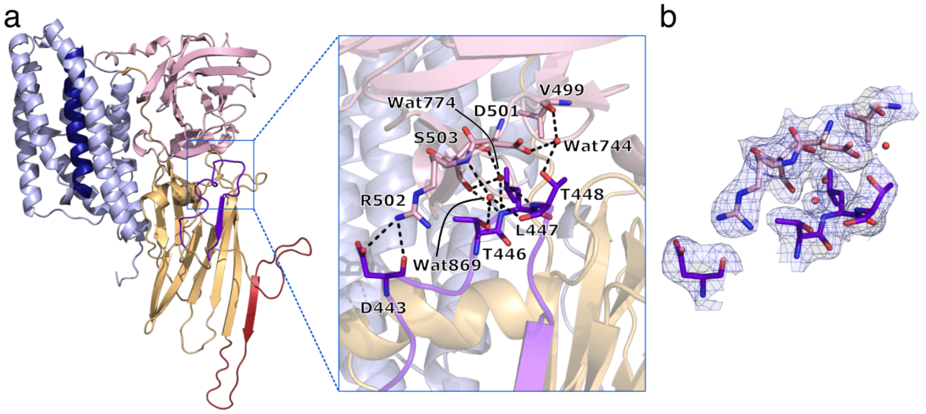


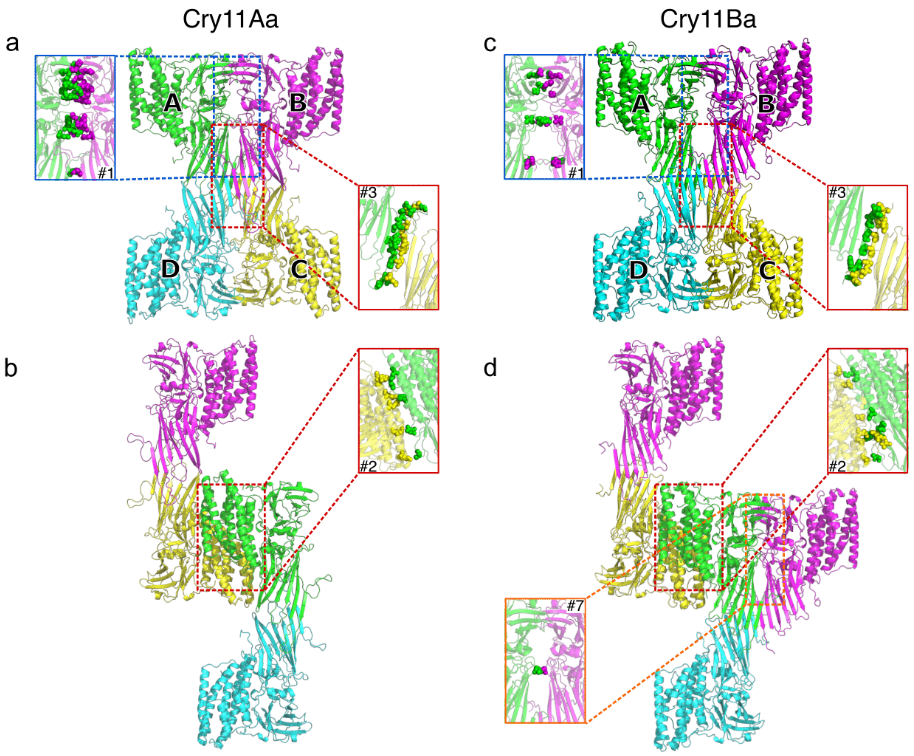
Fig. 2. Heavy atom locations after soaking of the Cry11 crystals. a, Location of Tb-Xo4 in Cry11Aa is clearly indicated the anomalous (left; highest peak at 33.1 σ) and isomorphous difference maps (right; highest peak at 24.5 σ), both contoured at ± 4 σ. **b-c**, Native Cry11Ba structure revealed a posteriori that soaking of Cry11Ba with salts of gadolinium and platinum (**b**) led to their successful binding to the crystalline Cry11Ba, as revealed by isomorphous peaks > 9.3 and 7.5 σ, respectively. However, no anomalous signal was detected (highest peaks at 5.5 and 5.0 σ, respectively). Soaking of Cry11Ba crystals with salts mercury or gold was fully infructuous, with significant peaks visible neither in the isomorphous (highest peaks at 5.3 and 5.5 σ, respectively) nor the anomalous (highest peaks at 4.9 and 5.5 σ, respectively) maps. In (**b**) and (**c**), the isomorphous difference maps are contoured at ± 4 σ.



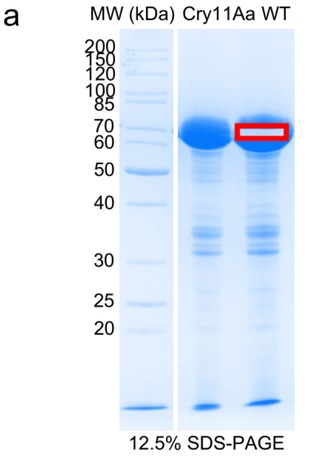
Supplementary Fig. 3. Secondary structure assignment of Cry11Aa, Cry11Ba, Cry2Aa and Cry3Aa. Secondary structures were assigned using DSSP[103] and colored according to sequence (from blue to red). α-helices and β-strand are shown by rods and arrows, respectively. Vertical black lines show the domain borders. Remarkable regions of difference between the Cry11 toxins and all other Cry toxins are indicated by red dashed boxes, whereas black boxes indicate differences between Cry11 and Cry2Aa toxin as compared to other toxins. Vertical red and blue arrows indicate trypsin and proteinase K digestion sites, respectively. Regions predicted to form short adhesive motifs of the Low Complexity, Amyloid-like Reversible Kinked Segments (LARKS) type are underlined in magenta.



Supplementary Fig. 4. Interactions by the $\alpha_h\beta_h$-handle in Cry11Aa. Color code as in Fig. 2. **a**, Dashes indicate hydrogen bonds (up to 3.2 Å) and the salt bridge between D443 and R502. **b**, $2F_{obs}$-DF_{calc} electron density map for the residues and water molecules shown in (**a**), contoured at 1 σ. Water molecule 744 is defined in the map with a contour level of 0.9 σ but not at a level of 1 σ.

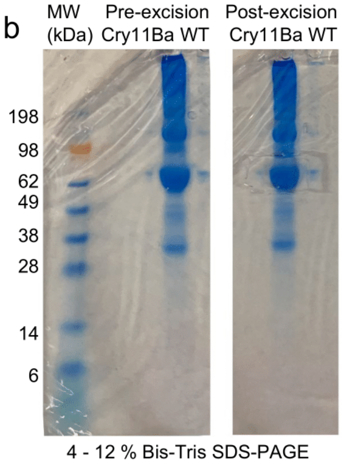


Supplementary Fig. 5. Hydrogen bonds and salt bridges between monomer in Cry11Aa and Cry11Ba crystals. Colour code as in Fig. 3. **a**, Cry11Aa tetramer with zoom on each of the interfaces identified by PISA that contain hydrogen bonds or salt bridges identified by PISA (interface #1 and #3), with the residues involved in these interactions depicted as spheres. **b**, In the Cry11Aa crystal assembly between neighbouring tetramers only interface #2 contains hydrogen and salt bridges. These are visualized as in (**a**). **c**, Cry11Ba tetramer with zoom on the interfaces that contain salt bridges and hydrogen bonds as in (**a**). **d**, Cry11Ba crystal assembly between neighbouring tetramers, visualized as in (**b**).



Database: SwissProt
Protein View: C11AA_BACTI (P21256)
Score (Mascot): 191
Protein coverage: 45% (matched peptides in bold red)

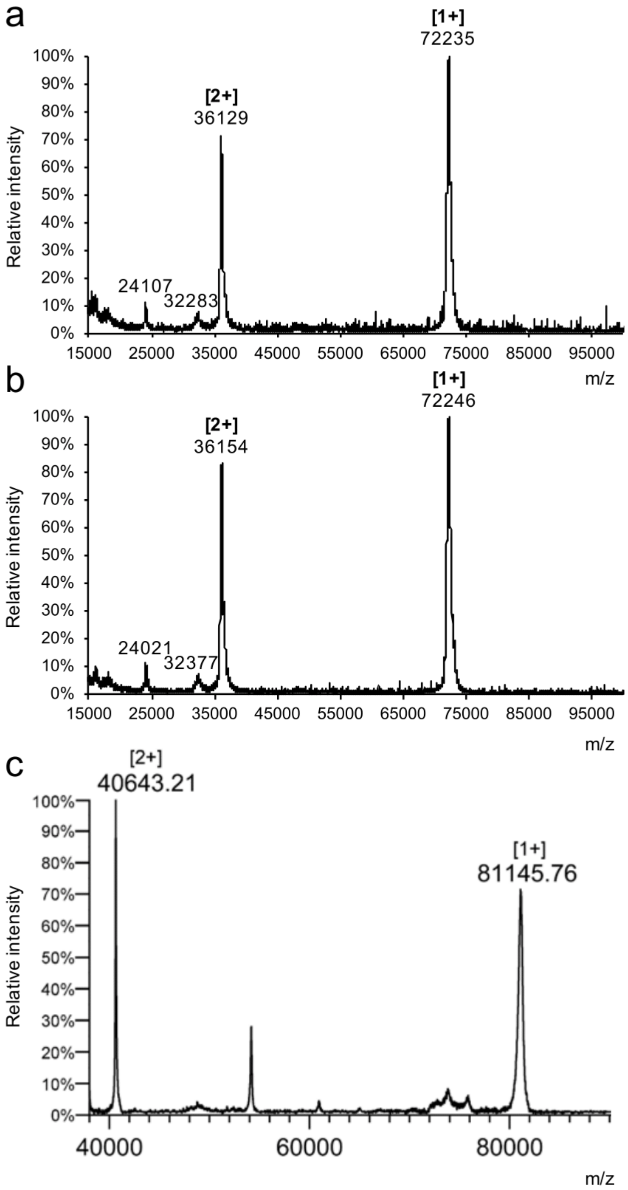
1 MEDSSLDTLS IVNETDFPLY NNYTEPTIAP ALIAVAPIAQ YLATAIGKWA
51 AKAAFSK**VLS LIFPGSQPAT MEK**VRTEVET LINQKLSQDR **VNILNAEYRG**
101 **IIEVSDVFDA YIK**QPGFTPA TAK**GYFLNLS GAIIQR**LPQF EVQTYEGVSI
151 ALFTQMCTLH LTLLK**DGILA GSAWGFTQAD VDSFIK**LFNQ KVLDYRTRLM
201 R**MYTEEFGR**L CKVSLK**DGLT FRNMCNLYVF PFAEAWSLMR** YEGLKLQSSL
251 SLWDYVGVSI PVNYNEWGGL VYK**LLMGEVN QR**LTTVK**FNY SFTNEPADIP**
301 **AR**ENIRGVHP IYDPSSGLTG WIGNGRTNNF NFADNNGNEI MEVR**TQTFYQ**
351 **NPNNEPIAPR DIINQILTAP APADLFFK**NA DINVK**FTQWF QSTLYGWNIK**
401 LGTQTVLSSR **TGTIPPNYLA YDGYYIR**AIS ACPR**GVSLAY NHDLTTLTYN**
451 **RIEYDSPTTE NIIVGFAPDN TK**DFYSK**KSH YLSETNDSYV IPALQFAEVS**
501 **DRSFLEDTPD QATDGSIK**FA RTFISNEAKY SIR**LNTGFNT ATR**YKLIIRV
551 RVPYR**LPAGI R**VQSQNSGNN RMLGSFTANA NPEWVDFVTD AFTFNDLGIT
601 TSSTNALFSI SSDSLNSGEE WYLSQLFLVK ESAFTTQINP LLK



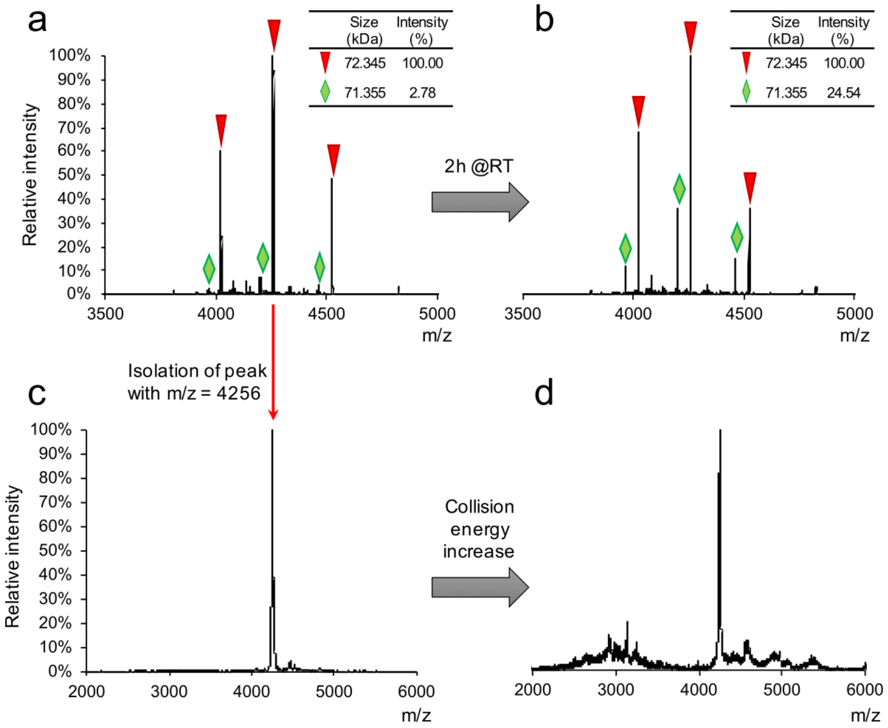
Database: SwissProt
Protein View: C11BA_BACTJ (Q45730)
Score (Mascot): 916.64
Protein coverage: 74% (matched peptides in bold red)

1 MQNNNFNTTE INNMINFPMY **NGRLEPSLAP ALIAVAPIAK YLATALAKWA**
51 **VKQGFAKLKS EIFPGNTPAT MDKVRIEVQT LLDQRLQDDR VKILEGEYKG**
101 **IIDVSKVFTD YVNQSKFETG TANRLFFDTS NQLISR**LPQF EIAGYEGVSI
151 SLFTQMCTFH **LGLLKDGILA GSDWGFAPAD KDALICQFNR FVNEYNTRLM**
201 **VLYSKEFGRL LAKNLNEALN FRNMCSLYVF PFSEAWSLLR YEGTKLENTL**
251 **SLWNFVGESI NNISPNDWKG ALYKLLMGAP NQRLNNVKFN YSYFSDTQAT**
301 **IHRENIHGVL PTYNGGPTIT GWIGNGRFSG LSFPCSNELE ITKIKQEITY**
351 **NDKGGNFNSI VPAATRNEIL TATVPTSADP FFKTADINWK YFSPGLYSGW**
401 **NIKFDDTVTL KSRVPSIIPS NILKYDDYYI RAVSACPKGV SLAYNHDFLT**
451 **LTYNKLEYDA PTTQNIIVGF SPDNTK**SFYR **SNSHYLSTTD DAYVIPALQF**
501 **STVSDRSFLE DTPDQATDGS IKFTDTVLGN EAKYSIRLNT GFNTATRYRL**
551 **IIRFKAPARL AAGIRVRSQN SGNNK**LLGGI PVEGNSGWID YITDSFTFDD
601 LGITTSSTNA FFSIDSDGVN **ASQQWYLSKL ILVKESSFTT QIPLKPYVIV**
651 **R**CPDTFFVSN NSSSTYEQGY NNNYNQNSSS MYDQGYNNSY NPNSGCTCNQ
701 DYNNSYNQNS GCTCNQGYNN NYPK

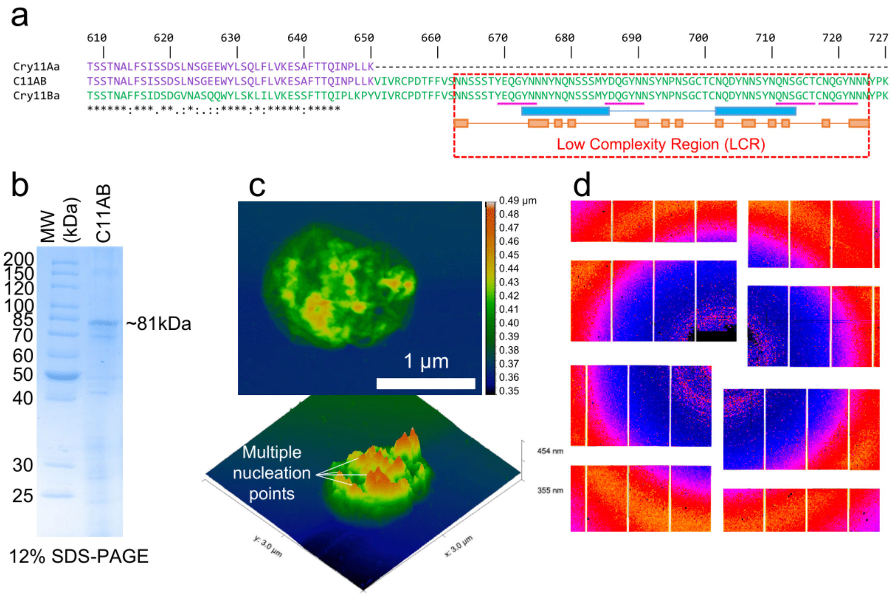
Supplementary Fig. 6. SDS PAGE and mass spectrometry confirm that Cry11Aa and Cry11Ba crystals are solely composed of these proteins. a, The most abundant band present at ~70 kDa from the proteomic profile of Cry11Aa purified crystal suspension was cut and digested with protease before being analyzed in MALDI. Analysis using Mascot Software allowed matching the majority of peptides identified, covering 45 of the Cry11Aa sequence (indicated in red). **b,** The most abundant band present at ~ 80 kDa from the SDS PAGE gel of the was digested with trypsin before being analyzed by LC-MS/MS. Analysis using Mascot Software allowed matching the majority of peptides identified, covering 74% of the Cry11Ba sequence (indicated in red). The MALDI experiment on Cry11Aa was performed once, and that on Cry11Ba was performed twice. Proteomic profiling by SDS-PAGE was performed more than ten times on both samples.



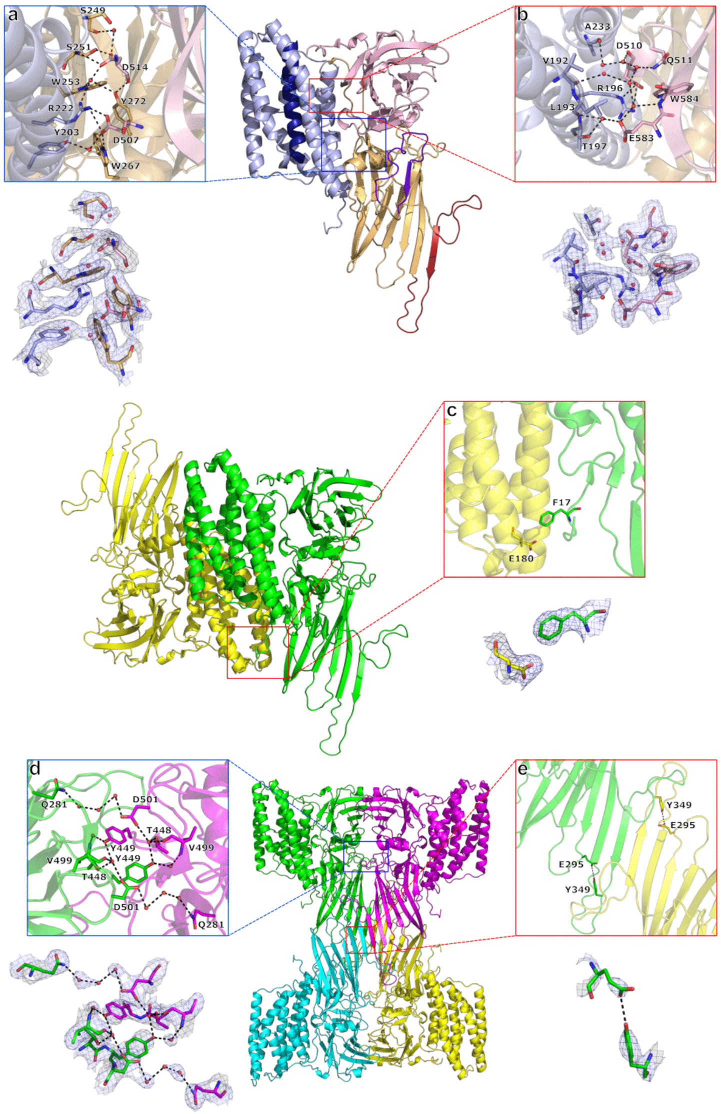
Supplementary Fig. 7. MALDI-ToF analysis confirms that Cry11Aa and Cry11Ba are present as a full-size monomer in the crystal. a-b, Cry11Aa crystals mixed with SA matrix in absence (**a**) or presence of DTT (**b**) showed the same profile with the most abundant peaks corresponding to a monomer of 72,235–72,246 kDa (expected mass: 72,349 kDa) monocharged or bicharged. **c**, Cry11Ba crystals mixed with DHAP showed the presence of full-length monomer of 81,145 kDa in agreement with the predicted mass of 81,344 kDa.



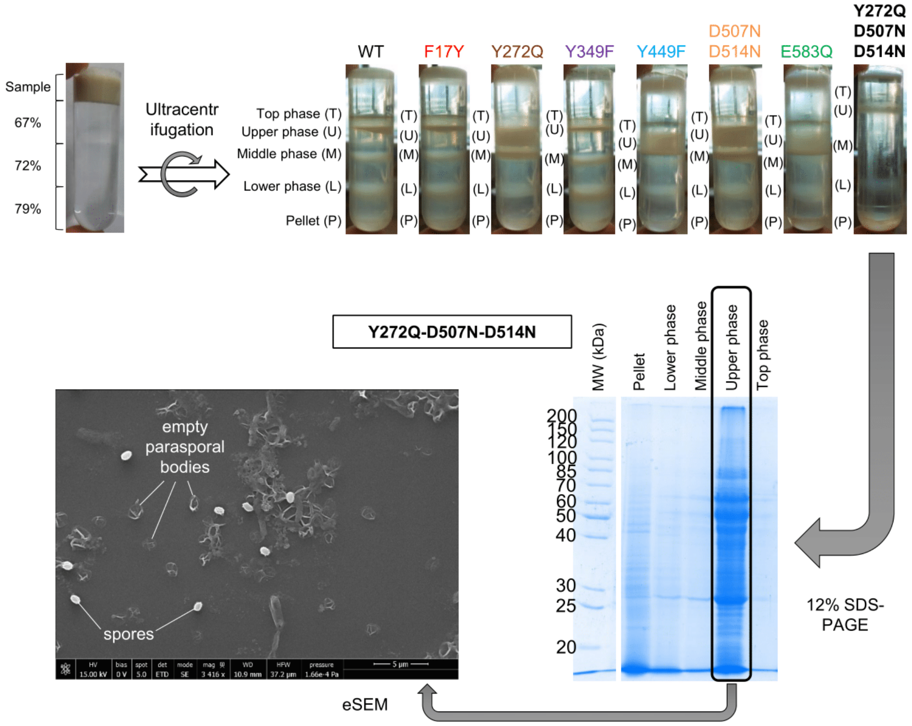
Supplementary Fig. 8. Native mass spectrometry confirms that Cry11Aa is solubilized as a full-size monomer. a, profile of soluble Cry11Aa in native MS shows three peaks corresponding to a full size toxin of 72.345 kDa (triangles, expected mass: 72.349 kDa) and three peaks for a ~1 kDa smaller form much less abundant (diamonds), likely resulting from a cleavage of the first 9 amino acids in N-terminal part of the toxin. **b**, incubation at room temperature leads to an increase in the smaller form, reaching one fourth of abundance of full size one, suggesting a targeted cleavage of few amino acids in the N- and/or C-terminal extremity. **c**, when isolating the most abundant peak (*i.e.,* at m/z = 4256, corresponding to Cry11Aa with a charge of 17) in the same condition as in panel a, the peak is sharp and well defined. **d**, increasing the collision energy leads to a fragmentation into species with different sizes. If the protein is an oligomer with multiple monomers through non-covalent bound, specific monomers would be easily separated and identified on the MS spectrum. This rather supports that Cry11Aa solubilizes as a full-sized monomer prone to degradation under certain conditions.



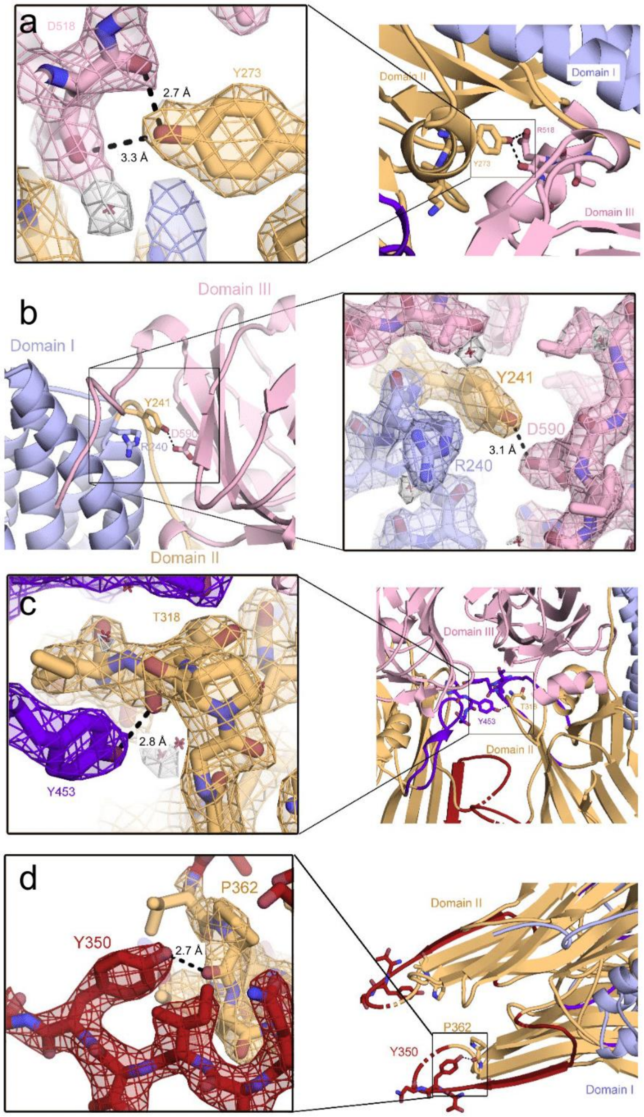
Supplementary Fig. 9. Addition of the Cry11Ba C-terminal Low Complexity Region (LCR) to Cry11Aa does not improve its crystallization. a, amino acid sequence alignment of Cry11Aa, Cry11Ba and Cry11Aa fused with the last 77 residues of Cry11Ba (named C11AB). Residues originating from Cry11Aa and Cry11Ba are indicated in purple and green, respectively, Blue and orange rectangles represent LCR domains identified by SEG and CAST programs, respectively. Segments predicted to form larks are underlined in magenta. **b**, C11AB crystals were purified using sucrose gradients and verified on 12% SDS-PAGE gels yielding a major band at a mean of 81.5 kDa, which matches the expected size for the fusion, although expressed at a lower level than Cry11Aa WT. **c**, purified inclusions analyzed by AFM revealed multiple nucleation points instead of a unique microcrystal. The experiment was performed once. **d**, maximum projection of the few diffraction patterns that were obtained using the C11AB inclusions at the EuXFEL indicating only limited diffraction.



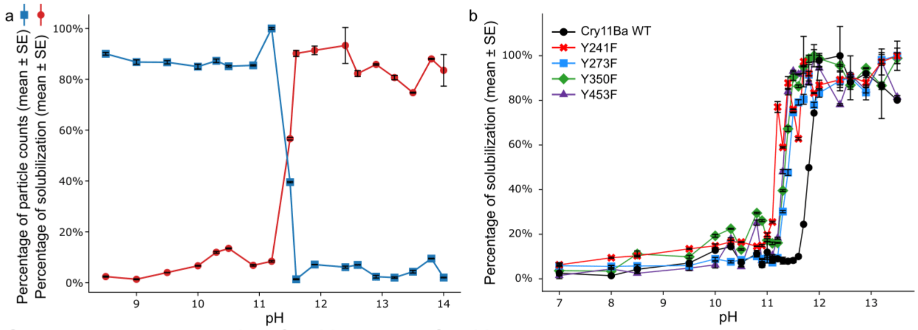
Supplementary Fig. 10. Cry11Aa mutation strategy. a, Hydrogen and salt bonding interaction network between domain I, II and III within a single monomer formed by Y272, D507 and D514. Color code as in Fig. 2. For clarity, a part of domain III is not shown in cartoon mode. Hydrogen bonds are shown up until a distance of 3.2 Å. The $2F_{obs}$-F_{calc} electron density map contouring the concerned residues and water molecules at 1 σ is shown at the bottom of each panel. **b**, Hydrogen and salt bonding interaction network between domain I and III within a single monomer involving E583. Color code as in Fig. 2. **c**, Location of F17 and E180 in interface #2 between two adjacent monomers. Monomers are color coded as in Fig. 3. **d**, Water-mediated hydrogen bonding network around Y449 in interface #1 between two adjacent monomers. For clarity, domain III of each of the monomer is omitted from the figure. Color code as in Fig. 3. **e**, Interaction between E295 and Y349 in the β_{pin} region which connects two adjacent monomers via interface #3. Color code as in Fig. 3.



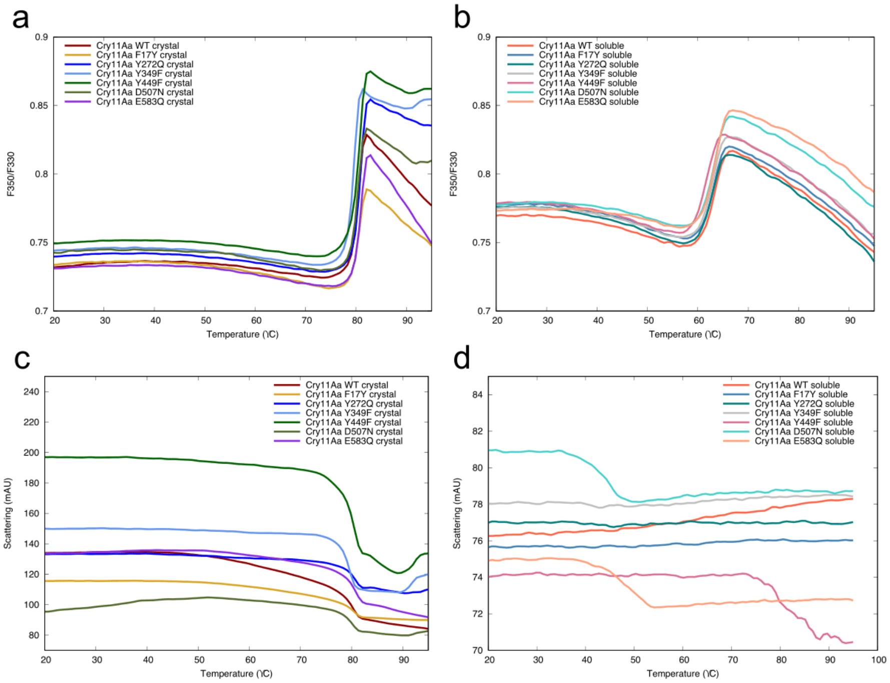
Supplementary Fig. 11. The triple mutation Y272Q-D507N-D514N of Cry11Aa affects crystal formation. All Cry11Aa mutants with one and two point-mutations (D507N-D514N) produced crystals that could be properly purified by sucrose gradient coupled with ultracentrifugation. In contrast, the triple mutant Y272Q-D507N-D514N showed a different purification profile. Most of proteins were contained in the upper phase of the sucrose gradient but no clear band at the ~70 kDa size could be seen. This is confirmed by environmental SEM (eSEM) experiments showing that this phase mostly contains empty parasporal bodies, suggesting that this mutation is deleterious for the proper production and/or crystallization of the toxin in the bacterium during its sporulation. The eSEM imaging of the Y272Q-D507N-D514N mutant was performed once.



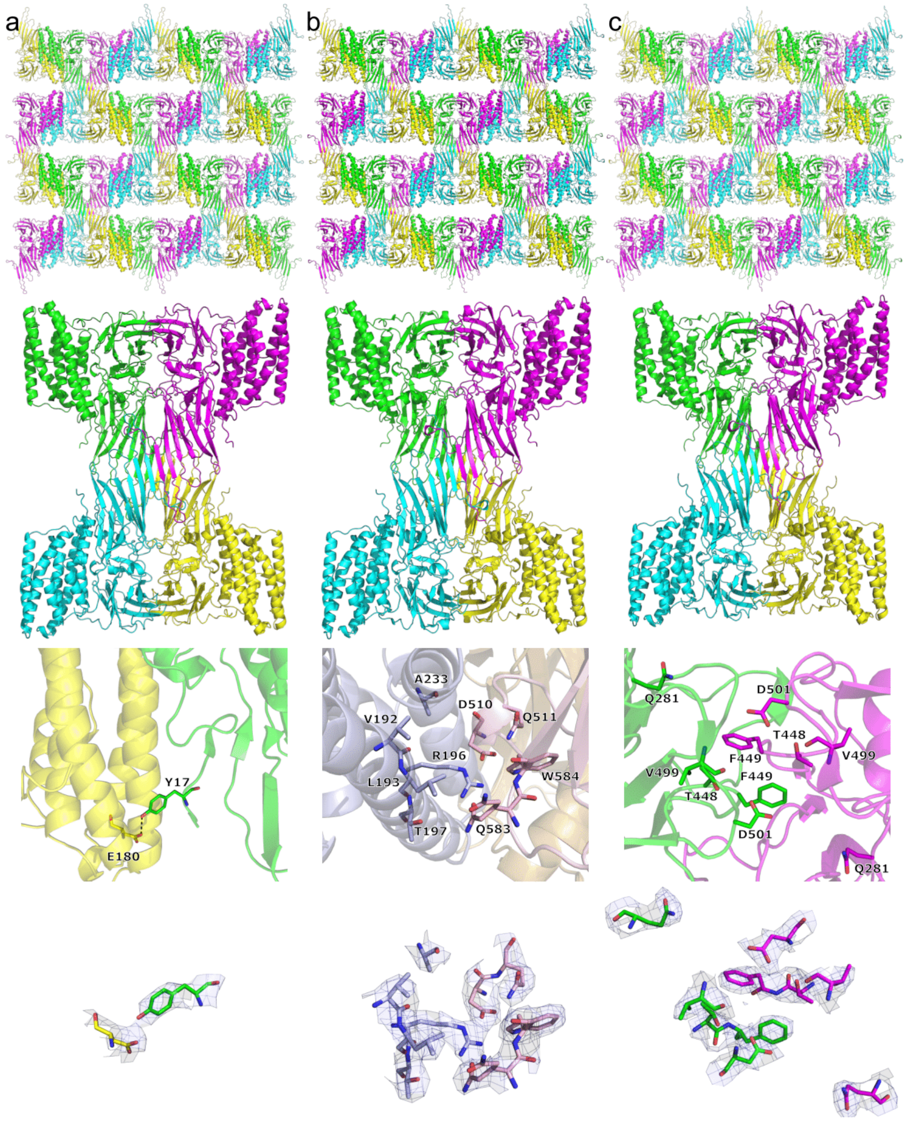
Supplementary Fig. 12. Cry11Ba structure with close-ups of electron density at selected mutation sites. Color code as in Fig. 2. **a**, The Y273F mutation was selected to disrupt the H-bonding of the Y273 to the R518 backbone carbonyl (~2.6 Å) and the intermittent H-bond of D517 side chain carbonyl (~3.3 Å) on the same chain between domain interfaces. **b**, The Y241F mutation causes the disruption of the H-bond (2.8 Å) between the Y241 hydroxyl and D590 carbonyl group, which is between two domain interfaces. **c**, The Y453F mutation caused a loss of the Y453's hydroxyl group H-bond to T318's carbonyl backbone of a different chain, which is located at an interface. **d**, Y350 is conserved in Cry11Ba where it H-bonds to P362(O), and the equivalent Y350F mutant solubilized at a lower pH.



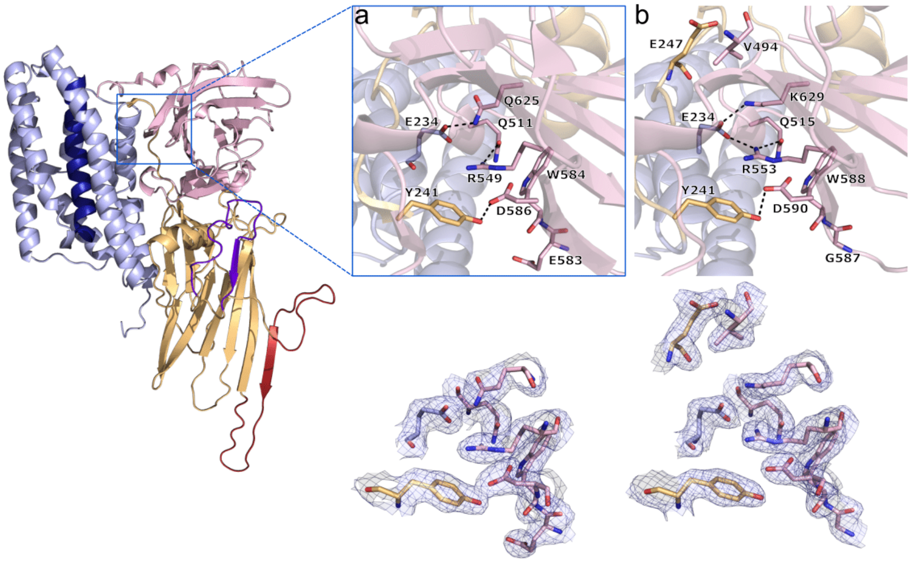
Supplementary Fig. 13. Cry11Ba and Cry11Ba mutants crystal solubilization and stabilization in function of pH. a, Solubility (red circles) and turbidity (blue squares) assay on Cry11Ba crystals indicate that 50 % of crystals solubilize at pH ~ 11.9 (n=3 independent measurements, data are presented as mean values +/- SEM). **b**, Solubility of Cry11Ba WT (black circles) and mutants (Y241F: red crosses, Y273F: blue squares, Y350F: green diamonds, Y453F: purple triangles) in function of pH show that the mutants solubilize at lower pH (~ 11.3) than the WT(n=3 independent measurements, data are presented as mean values +/- SEM).



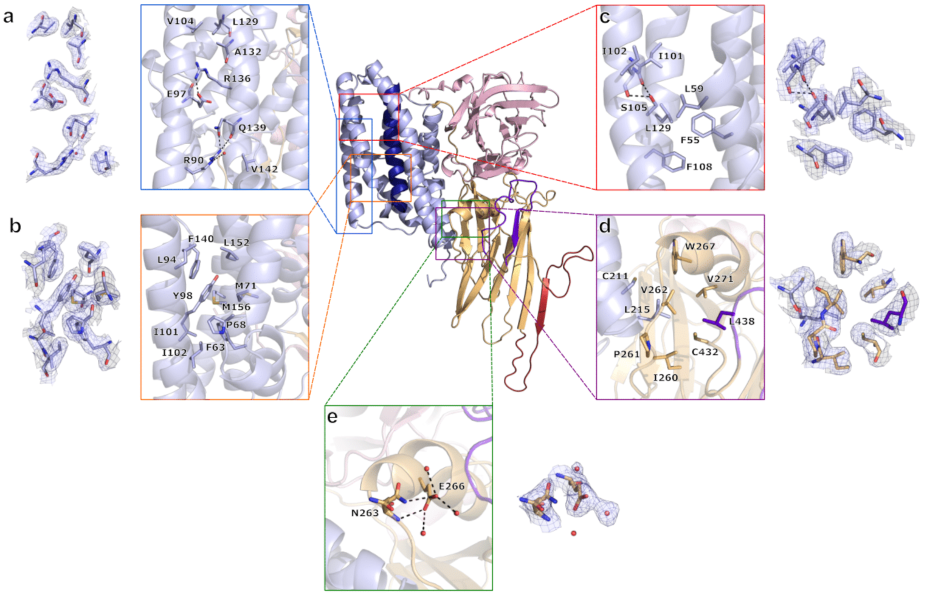
Supplementary Fig. 14. Toxin state (crystal *vs* soluble) but not point-mutations affected the thermal stability and aggregation propensity of Cry11Aa. a-b, Differential scanning fluorometry measurements indicate the thermal unfolding of crystals (**a**) and soluble (**b**) Cry11Aa WT and mutants in function of the temperature. **c-d**, Scattering measurement indicates aggregation propensity of crystals (**c**) and soluble toxins (**d**).



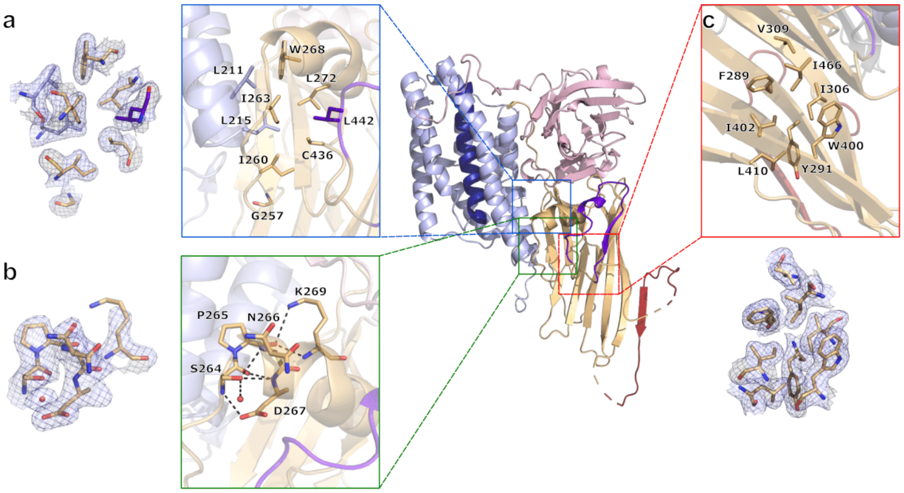
Supplementary Fig. 15. Cry11Aa mutant crystal structures indicating the crystal packing (top), tetramer (middle) and interaction region (bottom). Color code as in Fig. 2 and 3. **a**, Cry11Aa-F17Y; **b**, Cry11Aa-E583Q; **c**, Cry11Aa-Y449F. The dashes in (**a**) indicate the hydrogen bond made between E180(OE1) and Y17(OH) (2.4 Å). Due to the lower resolution as compared to Cry11Aa-WT, no water molecules are observed in the interaction region. Therefore, the specific interaction pattern is not shown in (**b**) and (**c**). The $2F_{obs}$-F_{calc} electron density map shown at the bottom of each panel is contoured at 1 σ.



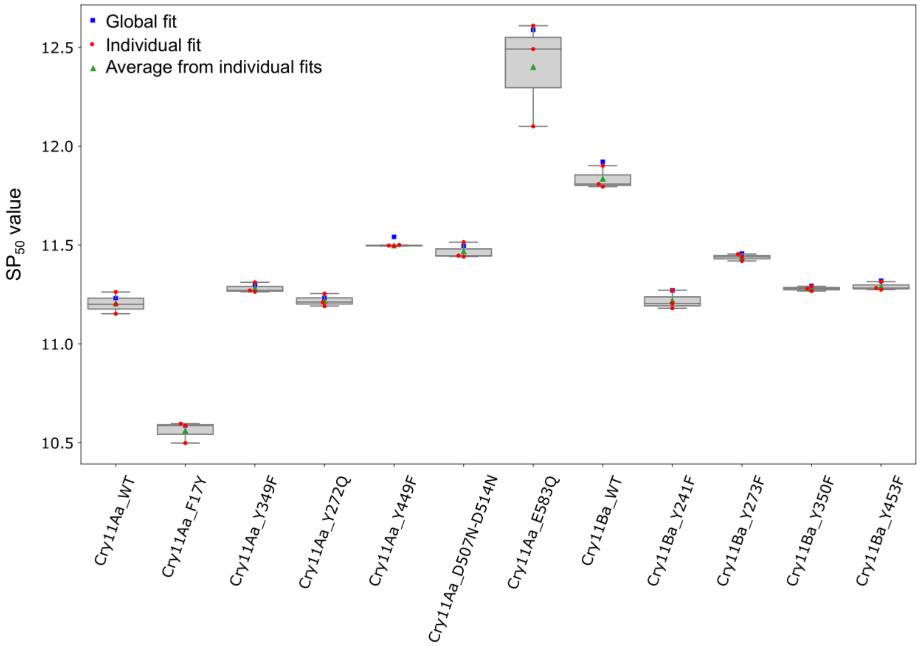
Supplementary Fig. 16. Potential mutation candidates to destabilize the domain I – domain III interface. Interactions between the concerned residues in chain A of Cry11Aa (**a**) and Cry11Ba (**b**), respectively. Color code as in Fig. 2. Dashes indicate hydrogen bonds or salt bridges (up to 3.2 Å). The $2F_{obs}$-F_{calc} map on the bottom of each panel is contoured at 1 σ.



Supplementary Fig. 17. Structural interpretation of Cry11Aa mutations described in literature. Color code as in Fig. 2. Dashes indicate hydrogen bonds (up to 3.2 Å). The 2F$_{obs}$-F$_{calc}$ electron density maps on the left and right of each panel are contoured at 1 σ. **a**, Position of and interactions formed by residues in Domain I which can explain the suppression of toxicity of the V104E mutation, and reduced toxicity by the E97A, R90E and V142E mutations. **b**, Hydrophobic pocket in which Y98 resides. **c**, Interactions made by S105 in domain I. **d**, Hydrophobic pocket formed by P261 and V262 at the interface between α8 and the β1-α8 loop. **e**, Polar interactions formed by E266 at the interface between α8 and the β1-α8 loop.



Supplementary Fig. 18. Structural interpretation of Cry11Ba mutations described in literature. Chain A of Cry11Ba is shown, colored as in Fig. 2. The $2F_{obs}$-F_{calc} electron density maps on the bottom of each panel are contoured at 1 σ. **a**, Hydrophobic pocket in which I263 resides and position of G257 in the turn between α8 and the β1-α8 loop. **b**, Polar interactions formed by S264 and K269 in the turn between α8 and the β1-α8 loop. **c**, Hydrophobic pocket in which I306 fits. Dashes indicate hydrogen bonds (up to 3.2 Å), and the salt bridge between K269(NZ) and P265(O) in (**b**).



Supplementary Fig. 19. SP_{50} values determined for Cry11Aa, Cry11Ba and their mutants. Blue squares show the SP_{50} values as determined from a global fit to the measurements (n=3 individual experiments). Red spheres and boxplots show the SP_{50} values determined from fitting the three individual measurements, with green triangles showing the average value. The boxes represent the lower and upper quartiles around the median. Whiskers indicate the minimum and maximum values measured amongst replicates.

Supplementary Table 1. Interactions in the mutated interfaces. Hydrogen bonds up to 3.2 Å are listed.

a. Direct and single water-mediated interactions formed by D515, Y272 and D507

Interaction partner 1	Interaction partner 2	Distance (Å)*
D514 (OD2)	S249 (O)	2.9 / 2.9
	S251 (OG)	2.4
D514 (OD1)	S251 (N)	3.1
	W253 (NE1)	2.4 / 3.0
	Y272 (OH)	2.7
	R222 (NH1)	2.4 / 3.1
Y272 (OH)	R222 (NH1)	3.2 / 3.1
D507 (OD1)	R222 (NH1)	2.5
D507 (OD2)	R222 (NH2)	3.2
	Y203 (OH)	2.7 / 2.7
	W267 (NE1)	2.9

* Multiple distances indicate water-mediated hydrogen bonds.

b, Extended water-mediated interactions formed by E583 and other residues involved in this network

Interaction partner 1	Interaction partner 2	Distance (Å)*
E583 (OE1)	T197 (OG1)	2.8 / 2.9
	L193 (O)	2.8 / 2.8
	V192 (O)	2.6 / 2.7 / 3.0 / 3.1
	A233 (O)	2.6 / 2.7 / 2.8 / 2.9 / 2.6 / 2.9
W584 (N)	V192 (O)	3.2 / 2.7 / 3.0 / 3.1
	A223 (O)	3.2 / 2.7 / 2.8 / 2.9 / 2.6 / 2.9
Q511 (NE2)	V192 (O)	3.0 / 2.8 / 3.0 / 3.1
	A223 (O)	3.0 / 2.9 / 2.6 / 2.9
D510 (O)	V192 (O)	3.0 / 2.9 / 2.8 / 3.0 / 3.1
	A223 (O)	3.0 / 2.6 / 2.9
D510 (OD2)	R196 (NE)	3.1
	R196 (NH2)	3.3

* Multiple distances indicate water-mediated hydrogen bonds.

c, Extended water-mediated interactions formed by Y449 and D501, connecting two monomers via interface #1

Interaction partner monomer 1	Interaction partner monomer 2	Distance (Å)*
Y449 (OH)	V499 (N)	2.7 / 2.8
	V499 (O)	2.9 / 2.8
	T448 (OG1)	2.9 / 2.6
	D501 (OD1)	2.9 / 2.8
D501 (OD2)	Q281 (NE2)	2.8 / 2.9 / 2.6 / 3.0

* Multiple distances indicate water-mediated hydrogen bonds.

d, Hydrogen bonds made by Y349

Interaction partner monomer 1	Interaction partner monomer 2	Distance (Å)
Y349 (OH)	E295 (OE1)	3.1

Supplementary Table 2. Effects of Cry11Aa and Cry11Ba point mutations on toxicity described in literature

Mutant		Toxicity			References§
		Aedes aegypti	*Anopheles stephensi*	*Culex quinquefasciatus*	
Cry11Aa	R90E	Not toxic	N.D.*	N.D.	1
	E97A	Not toxic	N.D.	N.D.	1, 2
	Y98E	Not toxic	N.D.	N.D.	1
	V104E	No crystal formation	No crystal formation	No crystal formation	1
	S105E	Not toxic	N.D.	N.D.	1
	V142D	Not toxic	N.D.	N.D.	2
	P261A	No change / 3-fold reduced	N.D.	N.D.	3, 4
	V262A	No change / 5-fold reduced	N.D.	N.D.	3, 4
	V262E	Not toxic	N.D.	N.D.	3
	E266A	No change / 21-fold reduced	N.D.	N.D.	3, 4
Cry11Ba	G257A	Strongly reduced	Reduced	Strongly reduced	5
	I263A	Not toxic	Not toxic	Strongly reduced	5
	S264A	Not toxic	No change	Reduced	5
	K269A	Strongly reduced	No change	Strongly reduced	5
	I306A	Not toxic	No change	Reduced	5

* N.D.: Not determined;

§ References:

1. Munoz-Garay, C. et al. Oligomerization of Cry11Aa from Bacillus thuringiensis Has an Important Role in Toxicity against Aedes aegypti. Appl. Environ. Microbiol. 75, 7548–7550 (2009).
2. Carmona, D. et al. Dominant Negative Phenotype of Bacillus thuringiensis Cry1Ab, Cry11Aa and Cry4Ba Mutants Suggest Hetero-Oligomer Formation among Different Cry Toxins. PLoS ONE 6, e19952 (2011).
3. Fernandez, L. E. et al. Cry11Aa toxin from Bacillus thuringiensis binds its receptor in Aedes aegypti mosquito larvae through loop alpha-8 of domain II. Febs Lett. 579, 3508–3514 (2005).
4. Perez, C. et al. Bacillus thuringiensis subsp israelensis Cyt1Aa synergizes Cry11Aa toxin by functioning as a membrane-bound receptor. Proc. Natl. Acad. Sci. U. S. A. 102, 18303–18308 (2005).
5. Likitvivatanavong, S., Aimanova, K. G. & Gill, S. S. Loop residues of the receptor binding domain of Bacillus thuringiensis Cry11Ba toxin are important for mosquitocidal activity. FEBS Lett. 583, 2021–2030 (2009).

Supplementary Table 3. Primers used to generate the Cry11Aa mutants.

Mutation	Forward / Reverse	Primer sequence (5'-3')	Comment
/	Forward	GCCGCAGTGTTATCACTCATGGTTATGGC	Amp_F1 was used with each reverse primer of Cry11Aa mutant construction. For each mutation primer, the inserted mutation is indicated as a capital bold letter. To generate the vector containing the mutation, the two fragments generated for the same mutations were assembled by Gibson assembly following the procedure described in the manuscript.
F17Y	Reverse	TTATATAATGGA**T**AGTCTGTTTCATTAACTATACTTAAAGTATCTAAAGAACTATCT	
Y272Q	Reverse	TAATAACTT**T**T**G**AACTAGTCCTCCCCATTCATTATAATTTACAGG	
Y349F	Reverse	TGGATTTTGA**A**AAAAAGTTTGTGTTCTAACTTCCATAATTT	
Y449F	Reverse	TCTATTCTATTA**A**ATGTTAGTGTTGTAAGATCGTGATTATATGCAAG	
D507N-D514N	Reverse	TTGCTTGATCTGGCGTAT**T**TTCTAAAAATGATCTATCT	
E583Q	Reverse	ATCCACCCATT**G**TGGATTAGCATTTG	
/	Reverse	AGTGCTGCCATAACCATGAGTGATAACACT	Amp_R1 was used with each forward primer of Cry11Aa mutant construction. For each mutation primer, the inserted mutation is indicated as a capital bold letter. To generate the vector containing the mutation, the two fragments generated for the same mutations were assembled by Gibson assembly following the procedure described in the manuscript.
F17Y	Forward	GTTAATGAAACAGACT**A**TCCATTATATAATAATTATACCGAACCTACT	
Y272Q	Forward	GGGAGGACTAGTT**C**A**A**AAGTTATTAATGGGGGA	
Y349F	Forward	CACAAACTTTTT**T**TCAAAATCCAAATAATGAGCCT	
Y449F	Forward	TTACAACACTAACAT**T**TAATAGAATAGAGTATGATTCACCTACTACAG	
D507N-D514N	Forward	ATACGCCAGATCAAGCAACA**A**ACGGCAGTATTAAATTTG	
E583Q	Forward	TAATCCA**C**AATGGGTGGATTTTGTCACAG	
C11AB chimera	Forward	CTATCCTAAATAGGCGATCGCACTCATTAGGC	This fragment was obtained using Amp_R1 as reverse primer.
C11AB chimera	Reverse	CGTACAATAACCTTTAGTAACGGATTAATTTGCGTCGTAAAGG	This fragment was obtained using Amp_F1 as forward primer.
C11AB chimera	Forward	CGTTACTAAAGGTTATTGTACGTTGTCCGGATACTTTTTTTGTG	This fragment was assembled with the two above-mentioned fragments by Gibson to create the shuffle vector encoding the C11AB chimera.
C11AB chimera	Reverse	CGATCGCCTATTTAGGATAGTTATTGTTATACCCTTGGTTACATGTACAG	

Supplementary Table 4. Primers used to generate the Cry11Ba mutants.

Mutation	Gblock sequence
Silent WT mutants to reduce repeats and secondary structures	GACCATGATTACGAATTGGTACCTTTTCGATTTCAAATTTTCCAAACTTAAATATGATTGAATGCCTGAGAAAGGTAATAGAGATGTTTTAGTTTATTATGAAGTATTAGGGGCGTCTTTTAAATTCAATCTATCAATTTGTGAAATATATTACTCAAAACCCAATACCATTCTAAAACTTATTCAAAATATATATTGCTTTAAAAGAGCATACATACTAAAAAAACAGGCATCTTTCGAACTATAGCGCATAGAATACTACGGTGAATCAAAAACAAATAAAATTTAGGAGGTATATTCAAGTATACAAAAAAACTTTAGTGTGAGGGGATTTAGATAAAAAGTATTCGTTATCCTTATAAATTAATTCTTAAACATGCACCAATGTATACATTAAATAATATTATGTGAATTAAGTCTATCAATTTAATTTATTATGTTACTTTATATTTGATTAATAATTGCAAGTTTAAAATCATAATTTAATGTTGAAAGGCCACTATTCTAATTAACTTAAGGAGTTGTTTATTTATGCAAAATAACAACTTTAATACCACAGAAATTAATAATATGATTAATTTCCCTATGTATAATGGTAGATTAGAACCTTCTCTAGCTCCAGCATTAATAGCAGTAGCTCCAATTGCTAAATATTTAGCAACAGCTCTTGCTAAATGGGCTGTAAAACAAGGGTTTGCAAAATTAAAATCCGAGATATTCCCCGGTAATACGCCTGCTACTATGGATAAGGTTCGTATTGAGGTACAAACACTTTTAGACCAAAGATTACAAGATGACAGAGTTAAGATTTTAGAAGGTGAATACAAAGGAATTATTGACGTGAGTAAAGTTTTTACTGATTATGTTAATCAATCTAAATTTGAGACTGGAACAGCTAATAGGCTTTTTTTTGATACAAGTAACCAATTAATAAGCAGATTGCCTCAATTTGAGATTGCAGGATATGAAGGAGTATCCATTTCACTTTTTACTCAGATGTGTACATTTCATTTGGGTTTATTAAAAGATGGAATTTTAGCAGGAAGCGATTGGGGATTTGCTCCTGCAGATAAAGACGCTCTTATTTGCCAATTTAATAGATTTGTCAATGAATATAATACTCGACTGATGGTATTGTACTCAAAAGAATTTGGACGGTTATTAGCAAAAAATCTTAATGAAGCCTTGAACTTTAGAAATATGTGTAGTTTATATGTCTTTCCTTTTTCTGAAGCATGGTCTTTATTAAGGTATGAAGGAACAAAATTAGAAAACACGCTTTCATTATGGAATTTTGTGGGTGAAAGTATCAATAATATATCTCCTAATGATTGGAAAGGTGCGCTTTATAAATTGTTAATGGGAGCACCTAATCAAAGATTAAACAATGTTAAGTTTAATTATAGTTATTTTTCTGATACTCAAGCGACAATACATCGTGAAAACATTCATGGTGTCCTGCCAACATATAATGGAGGACCAACAATTACAGGATGGATAGGGAATGGGCGTTTCAGCGGACTTAGTTTTCCTTGTAGTAATGAATTAGAAATTACAAAAATAAAACAGGAAATAACTTACAATGATAAAGGGGGAAATTTCAATTCAATAGTTCCTGCTGCTACGCGCAATGAAATTCTAACTGCTACCGTTCCAACATCAGCTGATCCATTTTTTAAAACCGCTGATATTAACTGGAAATATTTCTCTCCGGGTCTTTACTCTGGATGGAATATTAAATTTGATGATACAGTCACTTTAAAAAGTAGAGTACCAAGTATTATACCTTCAAATATATTAAAGTATGATGATTATTATATTCGTGCCGTTTCAGCCTGTCCAAAAGGCGTATCACTTGCATATAACCATGATTTTTTAACGTTAACATATAACAAATTAGAATATGATGCACCTACTACACAAAATATCATTGTAGGATTTTCACCAGATAATACTAAGAGTTTTTATAGGAGCAACTCTCATTATCTAAGTACAACAGATGATGCCTATGTAATTCCTGCTTTACAATTTTCTACAGTCTCAGATAGATCATTCTTAGAAGATACACCAGATCAAGCAACAGATGGCAGTATTAAATTTACGGATACTGTTCTTGGGAATGAGGCAAAATATTCTATTAGACTAAATACTGGATTTAATACAGCTACTAGGTATAGATTAATTATACGTTTTAAAGCGCCTGCTCGTTTGGCTGCTGGTATACGTGTACGTTCTCAAAATTCAGGGAATAATAAGTTATTAGGTGGTATTCCTGTAGAGGGTAATTCTGGATGGATAGATTATATTACAGATTCATTTACTTTTGATGACCTTGGGATTACAACTTCAAGTACAAATGCTTTCTTTAGTATTGATTCAGATGGTGTAAATGCTTCTCAACAATGGTATTTGTCTAAATTAATTTTAGTAAAAGAATCCAGTTTTACGACTCAGATTCCATTAAAACCATACGTTATTGTACGTTGTCCGGATACTTTTTTTGTGAGCAACAATTCAAGTAGTACGTACGAACAAGGCTATAACAACAATTACAACCAGAATTCTAGCAGTATGTACGATCAAGGATATAACAATAGCTATAATCCAAACTCTGGTTGTACGTGTAATCAAGACTATAATAACAGTTATAACCAAAACTCTGGCTGTACATGTAACCAAGGGTATAACAATAACTATCCTAAATAATCTTAGTAGCTATATTTATTAAATATGGTAATATCACAAGTATAAATACTTGTGGTATTACCTACCATTCTTAAATTATATCCAAAATCATGCGTTAATCTACATTCCCCTTTCTCTAAAATTTGTTCTTCACACATCCACATTTTTCGACTCGAGGCATGCAAGCTTGGC

(Continued)

Mutation	Gblock sequence*
Y241F	GACCATGATTACGAATTGGTACCTTTTCGATTTCAAATTTTCCAAACTTAAATATGATTGAATGCCTGAGAAAGGTAATAGAGATGTTTTAGTTTATTATGAAGTATTAGGGGCGTCTTTTAAATTCAATCTATCAATTTGTGAAATATATTACTCAAAACCCAATACCATTCTAAAACTTATTCAAAATATATATTGCTTTAAAAGAGCATACATACTAAAAAAACAGGCATCTTTCGAACTATAGCGCATAGAATACTACGGTGAATCAAAAACAAATAAAATTTAGGAGGTATATTCAAGTATACAAAAAAACTTTAGTGTGAGGGGATTTAGATAAAAAGTATTCGTTATCCTTATAAATTAATTCTTAAACATGCACCAATGTATACATTAAATAATATTATGTGAATTAAGTCTATCAATTTAATTTATTATGTTACTTTATATTTGATTAATAATTGCAAGTTTAAAATCATAATTTAATGTTGAAAGGCCACTATTCTAATTAACTTAAGGAGTTGTTTATTTATGCAAAATAACAACTTTAATACCACAGAAATTAATAATATGATTAATTTCCCTATGTATAATGGTAGATTAGAACCTTCTCTAGCTCCAGCATTAATAGCAGTAGCTCCAATTGCTAAATATTTAGCAACAGCTCTTGCTAAATGGGCTGTAAAACAAGGGTTTGCAAAATTAAAATCCGAGATATTCCCCGGTAATACGCCTGCTACTATGGATAAGGTTCGTATTGAGGTACAAACACTTTTAGACCAAAGATTACAAGATGACAGAGTTAAGATTTTAGAAGGTGAATACAAAGGAATTATTGACGTGAGTAAAGTTTTTACTGATTATGTTAATCAATCTAAATTTGAGACTGGAACAGCTAATAGGCTTTTTTTTGATACAAGTAACCAATTAATAAGCAGATTGCCTCAATTTGAGATTGCAGGATATGAAGGAGTATCCATTTCACTTTTTACTCAGATGTGTACATTTCATTTGGGTTTATTAAAAGATGGAATTTTAGCAGGAAGCGATTGGGGATTTGCTCCTGCAGATAAAGACGCTCTTATTTGCCAATTTAATAGATTTGTCAATGAATATAATACTCGACTGATGGTATTGTACTCAAAAGAATTTGGACGGTTATTAGCAAAAAATCTTAATGAAGCCTTGAACTTTAGAAATATGTGTAGTTTATATGTCTTTCCTTTTTCTGAAGCATGGTCTTTATTAAGGT**T**TGAAGGAACAAAATTAGAAAACACGCTTTCATTATGGAATTTTGTGGGTGAAAGTATCAATAATATATCTCCTAATGATTGGAAAGGTGCGCTTTATAAATTGTTAATGGGAGCACCTAATCAAAGATTAAACAATGTTAAGTTTAATTATAGTTATTTTTCTGATACTCAAGCGACAATACATCGTGAAAACATTCATGGTGTCCTGCCAACATATAATGGAGGACCAACAATTACAGGATGGATAGGGAATGGGCGTTTCAGCGGACTTAGTTTTCCTTGTAGTAATGAATTAGAAATTACAAAAATAAAACAGGAAATAACTTACAATGATAAAGGGGGAAATTTCAATTCAATAGTTCCTGCTGCTACGCGCAATGAAATTCTAACTGCTACCGTTCCAACATCAGCTGATCCATTTTTTAAAACCGCTGATATTAACTGGAAATATTTCTCTCCGGGTCTTTACTCTGGATGGAATATTAAATTTGATGATACAGTCACTTTAAAAAGTAGAGTACCAAGTATTATACCTTCAAATATATTAAAGTATGATGATTATTATATTCGTGCCGTTTCAGCCTGTCCAAAAGGCGTATCACTTGCATATAACCATGATTTTTTAACGTTAACATATAACAAATTAGAATATGATGCACCTACTACACAAAATATCATTGTAGGATTTTCACCAGATAATACTAAGAGTTTTTATAGGAGCAACTCTCATTATCTAAGTACAACAGATGATGCCTATGTAATTCCTGCTTTACAATTTTCTACAGTCTCAGATAGATCATTCTTAGAAGATACACCAGATCAAGCAACAGATGGCAGTATTAAATTTACGGATACTGTTCTTGGGAATGAGGCAAAATATTCTATTAGACTAAATACTGGATTTAATACAGCTACTAGGTATAGATTAATTATACGTTTTAAAGCGCCTGCTCGTTTGGCTGCTGGTATACGTGTACGTTCTCAAAATTCAGGGAATAATAAGTTATTAGGTGGTATTCCTGTAGAGGGTAATTCTGGATGGATAGATTATATTACAGATTCATTTACTTTTGATGACCTTGGGATTACAACTTCAAGTACAAATGCTTTCTTTAGTATTGATTCAGATGGTGTAAATGCTTCTCAACAATGGTATTTGTCTAAATTAATTTTAGTAAAAGAATCCAGTTTTACGACTCAGATTCCATTAAAACCATACGTTATTGTACGTTGTCCGGATACTTTTTTTGTGAGCAACAATTCAAGTAGTACGTACGAACAAGGCTATAACAACAATTACAACCAGAATTCTAGCAGTATGTACGATCAAGGATATAACAATAGCTATAATCCAAACTCTGGTTGTACGTGTAATCAAGACTATAATAACAGTTATAACCAAAACTCTGGCTGTACATGTAACCAAGGGTATAACAATAACTATCCTAAATAATCTTAGTAGCTATATTTATTAAATATGGTAATATCACAAGTATAAATACTTGTGGTATTACCTACCATTCTTAAATTATATCCAAAATCATGCGTTAATCTACATTCCCCTTTCTCTAAAATTTGTTCTTCACACATCCACATTTTTCGACTCGAGGCATGCAAGCTTGGC

*The inserted mutation is indicated as a capital bold letter and underlined

(Continued)

Mutation	Gblock sequence*
Y273F	GACCATGATTACGAATTGGTACCTTTTCGATTTCAAATTTTCCAAACTTAAATATGATTGAATGCCTGAGAAAGGTAATAGAGATGTTTTAGTTTATTATGAAGTATTAGGGGCGTCTTTTAAATTCAATCTATCAATTTGTGAAATATATTACTCAAAACCCAATACCATTCTAAAACTTATTCAAAATATATATTGCTTTAAAAGAGCATACATACTAAAAAAACAGGCATCTTTCGAACTATAGCGCATAGAATACTACGGTGAATCAAAAACAAATAAAATTTAGGAGGTATATTCAAGTATACAAAAAAACTTTAGTGTGAGGGGATTTAGATAAAAAGTATTCGTTATCCTTATAAATTAATTCTTAAACATGCACCAATGTATACATTAAATAATATTATGTGAATTAAGTCTATCAATTTAATTTATTATGTTACTTTATATTTGATTAATAATTGCAAGTTTAAAATCATAATTTAATGTTGAAAGGCCACTATTCTAATTAACTTAAGGAGTTGTTTATTTATGCAAAATAACAACTTTAATACCACAGAAATTAATAATATGATTAATTTCCCTATGTATAATGGTAGATTAGAACCTTCTCTAGCTCCAGCATTAATAGCAGTAGCTCCAATTGCTAAATATTTAGCAACAGCTCTTGCTAAATGGGCTGTAAAACAAGGGTTTGCAAAATTAAAATCCGAGATATTCCCCGGTAATACGCCTGCTACTATGGATAAGGTTCGTATTGAGGTACAAACACTTTTAGACCAAAGATTACAAGATGACAGAGTTAAGATTTTAGAAGGTGAATACAAAGGAATTATTGACGTGAGTAAAGTTTTTACTGATTATGTTAATCAATCTAAATTTGAGACTGGAACAGCTAATAGGCTTTTTTTTGATACAAGTAACCAATTAATAAGCAGATTGCCTCAATTTGAGATTGCAGGATATGAAGGAGTATCCATTTCACTTTTTACTCAGATGTGTACATTTCATTTGGGTTTATTAAAAGATGGAATTTTAGCAGGAAGCGATTGGGGATTTGCTCCTGCAGATAAAGACGCTCTTATTTGCCAATTTAATAGATTTGTCAATGAATATAATACTCGACTGATGGTATTGTACTCAAAAGAATTTGGACGGTTATTAGCAAAAAATCTTAATGAAGCCTTGAACTTTAGAAATATGTGTAGTTTATATGTCTTTCCTTTTTCTGAAGCATGGTCTTTATTAAGGTATGAAGGAACAAAATTAGAAAACACGCTTTCATTATGGAATTTTGTGGGTGAAAGTATCAATAATATATCTCCTAATGATTGGAAAGGTGCGCTTT**T**TAAATTGTTAATGGGAGCACCTAATCAAAGATTAAACAATGTTAAGTTTAATTATAGTTATTTTTCTGATACTCAAGCGACAATACATCGTGAAAACATTCATGGTGTCCTGCCAACATATAATGGAGGACCAACAATTACAGGATGGATAGGGAATGGGCGTTTCAGCGGACTTAGTTTTCCTTGTAGTAATGAATTAGAAATTACAAAAATAAAACAGGAAATAACTTACAATGATAAAGGGGGAAATTTCAATTCAATAGTTCCTGCTGCTACGCGCAATGAAATTCTAACTGCTACCGTTCCAACATCAGCTGATCCATTTTTTAAAACCGCTGATATTAACTGGAAATATTTCTCTCCGGGTCTTTACTCTGGATGGAATATTAAATTTGATGATACAGTCACTTTAAAAAGTAGAGTACCAAGTATTATACCTTCAAATATATTAAAGTATGATGATTATTATATTCGTGCCGTTTCAGCCTGTCCAAAAGGCGTATCACTTGCATATAACCATGATTTTTTAACGTTAACATATAACAAATTAGAATATGATGCACCTACTACACAAAATATCATTGTAGGATTTTCACCAGATAATACTAAGAGTTTTTATAGGAGCAACTCTCATTATCTAAGTACAACAGATGATGCCTATGTAATTCCTGCTTTACAATTTTCTACAGTCTCAGATAGATCATTCTTAGAAGATACACCAGATCAAGCAACAGATGGCAGTATTAAATTTACGGATACTGTTCTTGGGAATGAGGCAAAATATTCTATTAGACTAAATACTGGATTTAATACAGCTACTAGGTATAGATTAATTATACGTTTTAAAGCGCCTGCTCGTTTGGCTGCTGGTATACGTGTACGTTCTCAAAATTCAGGGAATAATAAGTTATTAGGTGGTATTCCTGTAGAGGGTAATTCTGGATGGATAGATTATATTACAGATTCATTTACTTTTGATGACCTTGGGATTACAACTTCAAGTACAAATGCTTTCTTTAGTATTGATTCAGATGGTGTAAATGCTTCTCAACAATGGTATTTGTCTAAATTAATTTTAGTAAAAGAATCCAGTTTTACGACTCAGATTCCATTAAAACCATACGTTATTGTACGTTGTCCGGATACTTTTTTTGTGAGCAACAATTCAAGTAGTACGTACGAACAAGGCTATAACAACAATTACAACCAGAATTCTAGCAGTATGTACGATCAAGGATATAACAATAGCTATAATCCAAACTCTGGTTGTACGTGTAATCAAGACTATAATAACAGTTATAACCAAAACTCTGGCTGTACATGTAACCAAGGGTATAACAATAACTATCCTAAATAATCTTAGTAGCTATATTTATTAAATATGGTAATATCACAAGTATAAATACTTGTGGTATTACCTACCATTCTTAAATTATATCCAAAATCATGCGTTAATCTACATTCCCCTTTCTCTAAAATTTGTTCTTCACACATCCACATTTTTCGACTCGAGGCATGCAAGCTTGGC

*The inserted mutation is indicated as a capital bold letter and underlined.

(Continued)

Mutation	Gblock sequence*
Y350F	GACCATGATTACGAATTGGTACCTTTTCGATTTCAAATTTTCCAAACTTAAATATGATTGAATGCCTGAGAAAGGTAATAGAGATGTTTTAGTTTATTATGAAGTATTAGGGGCGTCTTTTAAATTCAATCTATCAATTTGTGAAATATATTACTCAAAACCCAATACCATTCTAAAACTTATTCAAAATATATATTGCTTTAAAAGAGCATACATACTAAAAAAACAGGCATCTTTCGAACTATAGCGCATAGAATACTACGGTGAATCAAAAACAAATAAAATTTAGGAGGTATATTCAAGTATACAAAAAAACTTTAGTGTGAGGGGATTTAGATAAAAAGTATTCGTTATCCTTATAAATTAATTCTTAAACATGCACCAATGTATACATTAAATAATATTATGTGAATTAAGTCTATCAATTTAATTTATTATGTTACTTTATATTTGATTAATAATTGCAAGTTTAAAATCATAATTTAATGTTGAAAGGCCACTATTCTAATTAACTTAAGGAGTTGTTTATTTATGCAAAATAACAACTTTAATACCACAGAAATTAATAATATGATTAATTTCCCTATGTATAATGGTAGATTAGAACCTTCTCTAGCTCCAGCATTAATAGCAGTAGCTCCAATTGCTAAATATTTAGCAACAGCTCTTGCTAAATGGGCTGTAAAACAAGGGTTTGCAAAATTAAAATCCGAGATATTCCCCGGTAATACGCCTGCTACTATGGATAAGGTTCGTATTGAGGTACAAACACTTTTAGACCAAAGATTACAAGATGACAGAGTTAAGATTTTAGAAGGTGAATACAAAGGAATTATTGACGTGAGTAAAGTTTTTACTGATTATGTTAATCAATCTAAATTTGAGACTGGAACAGCTAATAGGCTTTTTTTTGATACAAGTAACCAATTAATAAGCAGATTGCCTCAATTTGAGATTGCAGGATATGAAGGAGTATCCATTTCACTTTTTACTCAGATGTGTACATTTCATTTGGGTTTATTAAAAGATGGAATTTTAGCAGGAAGCGATTGGGGATTTGCTCCTGCAGATAAAGACGCTCTTATTTGCCAATTTAATAGATTTGTCAATGAATATAATACTCGACTGATGGTATTGTACTCAAAAGAATTTGGACGGTTATTAGCAAAAAATCTTAATGAAGCCTTGAACTTTAGAAATATGTGTAGTTTATATGTCTTTCCTTTTTCTGAAGCATGGTCTTTATTAAGGTATGAAGGAACAAAATTAGAAAACACGCTTTCATTATGGAATTTTGTGGGTGAAAGTATCAATAATATATCTCCTAATGATTGGAAAGGTGCGCTTTATAAATTGTTAATGGGAGCACCTAATCAAAGATTAAACAATGTTAAGTTTAATTATAGTTATTTTTCTGATACTCAAGCGACAATACATCGTGAAAACATTCATGGTGTCCTGCCAACATATAATGGAGGACCAACAATTACAGGATGGATAGGGAATGGGCGTTTCAGCGGACTTAGTTTTCCTTGTAGTAATGAATTAGAAATTACAAAAATAAAACAGGAAATAACTT**T**CAATGATAAAGGGGGAAATTTCAATTCAATAGTTCCTGCTGCTACGCGCAATGAAATTCTAACTGCTACCGTTCCAACATCAGCTGATCCATTTTTTAAAACCGCTGATATTAACTGGAAATATTTCTCTCCGGGTCTTTACTCTGGATGGAATATTAAATTTGATGATACAGTCACTTTAAAAAGTAGAGTACCAAGTATTATACCTTCAAATATATTAAAGTATGATGATTATTATATTCGTGCCGTTTCAGCCTGTCCAAAAGGCGTATCACTTGCATATAACCATGATTTTTTAACGTTAACATATAACAAATTAGAATATGATGCACCTACTACACAAAATATCATTGTAGGATTTTCACCAGATAATACTAAGAGTTTTTATAGGAGCAACTCTCATTATCTAAGTACAACAGATGATGCCTATGTAATTCCTGCTTTACAATTTTCTACAGTCTCAGATAGATCATTCTTAGAAGATACACCAGATCAAGCAACAGATGGCAGTATTAAATTTACGGATACTGTTCTTGGGAATGAGGCAAAATATTCTATTAGACTAAATACTGGATTTAATACAGCTACTAGGTATAGATTAATTATACGTTTTAAAGCGCCTGCTCGTTTGGCTGCTGGTATACGTGTACGTTCTCAAAATTCAGGGAATAATAAGTTATTAGGTGGTATTCCTGTAGAGGGTAATTCTGGATGGATAGATTATATTACAGATTCATTTACTTTTGATGACCTTGGGATTACAACTTCAAGTACAAATGCTTTCTTTAGTATTGATTCAGATGGTGTAAATGCTTCTCAACAATGGTATTTGTCTAAATTAATTTTAGTAAAAGAATCCAGTTTTACGACTCAGATTCCATTAAAACCATACGTTATTGTACGTTGTCCGGATACTTTTTTTGTGAGCAACAATTCAAGTAGTACGTACGAACAAGGCTATAACAACAATTACAACCAGAATTCTAGCAGTATGTACGATCAAGGATATAACAATAGCTATAATCCAAACTCTGGTTGTACGTGTAATCAAGACTATAATAACAGTTATAACCAAAACTCTGGCTGTACATGTAACCAAGGGTATAACAATAACTATCCTAAATAATCTTAGTAGCTATATTTATTAAATATGGTAATATCACAAGTATAAATACTTGTGGTATTACCTACCATTCTTAAATTATATCCAAAATCATGCGTTAATCTACATTCCCCTTTCTCTAAAATTTGTTCTTCACACATCCACATTTTTCGACTCGAGGCATGCAAGCTTGGC

*The inserted mutation is indicated as a capital bold letter and underlined.

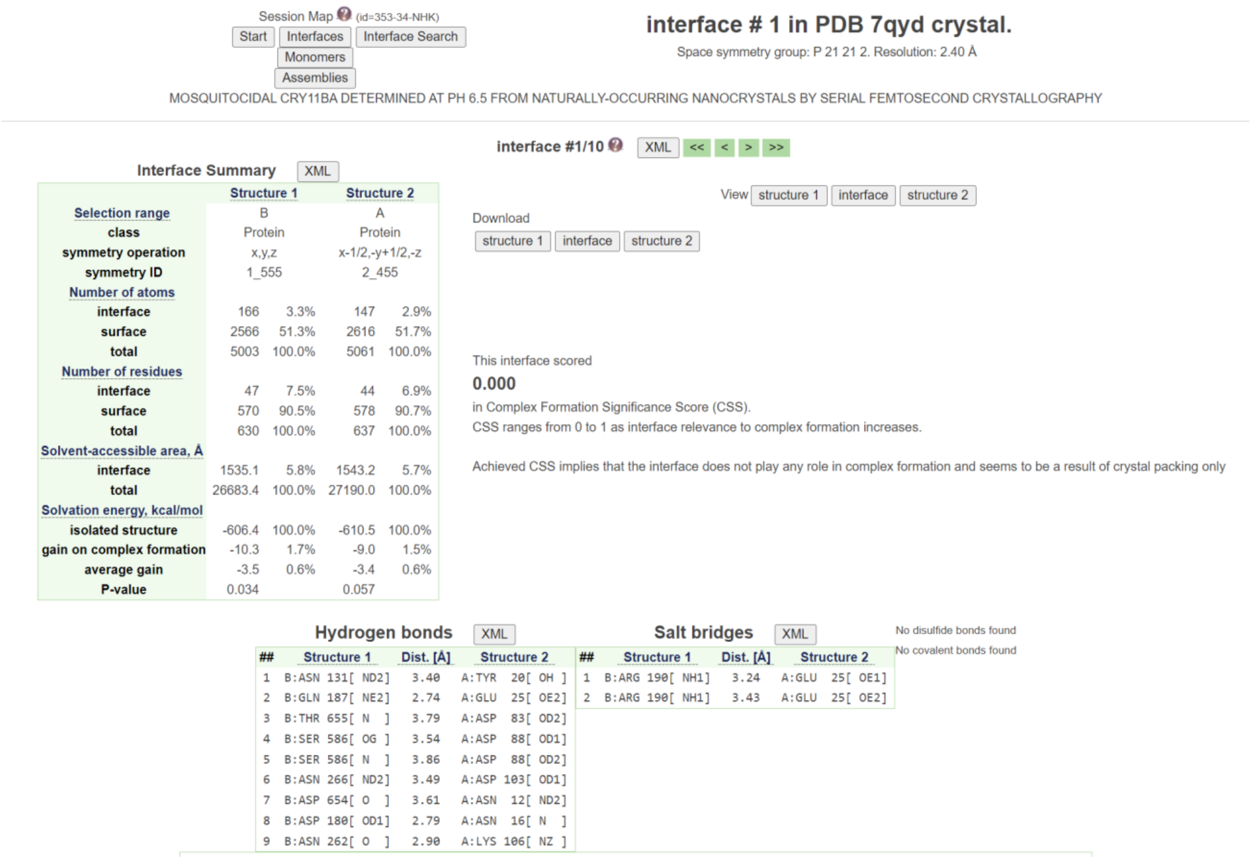
(Continued)

Mutation	Gblock sequence*
Y453F	GACCATGATTACGAATTGGTACCTTTTCGATTTCAAATTTTCCAAACTTAAATATGATTGAATGCCTGAGA AAGGTAATAGAGATGTTTTAGTTTATTATGAAGTATTAGGGGCGTCTTTTAAATTCAATCTATCAATTTGT GAAATATATTACTCAAAACCCAATACCATTCTAAAACTTATTCAAAATATATATTGCTTTAAAAGAGCATA CATACTAAAAAAACAGGCATCTTTCGAACTATAGCGCATAGAATACTACGGTGAATCAAAAACAAATAAA ATTTAGGAGGTATATTCAAGTATACAAAAAAACTTTAGTGTGAGGGGATTTAGATAAAAAGTATTCGTTAT CCTTATAAATTAATTCTTAAACATGCACCAATGTATACATTAAATAATATTATGTGAATTAAGTCTATCAAT TTAATTTATTATGTTACTTTATATTTGATTAATAATTGCAAGTTTAAAATCATAATTTAATGTTGAAAGGCC ACTATTCTAATTAACTTAAGGAGTTGTTTATTTATGCAAAATAACAACTTTAATACCACAGAAATTAATAAT ATGATTAATTTCCCTATGTATAATGGTAGATTAGAACCTTCTCTAGCTCCAGCATTAATAGCAGTAGCTC CAATTGCTAAATATTTAGCAACAGCTCTTGCTAAATGGGCTGTAAAACAAGGGTTTGCAAAATTAAAATC CGAGATATTCCCCGGTAATACGCCTGCTACTATGGATAAGGTTCGTATTGAGGTACAAACACTTTTAGA CCAAAGATTACAAGATGACAGAGTTAAGATTTTAGAAGGTGAATACAAAGGAATTATTGACGTGAGTAAA GTTTTTACTGATTATGTTAATCAATCTAAATTTGAGACTGGAACAGCTAATAGGCTTTTTTTTGATACAAG TAACCAATTAATAAGCAGATTGCCTCAATTTGAGATTGCAGGATATGAAGGAGTATCCATTTCACTTTTT ACTCAGATGTGTACATTTCATTTGGGTTTATTAAAAGATGGAATTTTAGCAGGAAGCGATTGGGGATTTG CTCCTGCAGATAAAGACGCTCTTATTTGCCAATTTAATAGATTTGTCAATGAATATAATACTCGACTGAT GGTATTGTACTCAAAAGAATTTGGACGGTTATTAGCAAAAAATCTTAATGAAGCCTTGAACTTTAGAAAT ATGTGTAGTTTATATGTCTTTCCTTTTTCTGAAGCATGGTCTTTATTAAGGTATGAAGGAACAAAATTAGA AAACACGCTTTCATTATGGAATTTTGTGGGTGAAAGTATCAATAATATATCTCCTAATGATTGGAAAGGT GCGCTTTATAAATTGTTAATGGGAGCACCTAATCAAAGATTAAACAATGTTAAGTTTAATTATAGTTATTT TTCTGATACTCAAGCGACAATACATCGTGAAAACATTCATGGTGTCCTGCCAACATATAATGGAGGACC AACAATTACAGGATGGATAGGGAATGGGCGTTTCAGCGGACTTAGTTTTCCTTGTAGTAATGAATTAGA AATTACAAAAATAAAACAGGAAATAACTTACAATGATAAAGGGGGAAATTTCAATTCAATAGTTCCTGCT GCTACGCGCAATGAAATTCTAACTGCTACCGTTCCAACATCAGCTGATCCATTTTTTAAAACCGCTGATA TTAACTGGAAATATTTCTCTCCGGGTCTTTACTCTGGATGGAATATTAAATTTGATGATACAGTCACTTTA AAAAGTAGAGTACCAAGTATTATACCTTCAAATATATTAAAGTATGATGATTATTATATTCGTGCCGTTTC AGCCTGTCCAAAAGGCGTATCACTTGCATATAACCATGATTTTTTAACGTTAACATT**T**TAACAAATTAGAAT ATGATGCACCTACTACACAAAATATCATTGTAGGATTTTCACCAGATAATACTAAGAGTTTTTATAGGAG CAACTCTCATTATCTAAGTACAACAGATGATGCCTATGTAATTCCTGCTTTACAATTTTCTACAGTCTCAG ATAGATCATTCTTAGAAGATACACCAGATCAAGCAACAGATGGCAGTATTAAATTTACGGATACTGTTCT TGGGAATGAGGCAAAATATTCTATTAGACTAAATACTGGATTTAATACAGCTACTAGGTATAGATTAATTA TACGTTTTAAAGCGCCTGCTCGTTTGGCTGCTGGTATACGTGTACGTTCTCAAAATTCAGGGAATAATAA GTTATTAGGTGGTATTCCTGTAGAGGGTAATTCTGGATGGATAGATTATATTACAGATTCATTTACTTTTG ATGACCTTGGGATTACAACTTCAAGTACAAATGCTTTCTTTAGTATTGATTCAGATGGTGTAAATGCTTCT CAACAATGGTATTTGTCTAAATTAATTTTAGTAAAAGAATCCAGTTTTACGACTCAGATTCCATTAAAACC ATACGTTATTGTACGTTGTCCGGATACTTTTTTTGTGAGCAACAATTCAAGTAGTACGTACGAACAAGGC TATAACAACAATTACAACCAGAATTCTAGCAGTATGTACGATCAAGGATATAACAATAGCTATAATCCAA ACTCTGGTTGTACGTGTAATCAAGACTATAATAACAGTTATAACCAAAACTCTGGCTGTACATGTAACCA AGGGTATAACAATAACTATCCTAAATAATCTTAGTAGCTATATTTATTAAATATGGTAATATCACAAGTAT AAATACTTGTGGTATTACCTACCATTCTTAAATTATATCCAAAATCATGCGTTAATCTACATTCCCCTTTC TCTAAAATTTGTTCTTCACACATCCACATTTTTCGACTCGAGGCATGCAAGCTTGGC

*The inserted mutation is indicated as a capital bold letter and underlined.

Appendix B

PBDePISA Interfaces



interface # 1 in PDB 7qyd crystal.

Space symmetry group: P 21 21 2. Resolution: 2.40 Å

MOSQUITOCIDAL CRY11BA DETERMINED AT PH 6.5 FROM NATURALLY-OCCURRING NANOCRYSTALS BY SERIAL FEMTOSECOND CRYSTALLOGRAPHY

interface #1/10

Interface Summary

	Structure 1		Structure 2	
Selection range	B		A	
class	Protein		Protein	
symmetry operation	x,y,z		x-1/2,-y+1/2,-z	
symmetry ID	1_555		2_455	
Number of atoms				
interface	166	3.3%	147	2.9%
surface	2566	51.3%	2616	51.7%
total	5003	100.0%	5061	100.0%
Number of residues				
interface	47	7.5%	44	6.9%
surface	570	90.5%	578	90.7%
total	630	100.0%	637	100.0%
Solvent-accessible area, Å				
interface	1535.1	5.8%	1543.2	5.7%
total	26683.4	100.0%	27190.0	100.0%
Solvation energy, kcal/mol				
isolated structure	-606.4	100.0%	-610.5	100.0%
gain on complex formation	-10.3	1.7%	-9.0	1.5%
average gain	-3.5	0.6%	-3.4	0.6%
P-value	0.034		0.057	

This interface scored

0.000

in Complex Formation Significance Score (CSS).

CSS ranges from 0 to 1 as interface relevance to complex formation increases.

Achieved CSS implies that the interface does not play any role in complex formation and seems to be a result of crystal packing only

Hydrogen bonds

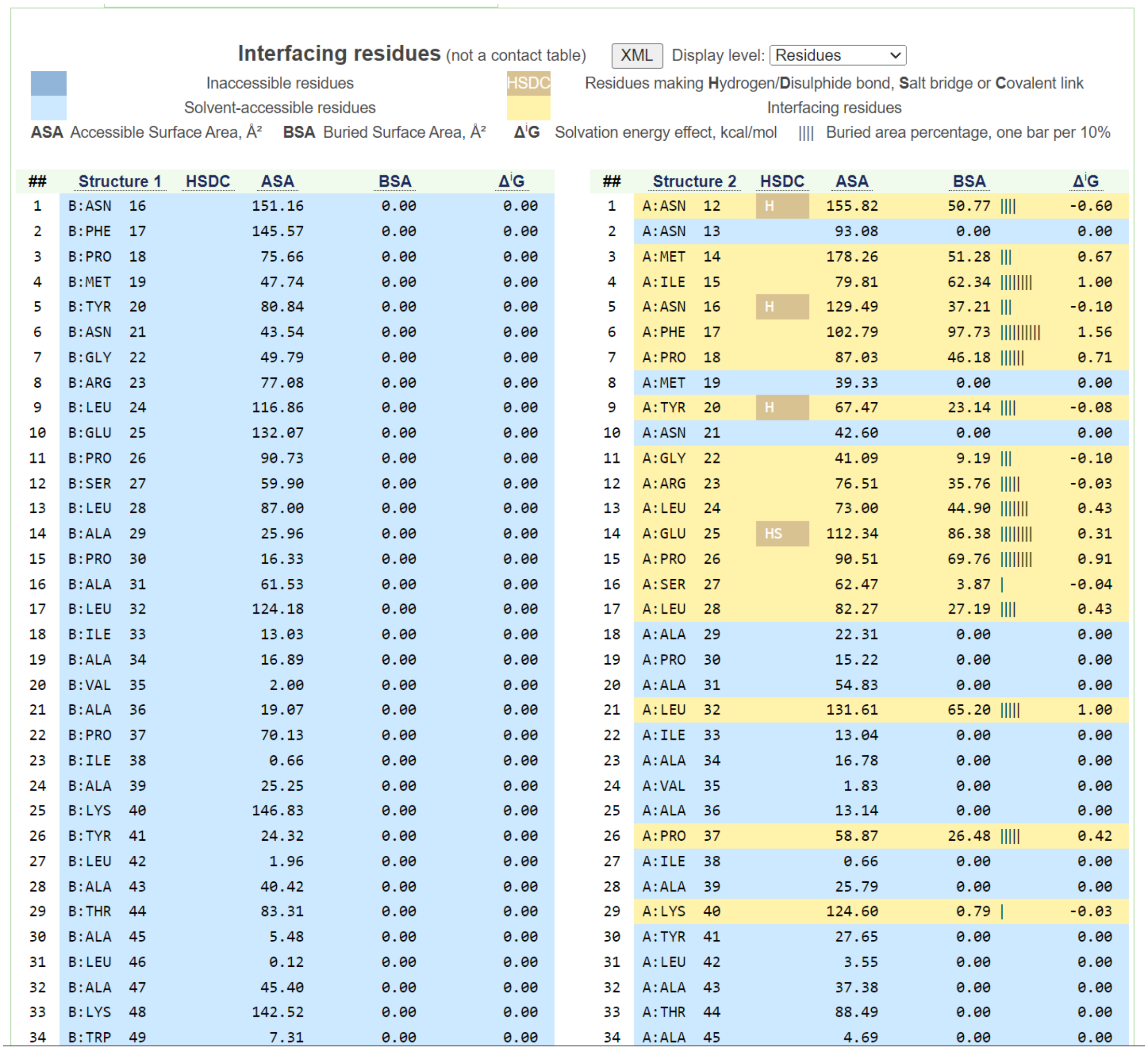
##	Structure 1	Dist. [Å]	Structure 2
1	B:ASN 131[ND2]	3.40	A:TYR 20[OH]
2	B:GLN 187[NE2]	2.74	A:GLU 25[OE2]
3	B:THR 655[N]	3.79	A:ASP 83[OD2]
4	B:SER 586[OG]	3.54	A:ASP 88[OD1]
5	B:SER 586[N]	3.86	A:ASP 88[OD2]
6	B:ASN 266[ND2]	3.49	A:ASP 103[OD1]
7	B:ASP 654[O]	3.61	A:ASN 12[ND2]
8	B:ASP 180[OD1]	2.79	A:ASN 16[N]
9	B:ASN 262[O]	2.90	A:LYS 106[NZ]

Salt bridges

##	Structure 1	Dist. [Å]	Structure 2
1	B:ARG 190[NH1]	3.24	A:GLU 25[OE1]
2	B:ARG 190[NH1]	3.43	A:GLU 25[OE2]

No disulfide bonds found

No covalent bonds found



Interfacing residues (not a contact table) Display level: Residues

Inaccessible residues

Solvent-accessible residues

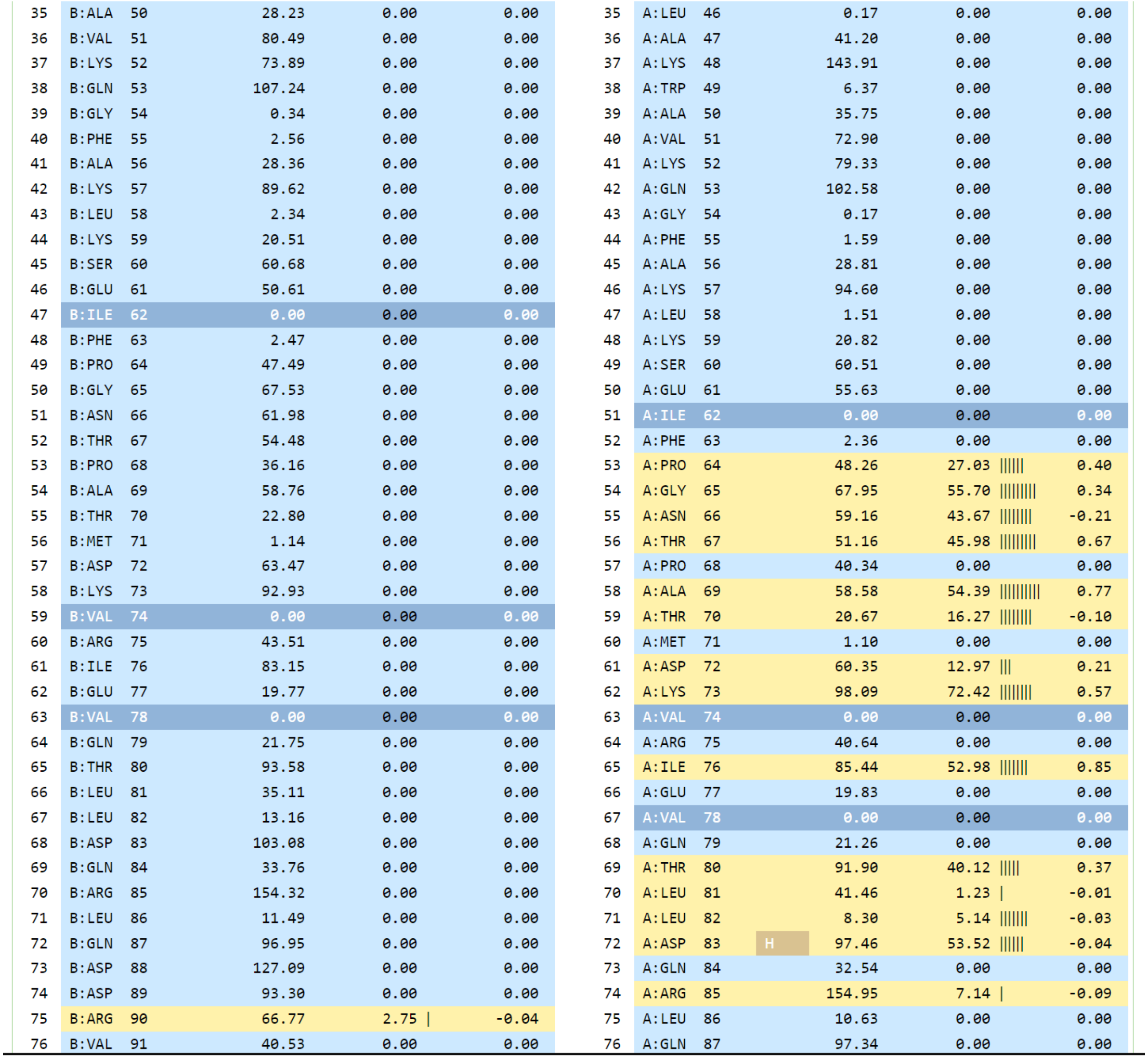
HSDC Residues making **H**ydrogen/**D**isulphide bond, **S**alt bridge or **C**ovalent link

Interfacing residues

ASA Accessible Surface Area, Å² **BSA** Buried Surface Area, Å² **Δ^{i}G** Solvation energy effect, kcal/mol |||| Buried area percentage, one bar per 10%

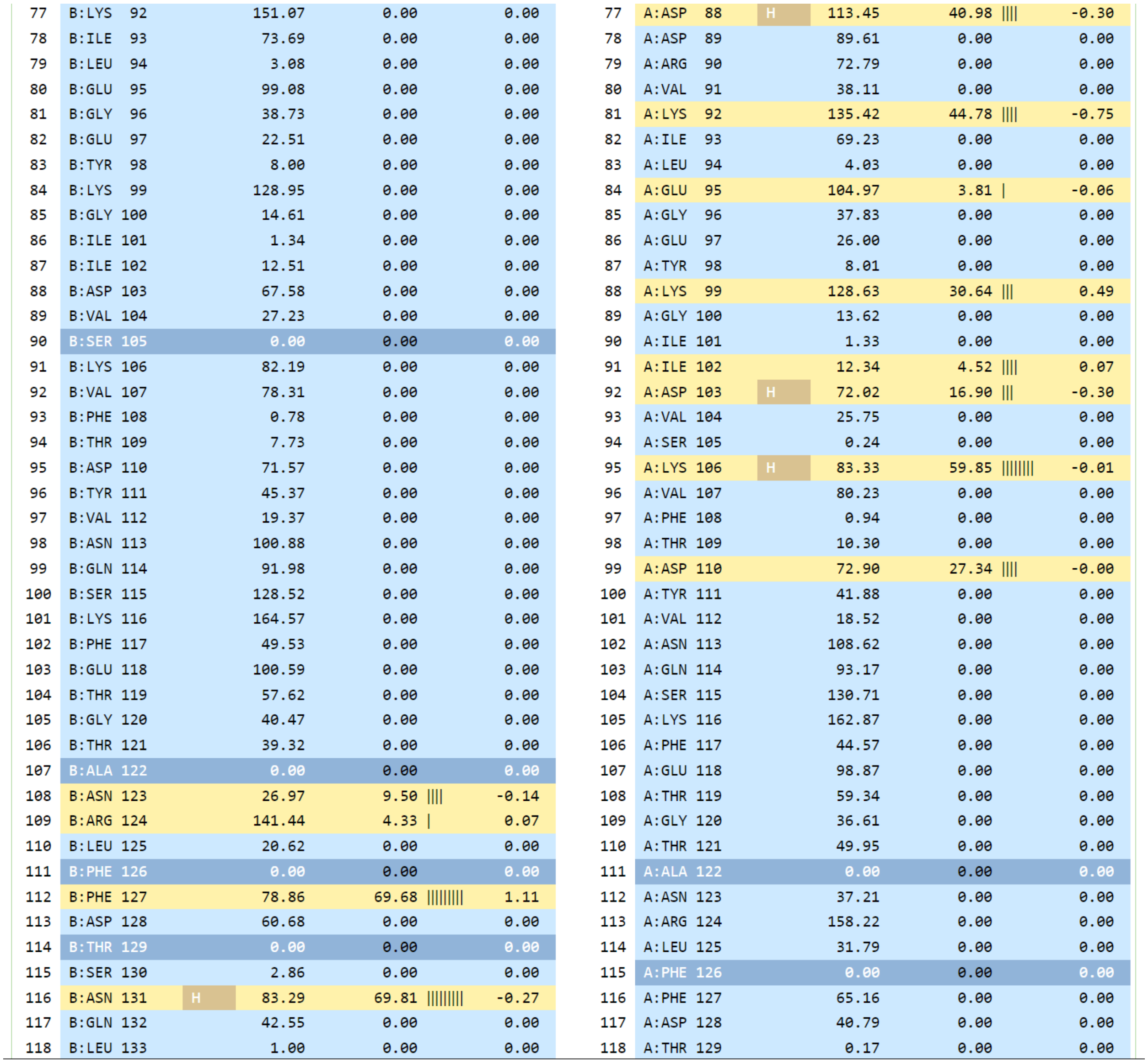
##	Structure 1		HSDC	ASA	BSA	Δ^{i}G
1	B:ASN	16		151.16	0.00	0.00
2	B:PHE	17		145.57	0.00	0.00
3	B:PRO	18		75.66	0.00	0.00
4	B:MET	19		47.74	0.00	0.00
5	B:TYR	20		80.84	0.00	0.00
6	B:ASN	21		43.54	0.00	0.00
7	B:GLY	22		49.79	0.00	0.00
8	B:ARG	23		77.08	0.00	0.00
9	B:LEU	24		116.86	0.00	0.00
10	B:GLU	25		132.07	0.00	0.00
11	B:PRO	26		90.73	0.00	0.00
12	B:SER	27		59.90	0.00	0.00
13	B:LEU	28		87.00	0.00	0.00
14	B:ALA	29		25.96	0.00	0.00
15	B:PRO	30		16.33	0.00	0.00
16	B:ALA	31		61.53	0.00	0.00
17	B:LEU	32		124.18	0.00	0.00
18	B:ILE	33		13.03	0.00	0.00
19	B:ALA	34		16.89	0.00	0.00
20	B:VAL	35		2.00	0.00	0.00
21	B:ALA	36		19.07	0.00	0.00
22	B:PRO	37		70.13	0.00	0.00
23	B:ILE	38		0.66	0.00	0.00
24	B:ALA	39		25.25	0.00	0.00
25	B:LYS	40		146.83	0.00	0.00
26	B:TYR	41		24.32	0.00	0.00
27	B:LEU	42		1.96	0.00	0.00
28	B:ALA	43		40.42	0.00	0.00
29	B:THR	44		83.31	0.00	0.00
30	B:ALA	45		5.48	0.00	0.00
31	B:LEU	46		0.12	0.00	0.00
32	B:ALA	47		45.40	0.00	0.00
33	B:LYS	48		142.52	0.00	0.00
34	B:TRP	49		7.31	0.00	0.00

##	Structure 2		HSDC	ASA	BSA	Δ^{i}G
1	A:ASN	12	H	155.82	50.77 ||||	-0.60
2	A:ASN	13		93.08	0.00	0.00
3	A:MET	14		178.26	51.28 |||	0.67
4	A:ILE	15		79.81	62.34 ||||||||	1.00
5	A:ASN	16	H	129.49	37.21 |||	-0.10
6	A:PHE	17		102.79	97.73 ||||||||||	1.56
7	A:PRO	18		87.03	46.18 ||||||	0.71
8	A:MET	19		39.33	0.00	0.00
9	A:TYR	20	H	67.47	23.14 ||||	-0.08
10	A:ASN	21		42.60	0.00	0.00
11	A:GLY	22		41.09	9.19 |||	-0.10
12	A:ARG	23		76.51	35.76 |||||	-0.03
13	A:LEU	24		73.00	44.90 |||||||	0.43
14	A:GLU	25	HS	112.34	86.38 ||||||||	0.31
15	A:PRO	26		90.51	69.76 ||||||||	0.91
16	A:SER	27		62.47	3.87 |	-0.04
17	A:LEU	28		82.27	27.19 ||||	0.43
18	A:ALA	29		22.31	0.00	0.00
19	A:PRO	30		15.22	0.00	0.00
20	A:ALA	31		54.83	0.00	0.00
21	A:LEU	32		131.61	65.20 |||||	1.00
22	A:ILE	33		13.04	0.00	0.00
23	A:ALA	34		16.78	0.00	0.00
24	A:VAL	35		1.83	0.00	0.00
25	A:ALA	36		13.14	0.00	0.00
26	A:PRO	37		58.87	26.48 |||||	0.42
27	A:ILE	38		0.66	0.00	0.00
28	A:ALA	39		25.79	0.00	0.00
29	A:LYS	40		124.60	0.79 |	-0.03
30	A:TYR	41		27.65	0.00	0.00
31	A:LEU	42		3.55	0.00	0.00
32	A:ALA	43		37.38	0.00	0.00
33	A:THR	44		88.49	0.00	0.00
34	A:ALA	45		4.69	0.00	0.00

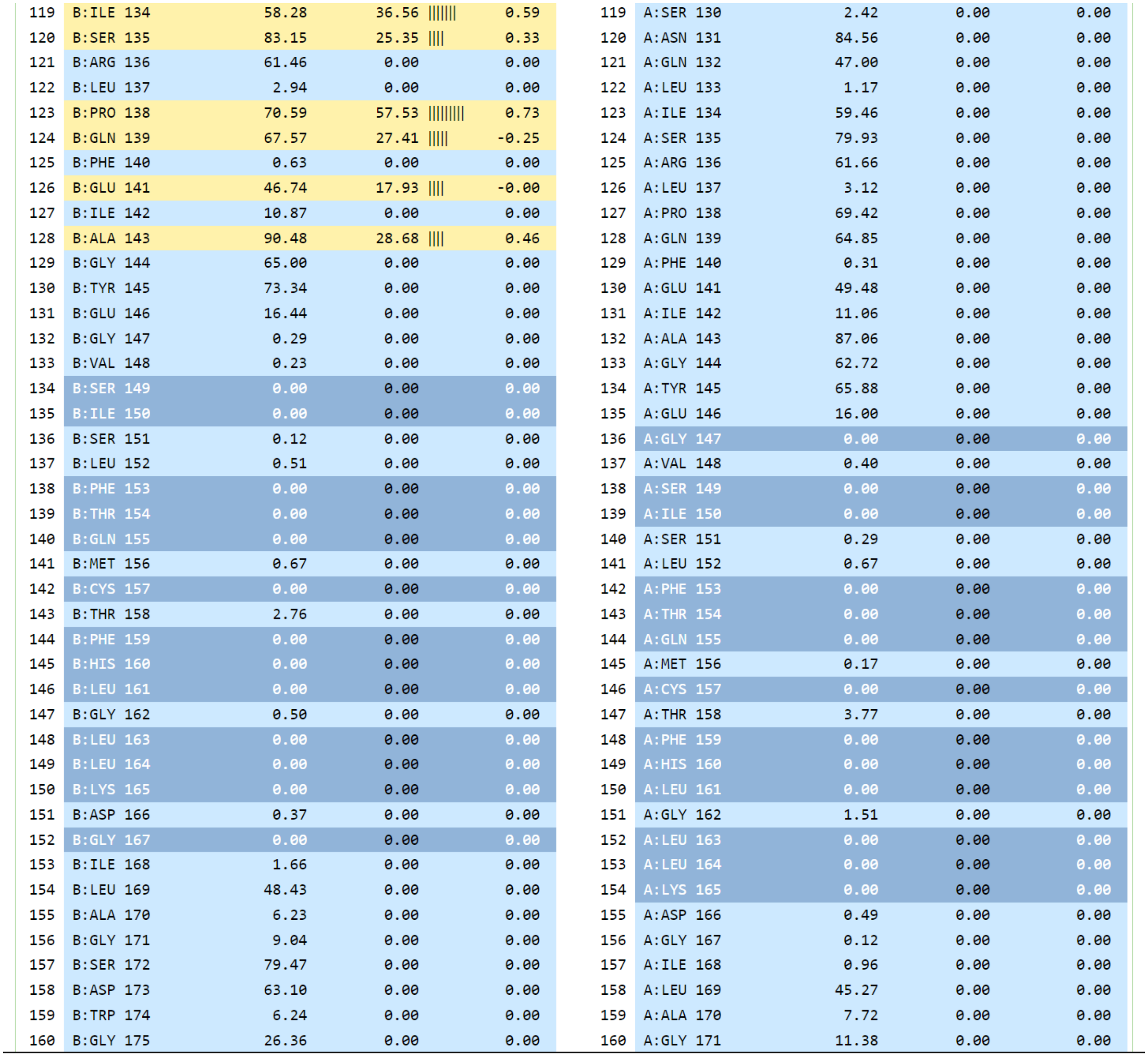


<table>
<tr><td>35</td><td>B:ALA</td><td>50</td><td></td><td>28.23</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>36</td><td>B:VAL</td><td>51</td><td></td><td>80.49</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>37</td><td>B:LYS</td><td>52</td><td></td><td>73.89</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>38</td><td>B:GLN</td><td>53</td><td></td><td>107.24</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>39</td><td>B:GLY</td><td>54</td><td></td><td>0.34</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>40</td><td>B:PHE</td><td>55</td><td></td><td>2.56</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>41</td><td>B:ALA</td><td>56</td><td></td><td>28.36</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>42</td><td>B:LYS</td><td>57</td><td></td><td>89.62</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>43</td><td>B:LEU</td><td>58</td><td></td><td>2.34</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>44</td><td>B:LYS</td><td>59</td><td></td><td>20.51</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>45</td><td>B:SER</td><td>60</td><td></td><td>60.68</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>46</td><td>B:GLU</td><td>61</td><td></td><td>50.61</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>47</td><td>B:ILE</td><td>62</td><td></td><td>0.00</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>48</td><td>B:PHE</td><td>63</td><td></td><td>2.47</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>49</td><td>B:PRO</td><td>64</td><td></td><td>47.49</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>50</td><td>B:GLY</td><td>65</td><td></td><td>67.53</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>51</td><td>B:ASN</td><td>66</td><td></td><td>61.98</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>52</td><td>B:THR</td><td>67</td><td></td><td>54.48</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>53</td><td>B:PRO</td><td>68</td><td></td><td>36.16</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>54</td><td>B:ALA</td><td>69</td><td></td><td>58.76</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>55</td><td>B:THR</td><td>70</td><td></td><td>22.80</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>56</td><td>B:MET</td><td>71</td><td></td><td>1.14</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>57</td><td>B:ASP</td><td>72</td><td></td><td>63.47</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>58</td><td>B:LYS</td><td>73</td><td></td><td>92.93</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>59</td><td>B:VAL</td><td>74</td><td></td><td>0.00</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>60</td><td>B:ARG</td><td>75</td><td></td><td>43.51</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>61</td><td>B:ILE</td><td>76</td><td></td><td>83.15</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>62</td><td>B:GLU</td><td>77</td><td></td><td>19.77</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>63</td><td>B:VAL</td><td>78</td><td></td><td>0.00</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>64</td><td>B:GLN</td><td>79</td><td></td><td>21.75</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>65</td><td>B:THR</td><td>80</td><td></td><td>93.58</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>66</td><td>B:LEU</td><td>81</td><td></td><td>35.11</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>67</td><td>B:LEU</td><td>82</td><td></td><td>13.16</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>68</td><td>B:ASP</td><td>83</td><td></td><td>103.08</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>69</td><td>B:GLN</td><td>84</td><td></td><td>33.76</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>70</td><td>B:ARG</td><td>85</td><td></td><td>154.32</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>71</td><td>B:LEU</td><td>86</td><td></td><td>11.49</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>72</td><td>B:GLN</td><td>87</td><td></td><td>96.95</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>73</td><td>B:ASP</td><td>88</td><td></td><td>127.09</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>74</td><td>B:ASP</td><td>89</td><td></td><td>93.30</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>75</td><td>B:ARG</td><td>90</td><td></td><td>66.77</td><td>2.75</td><td>|</td><td>-0.04</td></tr>
<tr><td>76</td><td>B:VAL</td><td>91</td><td></td><td>40.53</td><td>0.00</td><td></td><td>0.00</td></tr>
</table>

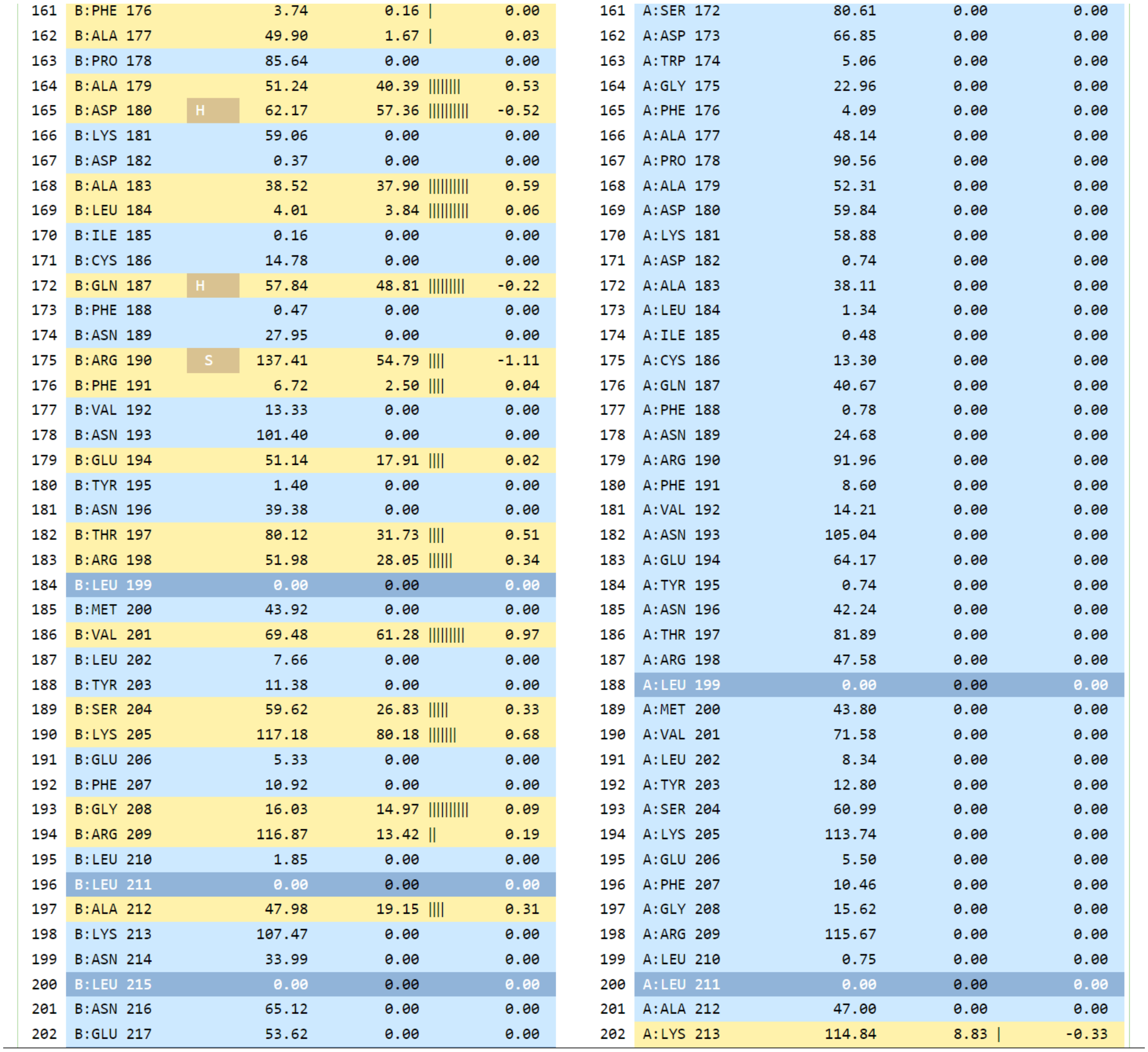
<table>
<tr><td>35</td><td>A:LEU</td><td>46</td><td></td><td>0.17</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>36</td><td>A:ALA</td><td>47</td><td></td><td>41.20</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>37</td><td>A:LYS</td><td>48</td><td></td><td>143.91</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>38</td><td>A:TRP</td><td>49</td><td></td><td>6.37</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>39</td><td>A:ALA</td><td>50</td><td></td><td>35.75</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>40</td><td>A:VAL</td><td>51</td><td></td><td>72.90</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>41</td><td>A:LYS</td><td>52</td><td></td><td>79.33</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>42</td><td>A:GLN</td><td>53</td><td></td><td>102.58</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>43</td><td>A:GLY</td><td>54</td><td></td><td>0.17</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>44</td><td>A:PHE</td><td>55</td><td></td><td>1.59</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>45</td><td>A:ALA</td><td>56</td><td></td><td>28.81</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>46</td><td>A:LYS</td><td>57</td><td></td><td>94.60</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>47</td><td>A:LEU</td><td>58</td><td></td><td>1.51</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>48</td><td>A:LYS</td><td>59</td><td></td><td>20.82</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>49</td><td>A:SER</td><td>60</td><td></td><td>60.51</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>50</td><td>A:GLU</td><td>61</td><td></td><td>55.63</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>51</td><td>A:ILE</td><td>62</td><td></td><td>0.00</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>52</td><td>A:PHE</td><td>63</td><td></td><td>2.36</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>53</td><td>A:PRO</td><td>64</td><td></td><td>48.26</td><td>27.03</td><td>||||||</td><td>0.40</td></tr>
<tr><td>54</td><td>A:GLY</td><td>65</td><td></td><td>67.95</td><td>55.70</td><td>|||||||||</td><td>0.34</td></tr>
<tr><td>55</td><td>A:ASN</td><td>66</td><td></td><td>59.16</td><td>43.67</td><td>||||||||</td><td>-0.21</td></tr>
<tr><td>56</td><td>A:THR</td><td>67</td><td></td><td>51.16</td><td>45.98</td><td>|||||||||</td><td>0.67</td></tr>
<tr><td>57</td><td>A:PRO</td><td>68</td><td></td><td>40.34</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>58</td><td>A:ALA</td><td>69</td><td></td><td>58.58</td><td>54.39</td><td>||||||||||</td><td>0.77</td></tr>
<tr><td>59</td><td>A:THR</td><td>70</td><td></td><td>20.67</td><td>16.27</td><td>||||||||</td><td>-0.10</td></tr>
<tr><td>60</td><td>A:MET</td><td>71</td><td></td><td>1.10</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>61</td><td>A:ASP</td><td>72</td><td></td><td>60.35</td><td>12.97</td><td>|||</td><td>0.21</td></tr>
<tr><td>62</td><td>A:LYS</td><td>73</td><td></td><td>98.09</td><td>72.42</td><td>||||||||</td><td>0.57</td></tr>
<tr><td>63</td><td>A:VAL</td><td>74</td><td></td><td>0.00</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>64</td><td>A:ARG</td><td>75</td><td></td><td>40.64</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>65</td><td>A:ILE</td><td>76</td><td></td><td>85.44</td><td>52.98</td><td>|||||||</td><td>0.85</td></tr>
<tr><td>66</td><td>A:GLU</td><td>77</td><td></td><td>19.83</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>67</td><td>A:VAL</td><td>78</td><td></td><td>0.00</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>68</td><td>A:GLN</td><td>79</td><td></td><td>21.26</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>69</td><td>A:THR</td><td>80</td><td></td><td>91.90</td><td>40.12</td><td>|||||</td><td>0.37</td></tr>
<tr><td>70</td><td>A:LEU</td><td>81</td><td></td><td>41.46</td><td>1.23</td><td>|</td><td>-0.01</td></tr>
<tr><td>71</td><td>A:LEU</td><td>82</td><td></td><td>8.30</td><td>5.14</td><td>|||||||</td><td>-0.03</td></tr>
<tr><td>72</td><td>A:ASP</td><td>83</td><td>H</td><td>97.46</td><td>53.52</td><td>||||||</td><td>-0.04</td></tr>
<tr><td>73</td><td>A:GLN</td><td>84</td><td></td><td>32.54</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>74</td><td>A:ARG</td><td>85</td><td></td><td>154.95</td><td>7.14</td><td>|</td><td>-0.09</td></tr>
<tr><td>75</td><td>A:LEU</td><td>86</td><td></td><td>10.63</td><td>0.00</td><td></td><td>0.00</td></tr>
<tr><td>76</td><td>A:GLN</td><td>87</td><td></td><td>97.34</td><td>0.00</td><td></td><td>0.00</td></tr>
</table>



77	B:LYS 92		151.07	0.00	0.00	77	A:ASP 88	H	113.45	40.98 ||||	-0.30
78	B:ILE 93		73.69	0.00	0.00	78	A:ASP 89		89.61	0.00	0.00
79	B:LEU 94		3.08	0.00	0.00	79	A:ARG 90		72.79	0.00	0.00
80	B:GLU 95		99.08	0.00	0.00	80	A:VAL 91		38.11	0.00	0.00
81	B:GLY 96		38.73	0.00	0.00	81	A:LYS 92		135.42	44.78 ||||	-0.75
82	B:GLU 97		22.51	0.00	0.00	82	A:ILE 93		69.23	0.00	0.00
83	B:TYR 98		8.00	0.00	0.00	83	A:LEU 94		4.03	0.00	0.00
84	B:LYS 99		128.95	0.00	0.00	84	A:GLU 95		104.97	3.81 |	-0.06
85	B:GLY 100		14.61	0.00	0.00	85	A:GLY 96		37.83	0.00	0.00
86	B:ILE 101		1.34	0.00	0.00	86	A:GLU 97		26.00	0.00	0.00
87	B:ILE 102		12.51	0.00	0.00	87	A:TYR 98		8.01	0.00	0.00
88	B:ASP 103		67.58	0.00	0.00	88	A:LYS 99		128.63	30.64 |||	0.49
89	B:VAL 104		27.23	0.00	0.00	89	A:GLY 100		13.62	0.00	0.00
90	B:SER 105		0.00	0.00	0.00	90	A:ILE 101		1.33	0.00	0.00
91	B:LYS 106		82.19	0.00	0.00	91	A:ILE 102		12.34	4.52 ||||	0.07
92	B:VAL 107		78.31	0.00	0.00	92	A:ASP 103	H	72.02	16.90 |||	-0.30
93	B:PHE 108		0.78	0.00	0.00	93	A:VAL 104		25.75	0.00	0.00
94	B:THR 109		7.73	0.00	0.00	94	A:SER 105		0.24	0.00	0.00
95	B:ASP 110		71.57	0.00	0.00	95	A:LYS 106	H	83.33	59.85 ||||||||	-0.01
96	B:TYR 111		45.37	0.00	0.00	96	A:VAL 107		80.23	0.00	0.00
97	B:VAL 112		19.37	0.00	0.00	97	A:PHE 108		0.94	0.00	0.00
98	B:ASN 113		100.88	0.00	0.00	98	A:THR 109		10.30	0.00	0.00
99	B:GLN 114		91.98	0.00	0.00	99	A:ASP 110		72.90	27.34 ||||	-0.00
100	B:SER 115		128.52	0.00	0.00	100	A:TYR 111		41.88	0.00	0.00
101	B:LYS 116		164.57	0.00	0.00	101	A:VAL 112		18.52	0.00	0.00
102	B:PHE 117		49.53	0.00	0.00	102	A:ASN 113		108.62	0.00	0.00
103	B:GLU 118		100.59	0.00	0.00	103	A:GLN 114		93.17	0.00	0.00
104	B:THR 119		57.62	0.00	0.00	104	A:SER 115		130.71	0.00	0.00
105	B:GLY 120		40.47	0.00	0.00	105	A:LYS 116		162.87	0.00	0.00
106	B:THR 121		39.32	0.00	0.00	106	A:PHE 117		44.57	0.00	0.00
107	B:ALA 122		0.00	0.00	0.00	107	A:GLU 118		98.87	0.00	0.00
108	B:ASN 123		26.97	9.50 ||||	-0.14	108	A:THR 119		59.34	0.00	0.00
109	B:ARG 124		141.44	4.33 |	0.07	109	A:GLY 120		36.61	0.00	0.00
110	B:LEU 125		20.62	0.00	0.00	110	A:THR 121		49.95	0.00	0.00
111	B:PHE 126		0.00	0.00	0.00	111	A:ALA 122		0.00	0.00	0.00
112	B:PHE 127		78.86	69.68 |||||||||	1.11	112	A:ASN 123		37.21	0.00	0.00
113	B:ASP 128		60.68	0.00	0.00	113	A:ARG 124		158.22	0.00	0.00
114	B:THR 129		0.00	0.00	0.00	114	A:LEU 125		31.79	0.00	0.00
115	B:SER 130		2.86	0.00	0.00	115	A:PHE 126		0.00	0.00	0.00
116	B:ASN 131	H	83.29	69.81 |||||||||	-0.27	116	A:PHE 127		65.16	0.00	0.00
117	B:GLN 132		42.55	0.00	0.00	117	A:ASP 128		40.79	0.00	0.00
118	B:LEU 133		1.00	0.00	0.00	118	A:THR 129		0.17	0.00	0.00

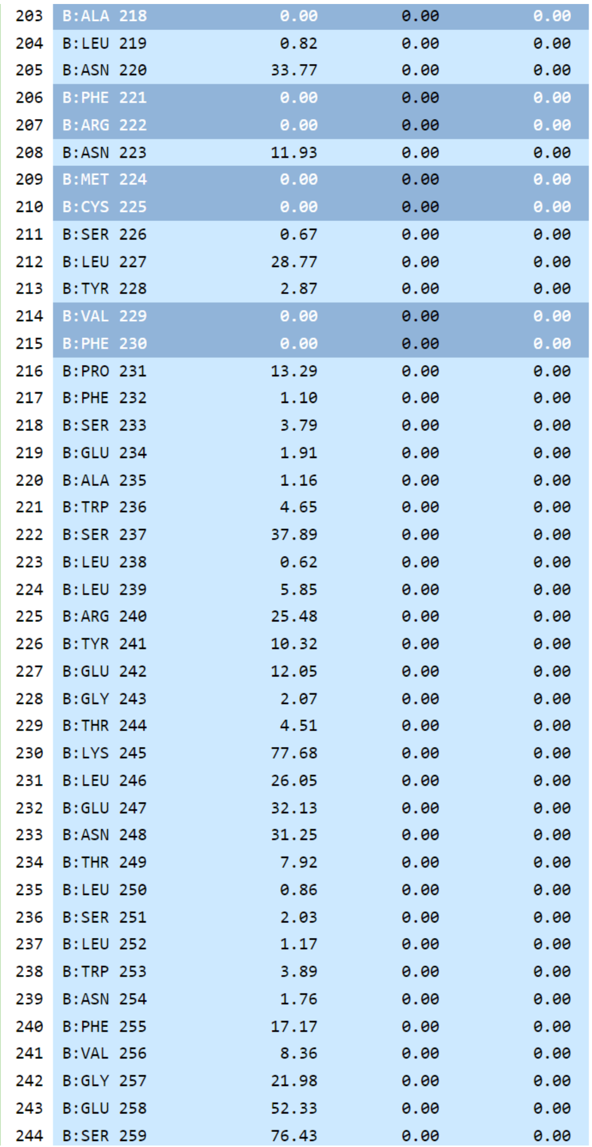


119	B:ILE 134	58.28	36.56	|||||||	0.59	119	A:SER 130	2.42	0.00	0.00
120	B:SER 135	83.15	25.35	||||	0.33	120	A:ASN 131	84.56	0.00	0.00
121	B:ARG 136	61.46	0.00		0.00	121	A:GLN 132	47.00	0.00	0.00
122	B:LEU 137	2.94	0.00		0.00	122	A:LEU 133	1.17	0.00	0.00
123	B:PRO 138	70.59	57.53	|||||||||	0.73	123	A:ILE 134	59.46	0.00	0.00
124	B:GLN 139	67.57	27.41	|||||	-0.25	124	A:SER 135	79.93	0.00	0.00
125	B:PHE 140	0.63	0.00		0.00	125	A:ARG 136	61.66	0.00	0.00
126	B:GLU 141	46.74	17.93	||||	-0.00	126	A:LEU 137	3.12	0.00	0.00
127	B:ILE 142	10.87	0.00		0.00	127	A:PRO 138	69.42	0.00	0.00
128	B:ALA 143	90.48	28.68	||||	0.46	128	A:GLN 139	64.85	0.00	0.00
129	B:GLY 144	65.00	0.00		0.00	129	A:PHE 140	0.31	0.00	0.00
130	B:TYR 145	73.34	0.00		0.00	130	A:GLU 141	49.48	0.00	0.00
131	B:GLU 146	16.44	0.00		0.00	131	A:ILE 142	11.06	0.00	0.00
132	B:GLY 147	0.29	0.00		0.00	132	A:ALA 143	87.06	0.00	0.00
133	B:VAL 148	0.23	0.00		0.00	133	A:GLY 144	62.72	0.00	0.00
134	B:SER 149	0.00	0.00		0.00	134	A:TYR 145	65.88	0.00	0.00
135	B:ILE 150	0.00	0.00		0.00	135	A:GLU 146	16.00	0.00	0.00
136	B:SER 151	0.12	0.00		0.00	136	A:GLY 147	0.00	0.00	0.00
137	B:LEU 152	0.51	0.00		0.00	137	A:VAL 148	0.40	0.00	0.00
138	B:PHE 153	0.00	0.00		0.00	138	A:SER 149	0.00	0.00	0.00
139	B:THR 154	0.00	0.00		0.00	139	A:ILE 150	0.00	0.00	0.00
140	B:GLN 155	0.00	0.00		0.00	140	A:SER 151	0.29	0.00	0.00
141	B:MET 156	0.67	0.00		0.00	141	A:LEU 152	0.67	0.00	0.00
142	B:CYS 157	0.00	0.00		0.00	142	A:PHE 153	0.00	0.00	0.00
143	B:THR 158	2.76	0.00		0.00	143	A:THR 154	0.00	0.00	0.00
144	B:PHE 159	0.00	0.00		0.00	144	A:GLN 155	0.00	0.00	0.00
145	B:HIS 160	0.00	0.00		0.00	145	A:MET 156	0.17	0.00	0.00
146	B:LEU 161	0.00	0.00		0.00	146	A:CYS 157	0.00	0.00	0.00
147	B:GLY 162	0.50	0.00		0.00	147	A:THR 158	3.77	0.00	0.00
148	B:LEU 163	0.00	0.00		0.00	148	A:PHE 159	0.00	0.00	0.00
149	B:LEU 164	0.00	0.00		0.00	149	A:HIS 160	0.00	0.00	0.00
150	B:LYS 165	0.00	0.00		0.00	150	A:LEU 161	0.00	0.00	0.00
151	B:ASP 166	0.37	0.00		0.00	151	A:GLY 162	1.51	0.00	0.00
152	B:GLY 167	0.00	0.00		0.00	152	A:LEU 163	0.00	0.00	0.00
153	B:ILE 168	1.66	0.00		0.00	153	A:LEU 164	0.00	0.00	0.00
154	B:LEU 169	48.43	0.00		0.00	154	A:LYS 165	0.00	0.00	0.00
155	B:ALA 170	6.23	0.00		0.00	155	A:ASP 166	0.49	0.00	0.00
156	B:GLY 171	9.04	0.00		0.00	156	A:GLY 167	0.12	0.00	0.00
157	B:SER 172	79.47	0.00		0.00	157	A:ILE 168	0.96	0.00	0.00
158	B:ASP 173	63.10	0.00		0.00	158	A:LEU 169	45.27	0.00	0.00
159	B:TRP 174	6.24	0.00		0.00	159	A:ALA 170	7.72	0.00	0.00
160	B:GLY 175	26.36	0.00		0.00	160	A:GLY 171	11.38	0.00	0.00

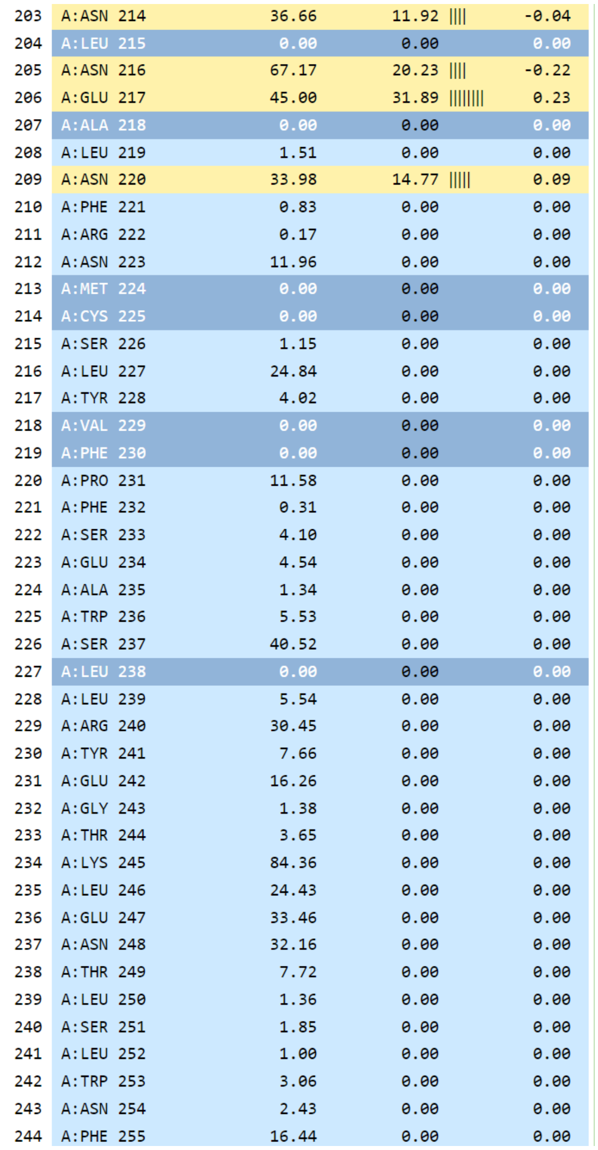


161	B:PHE 176		3.74	0.16	|	0.00
162	B:ALA 177		49.90	1.67	|	0.03
163	B:PRO 178		85.64	0.00		0.00
164	B:ALA 179		51.24	40.39	||||||||	0.53
165	B:ASP 180	H	62.17	57.36	||||||||||	-0.52
166	B:LYS 181		59.06	0.00		0.00
167	B:ASP 182		0.37	0.00		0.00
168	B:ALA 183		38.52	37.90	||||||||||	0.59
169	B:LEU 184		4.01	3.84	||||||||||	0.06
170	B:ILE 185		0.16	0.00		0.00
171	B:CYS 186		14.78	0.00		0.00
172	B:GLN 187	H	57.84	48.81	|||||||||	-0.22
173	B:PHE 188		0.47	0.00		0.00
174	B:ASN 189		27.95	0.00		0.00
175	B:ARG 190	S	137.41	54.79	||||	-1.11
176	B:PHE 191		6.72	2.50	||||	0.04
177	B:VAL 192		13.33	0.00		0.00
178	B:ASN 193		101.40	0.00		0.00
179	B:GLU 194		51.14	17.91	||||	0.02
180	B:TYR 195		1.40	0.00		0.00
181	B:ASN 196		39.38	0.00		0.00
182	B:THR 197		80.12	31.73	||||	0.51
183	B:ARG 198		51.98	28.05	||||||	0.34
184	B:LEU 199		0.00	0.00		0.00
185	B:MET 200		43.92	0.00		0.00
186	B:VAL 201		69.48	61.28	|||||||||	0.97
187	B:LEU 202		7.66	0.00		0.00
188	B:TYR 203		11.38	0.00		0.00
189	B:SER 204		59.62	26.83	|||||	0.33
190	B:LYS 205		117.18	80.18	|||||||	0.68
191	B:GLU 206		5.33	0.00		0.00
192	B:PHE 207		10.92	0.00		0.00
193	B:GLY 208		16.03	14.97	||||||||||	0.09
194	B:ARG 209		116.87	13.42	||	0.19
195	B:LEU 210		1.85	0.00		0.00
196	B:LEU 211		0.00	0.00		0.00
197	B:ALA 212		47.98	19.15	||||	0.31
198	B:LYS 213		107.47	0.00		0.00
199	B:ASN 214		33.99	0.00		0.00
200	B:LEU 215		0.00	0.00		0.00
201	B:ASN 216		65.12	0.00		0.00
202	B:GLU 217		53.62	0.00		0.00

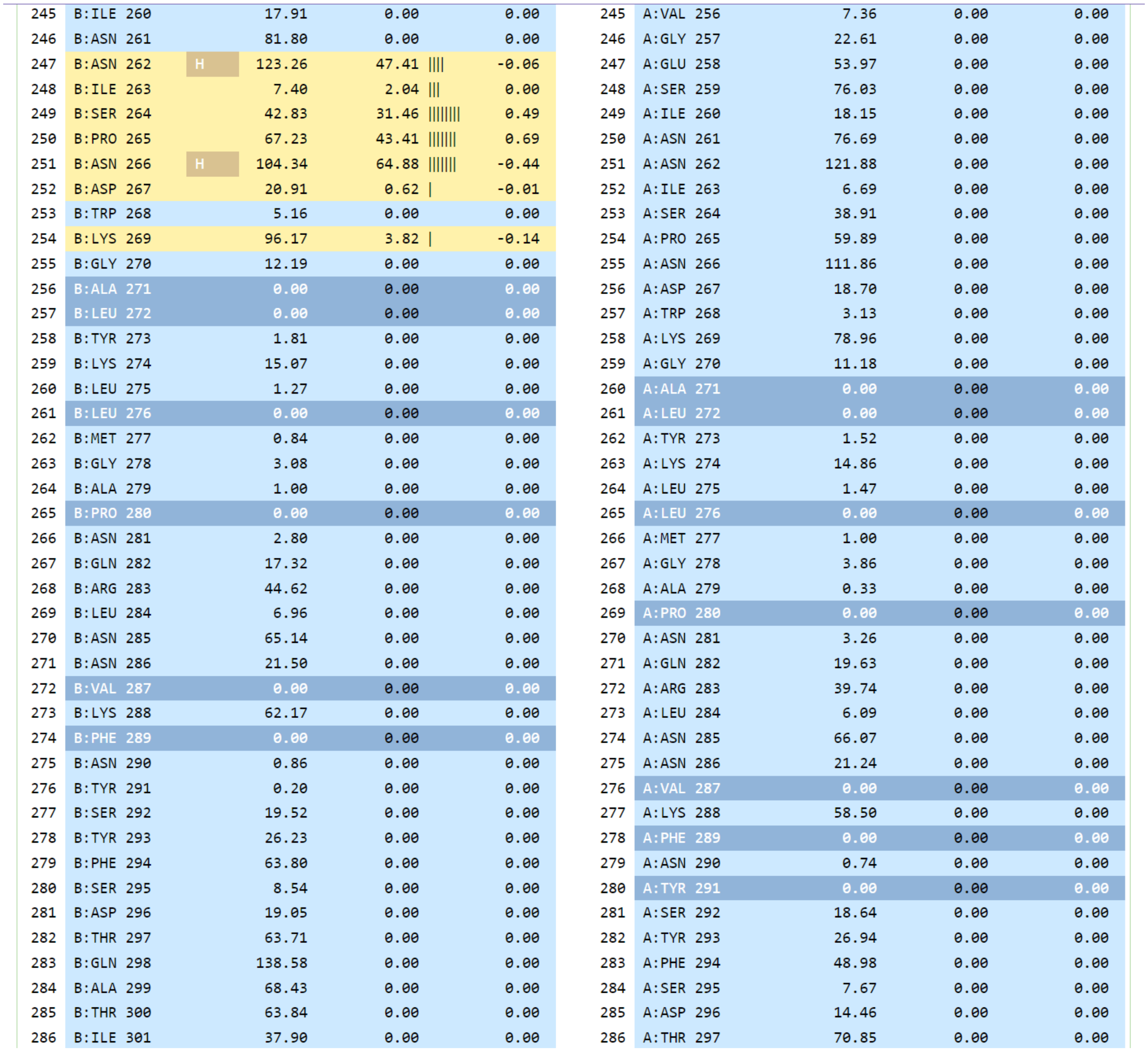
161	A:SER 172		80.61	0.00		0.00
162	A:ASP 173		66.85	0.00		0.00
163	A:TRP 174		5.06	0.00		0.00
164	A:GLY 175		22.96	0.00		0.00
165	A:PHE 176		4.09	0.00		0.00
166	A:ALA 177		48.14	0.00		0.00
167	A:PRO 178		90.56	0.00		0.00
168	A:ALA 179		52.31	0.00		0.00
169	A:ASP 180		59.84	0.00		0.00
170	A:LYS 181		58.88	0.00		0.00
171	A:ASP 182		0.74	0.00		0.00
172	A:ALA 183		38.11	0.00		0.00
173	A:LEU 184		1.34	0.00		0.00
174	A:ILE 185		0.48	0.00		0.00
175	A:CYS 186		13.30	0.00		0.00
176	A:GLN 187		40.67	0.00		0.00
177	A:PHE 188		0.78	0.00		0.00
178	A:ASN 189		24.68	0.00		0.00
179	A:ARG 190		91.96	0.00		0.00
180	A:PHE 191		8.60	0.00		0.00
181	A:VAL 192		14.21	0.00		0.00
182	A:ASN 193		105.04	0.00		0.00
183	A:GLU 194		64.17	0.00		0.00
184	A:TYR 195		0.74	0.00		0.00
185	A:ASN 196		42.24	0.00		0.00
186	A:THR 197		81.89	0.00		0.00
187	A:ARG 198		47.58	0.00		0.00
188	A:LEU 199		0.00	0.00		0.00
189	A:MET 200		43.80	0.00		0.00
190	A:VAL 201		71.58	0.00		0.00
191	A:LEU 202		8.34	0.00		0.00
192	A:TYR 203		12.80	0.00		0.00
193	A:SER 204		60.99	0.00		0.00
194	A:LYS 205		113.74	0.00		0.00
195	A:GLU 206		5.50	0.00		0.00
196	A:PHE 207		10.46	0.00		0.00
197	A:GLY 208		15.62	0.00		0.00
198	A:ARG 209		115.67	0.00		0.00
199	A:LEU 210		0.75	0.00		0.00
200	A:LEU 211		0.00	0.00		0.00
201	A:ALA 212		47.00	0.00		0.00
202	A:LYS 213		114.84	8.83	|	-0.33



```
203  B:ALA 218      0.00      0.00      0.00
204  B:LEU 219      0.82      0.00      0.00
205  B:ASN 220     33.77      0.00      0.00
206  B:PHE 221      0.00      0.00      0.00
207  B:ARG 222      0.00      0.00      0.00
208  B:ASN 223     11.93      0.00      0.00
209  B:MET 224      0.00      0.00      0.00
210  B:CYS 225      0.00      0.00      0.00
211  B:SER 226      0.67      0.00      0.00
212  B:LEU 227     28.77      0.00      0.00
213  B:TYR 228      2.87      0.00      0.00
214  B:VAL 229      0.00      0.00      0.00
215  B:PHE 230      0.00      0.00      0.00
216  B:PRO 231     13.29      0.00      0.00
217  B:PHE 232      1.10      0.00      0.00
218  B:SER 233      3.79      0.00      0.00
219  B:GLU 234      1.91      0.00      0.00
220  B:ALA 235      1.16      0.00      0.00
221  B:TRP 236      4.65      0.00      0.00
222  B:SER 237     37.89      0.00      0.00
223  B:LEU 238      0.62      0.00      0.00
224  B:LEU 239      5.85      0.00      0.00
225  B:ARG 240     25.48      0.00      0.00
226  B:TYR 241     10.32      0.00      0.00
227  B:GLU 242     12.05      0.00      0.00
228  B:GLY 243      2.07      0.00      0.00
229  B:THR 244      4.51      0.00      0.00
230  B:LYS 245     77.68      0.00      0.00
231  B:LEU 246     26.05      0.00      0.00
232  B:GLU 247     32.13      0.00      0.00
233  B:ASN 248     31.25      0.00      0.00
234  B:THR 249      7.92      0.00      0.00
235  B:LEU 250      0.86      0.00      0.00
236  B:SER 251      2.03      0.00      0.00
237  B:LEU 252      1.17      0.00      0.00
238  B:TRP 253      3.89      0.00      0.00
239  B:ASN 254      1.76      0.00      0.00
240  B:PHE 255     17.17      0.00      0.00
241  B:VAL 256      8.36      0.00      0.00
242  B:GLY 257     21.98      0.00      0.00
243  B:GLU 258     52.33      0.00      0.00
244  B:SER 259     76.43      0.00      0.00
```

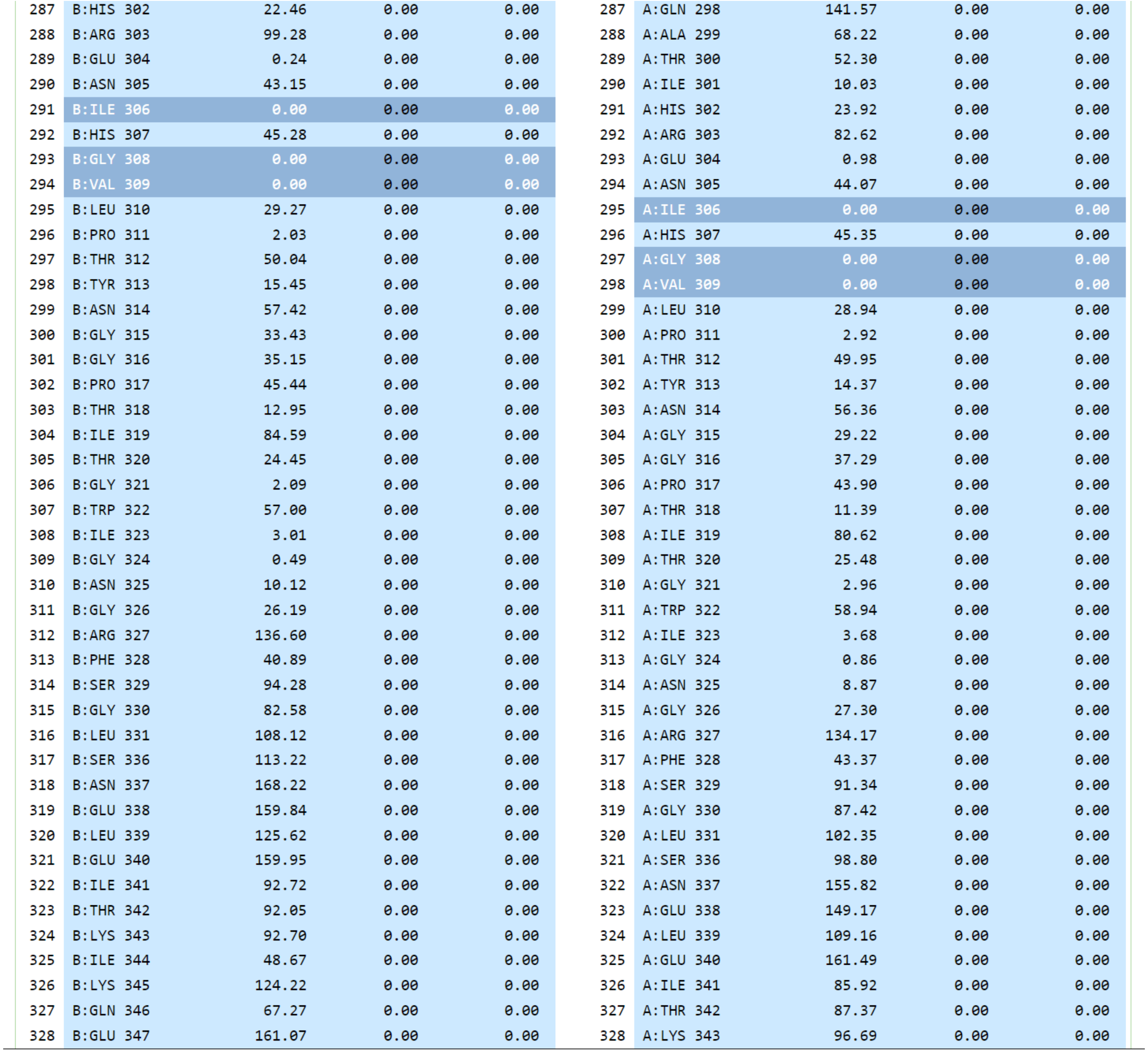


```
203  A:ASN 214     36.66     11.92 ||||      -0.04
204  A:LEU 215      0.00      0.00            0.00
205  A:ASN 216     67.17     20.23 ||||      -0.22
206  A:GLU 217     45.00     31.89 ||||||||   0.23
207  A:ALA 218      0.00      0.00            0.00
208  A:LEU 219      1.51      0.00            0.00
209  A:ASN 220     33.98     14.77 |||||      0.09
210  A:PHE 221      0.83      0.00            0.00
211  A:ARG 222      0.17      0.00            0.00
212  A:ASN 223     11.96      0.00            0.00
213  A:MET 224      0.00      0.00            0.00
214  A:CYS 225      0.00      0.00            0.00
215  A:SER 226      1.15      0.00            0.00
216  A:LEU 227     24.84      0.00            0.00
217  A:TYR 228      4.02      0.00            0.00
218  A:VAL 229      0.00      0.00            0.00
219  A:PHE 230      0.00      0.00            0.00
220  A:PRO 231     11.58      0.00            0.00
221  A:PHE 232      0.31      0.00            0.00
222  A:SER 233      4.10      0.00            0.00
223  A:GLU 234      4.54      0.00            0.00
224  A:ALA 235      1.34      0.00            0.00
225  A:TRP 236      5.53      0.00            0.00
226  A:SER 237     40.52      0.00            0.00
227  A:LEU 238      0.00      0.00            0.00
228  A:LEU 239      5.54      0.00            0.00
229  A:ARG 240     30.45      0.00            0.00
230  A:TYR 241      7.66      0.00            0.00
231  A:GLU 242     16.26      0.00            0.00
232  A:GLY 243      1.38      0.00            0.00
233  A:THR 244      3.65      0.00            0.00
234  A:LYS 245     84.36      0.00            0.00
235  A:LEU 246     24.43      0.00            0.00
236  A:GLU 247     33.46      0.00            0.00
237  A:ASN 248     32.16      0.00            0.00
238  A:THR 249      7.72      0.00            0.00
239  A:LEU 250      1.36      0.00            0.00
240  A:SER 251      1.85      0.00            0.00
241  A:LEU 252      1.00      0.00            0.00
242  A:TRP 253      3.06      0.00            0.00
243  A:ASN 254      2.43      0.00            0.00
244  A:PHE 255     16.44      0.00            0.00
```

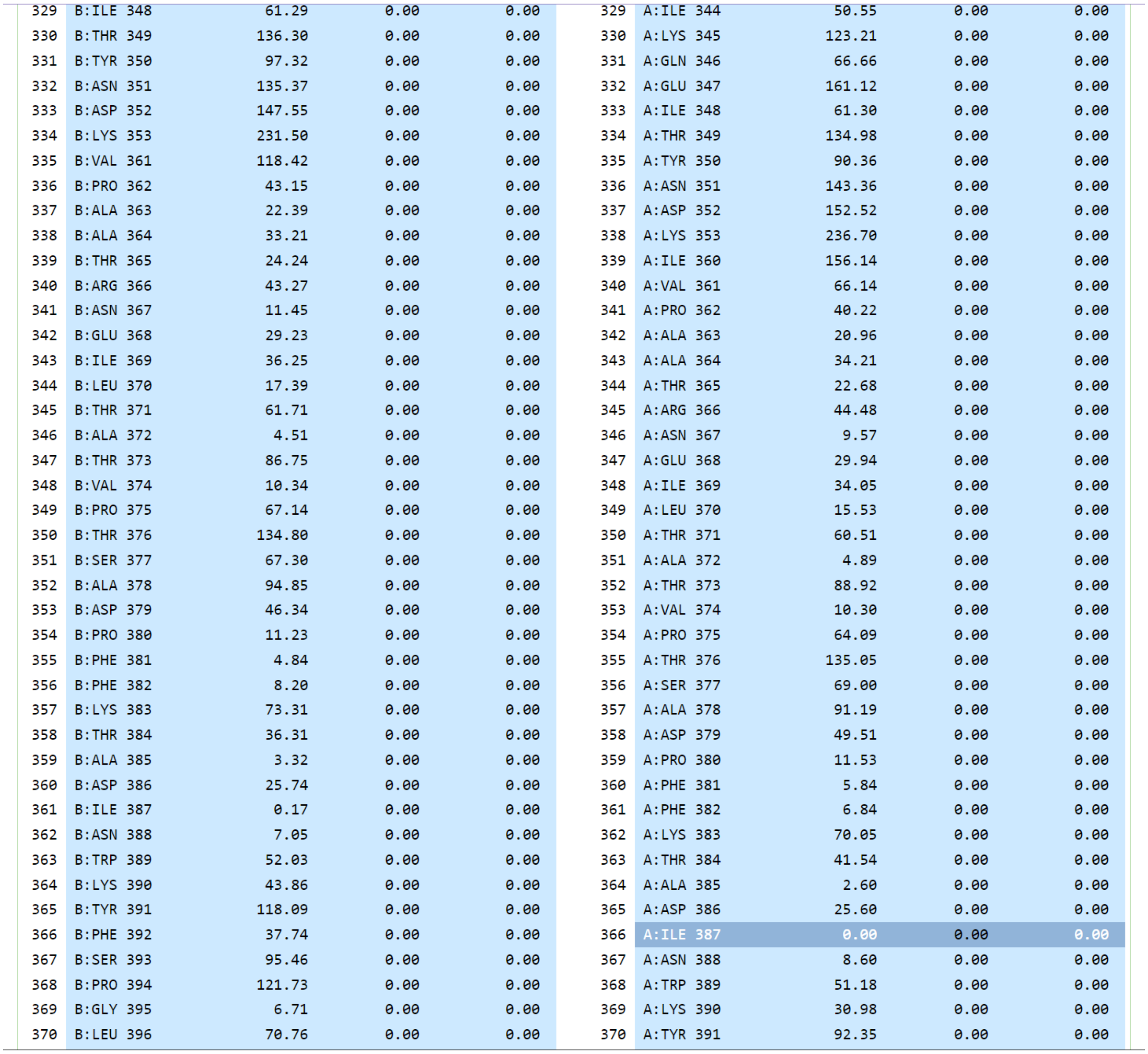


245	B:ILE 260		17.91	0.00	0.00
246	B:ASN 261		81.80	0.00	0.00
247	B:ASN 262	H	123.26	47.41 ||||	-0.06
248	B:ILE 263		7.40	2.04 |||	0.00
249	B:SER 264		42.83	31.46 ||||||||	0.49
250	B:PRO 265		67.23	43.41 |||||||	0.69
251	B:ASN 266	H	104.34	64.88 |||||||	-0.44
252	B:ASP 267		20.91	0.62 |	-0.01
253	B:TRP 268		5.16	0.00	0.00
254	B:LYS 269		96.17	3.82 |	-0.14
255	B:GLY 270		12.19	0.00	0.00
256	B:ALA 271		0.00	0.00	0.00
257	B:LEU 272		0.00	0.00	0.00
258	B:TYR 273		1.81	0.00	0.00
259	B:LYS 274		15.07	0.00	0.00
260	B:LEU 275		1.27	0.00	0.00
261	B:LEU 276		0.00	0.00	0.00
262	B:MET 277		0.84	0.00	0.00
263	B:GLY 278		3.08	0.00	0.00
264	B:ALA 279		1.00	0.00	0.00
265	B:PRO 280		0.00	0.00	0.00
266	B:ASN 281		2.80	0.00	0.00
267	B:GLN 282		17.32	0.00	0.00
268	B:ARG 283		44.62	0.00	0.00
269	B:LEU 284		6.96	0.00	0.00
270	B:ASN 285		65.14	0.00	0.00
271	B:ASN 286		21.50	0.00	0.00
272	B:VAL 287		0.00	0.00	0.00
273	B:LYS 288		62.17	0.00	0.00
274	B:PHE 289		0.00	0.00	0.00
275	B:ASN 290		0.86	0.00	0.00
276	B:TYR 291		0.20	0.00	0.00
277	B:SER 292		19.52	0.00	0.00
278	B:TYR 293		26.23	0.00	0.00
279	B:PHE 294		63.80	0.00	0.00
280	B:SER 295		8.54	0.00	0.00
281	B:ASP 296		19.05	0.00	0.00
282	B:THR 297		63.71	0.00	0.00
283	B:GLN 298		138.58	0.00	0.00
284	B:ALA 299		68.43	0.00	0.00
285	B:THR 300		63.84	0.00	0.00
286	B:ILE 301		37.90	0.00	0.00

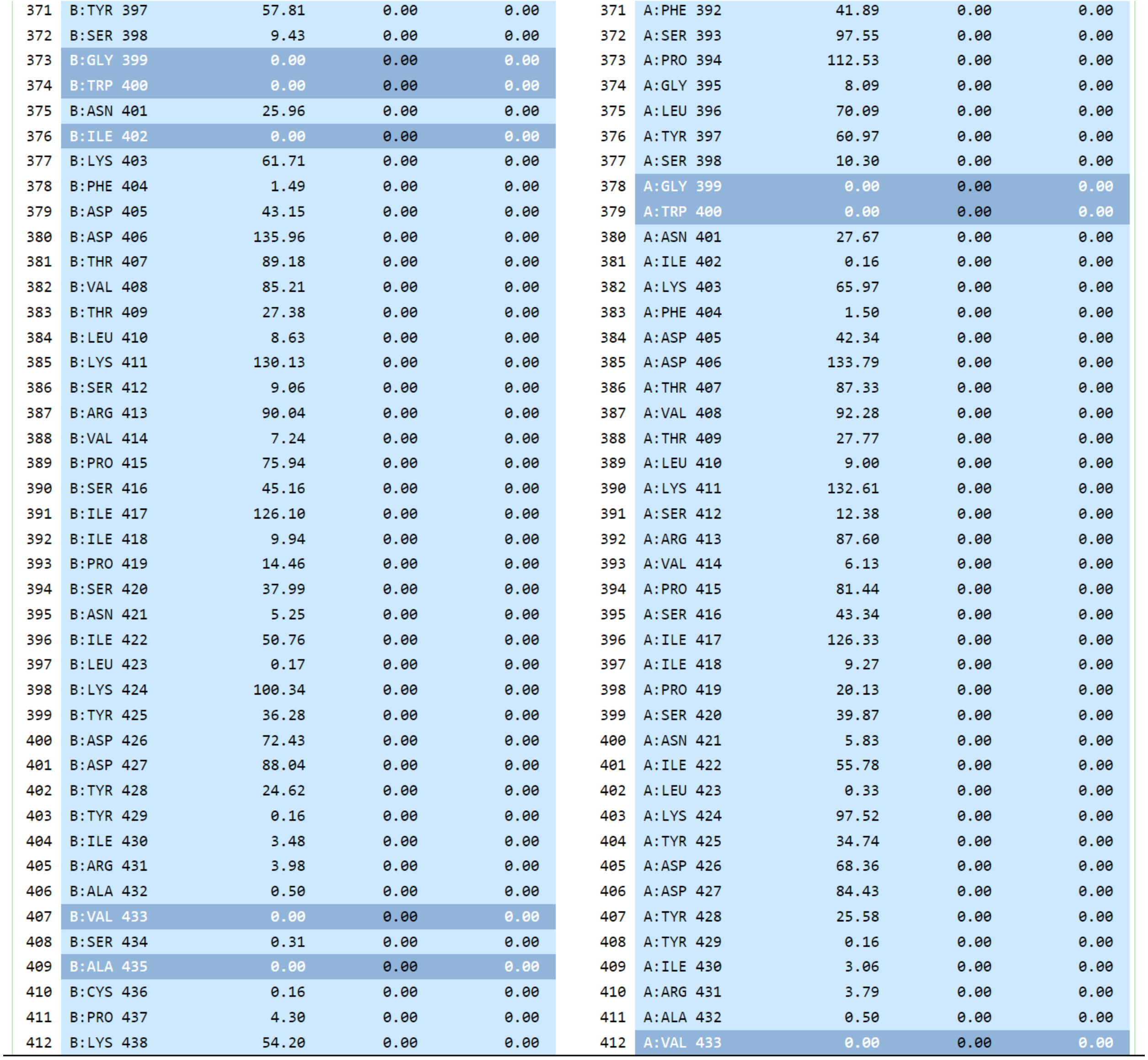
245	A:VAL 256	7.36	0.00	0.00
246	A:GLY 257	22.61	0.00	0.00
247	A:GLU 258	53.97	0.00	0.00
248	A:SER 259	76.03	0.00	0.00
249	A:ILE 260	18.15	0.00	0.00
250	A:ASN 261	76.69	0.00	0.00
251	A:ASN 262	121.88	0.00	0.00
252	A:ILE 263	6.69	0.00	0.00
253	A:SER 264	38.91	0.00	0.00
254	A:PRO 265	59.89	0.00	0.00
255	A:ASN 266	111.86	0.00	0.00
256	A:ASP 267	18.70	0.00	0.00
257	A:TRP 268	3.13	0.00	0.00
258	A:LYS 269	78.96	0.00	0.00
259	A:GLY 270	11.18	0.00	0.00
260	A:ALA 271	0.00	0.00	0.00
261	A:LEU 272	0.00	0.00	0.00
262	A:TYR 273	1.52	0.00	0.00
263	A:LYS 274	14.86	0.00	0.00
264	A:LEU 275	1.47	0.00	0.00
265	A:LEU 276	0.00	0.00	0.00
266	A:MET 277	1.00	0.00	0.00
267	A:GLY 278	3.86	0.00	0.00
268	A:ALA 279	0.33	0.00	0.00
269	A:PRO 280	0.00	0.00	0.00
270	A:ASN 281	3.26	0.00	0.00
271	A:GLN 282	19.63	0.00	0.00
272	A:ARG 283	39.74	0.00	0.00
273	A:LEU 284	6.09	0.00	0.00
274	A:ASN 285	66.07	0.00	0.00
275	A:ASN 286	21.24	0.00	0.00
276	A:VAL 287	0.00	0.00	0.00
277	A:LYS 288	58.50	0.00	0.00
278	A:PHE 289	0.00	0.00	0.00
279	A:ASN 290	0.74	0.00	0.00
280	A:TYR 291	0.00	0.00	0.00
281	A:SER 292	18.64	0.00	0.00
282	A:TYR 293	26.94	0.00	0.00
283	A:PHE 294	48.98	0.00	0.00
284	A:SER 295	7.67	0.00	0.00
285	A:ASP 296	14.46	0.00	0.00
286	A:THR 297	70.85	0.00	0.00



287	B:HIS 302	22.46	0.00	0.00	287	A:GLN 298	141.57	0.00	0.00
288	B:ARG 303	99.28	0.00	0.00	288	A:ALA 299	68.22	0.00	0.00
289	B:GLU 304	0.24	0.00	0.00	289	A:THR 300	52.30	0.00	0.00
290	B:ASN 305	43.15	0.00	0.00	290	A:ILE 301	10.03	0.00	0.00
291	B:ILE 306	0.00	0.00	0.00	291	A:HIS 302	23.92	0.00	0.00
292	B:HIS 307	45.28	0.00	0.00	292	A:ARG 303	82.62	0.00	0.00
293	B:GLY 308	0.00	0.00	0.00	293	A:GLU 304	0.98	0.00	0.00
294	B:VAL 309	0.00	0.00	0.00	294	A:ASN 305	44.07	0.00	0.00
295	B:LEU 310	29.27	0.00	0.00	295	A:ILE 306	0.00	0.00	0.00
296	B:PRO 311	2.03	0.00	0.00	296	A:HIS 307	45.35	0.00	0.00
297	B:THR 312	50.04	0.00	0.00	297	A:GLY 308	0.00	0.00	0.00
298	B:TYR 313	15.45	0.00	0.00	298	A:VAL 309	0.00	0.00	0.00
299	B:ASN 314	57.42	0.00	0.00	299	A:LEU 310	28.94	0.00	0.00
300	B:GLY 315	33.43	0.00	0.00	300	A:PRO 311	2.92	0.00	0.00
301	B:GLY 316	35.15	0.00	0.00	301	A:THR 312	49.95	0.00	0.00
302	B:PRO 317	45.44	0.00	0.00	302	A:TYR 313	14.37	0.00	0.00
303	B:THR 318	12.95	0.00	0.00	303	A:ASN 314	56.36	0.00	0.00
304	B:ILE 319	84.59	0.00	0.00	304	A:GLY 315	29.22	0.00	0.00
305	B:THR 320	24.45	0.00	0.00	305	A:GLY 316	37.29	0.00	0.00
306	B:GLY 321	2.09	0.00	0.00	306	A:PRO 317	43.90	0.00	0.00
307	B:TRP 322	57.00	0.00	0.00	307	A:THR 318	11.39	0.00	0.00
308	B:ILE 323	3.01	0.00	0.00	308	A:ILE 319	80.62	0.00	0.00
309	B:GLY 324	0.49	0.00	0.00	309	A:THR 320	25.48	0.00	0.00
310	B:ASN 325	10.12	0.00	0.00	310	A:GLY 321	2.96	0.00	0.00
311	B:GLY 326	26.19	0.00	0.00	311	A:TRP 322	58.94	0.00	0.00
312	B:ARG 327	136.60	0.00	0.00	312	A:ILE 323	3.68	0.00	0.00
313	B:PHE 328	40.89	0.00	0.00	313	A:GLY 324	0.86	0.00	0.00
314	B:SER 329	94.28	0.00	0.00	314	A:ASN 325	8.87	0.00	0.00
315	B:GLY 330	82.58	0.00	0.00	315	A:GLY 326	27.30	0.00	0.00
316	B:LEU 331	108.12	0.00	0.00	316	A:ARG 327	134.17	0.00	0.00
317	B:SER 336	113.22	0.00	0.00	317	A:PHE 328	43.37	0.00	0.00
318	B:ASN 337	168.22	0.00	0.00	318	A:SER 329	91.34	0.00	0.00
319	B:GLU 338	159.84	0.00	0.00	319	A:GLY 330	87.42	0.00	0.00
320	B:LEU 339	125.62	0.00	0.00	320	A:LEU 331	102.35	0.00	0.00
321	B:GLU 340	159.95	0.00	0.00	321	A:SER 336	98.80	0.00	0.00
322	B:ILE 341	92.72	0.00	0.00	322	A:ASN 337	155.82	0.00	0.00
323	B:THR 342	92.05	0.00	0.00	323	A:GLU 338	149.17	0.00	0.00
324	B:LYS 343	92.70	0.00	0.00	324	A:LEU 339	109.16	0.00	0.00
325	B:ILE 344	48.67	0.00	0.00	325	A:GLU 340	161.49	0.00	0.00
326	B:LYS 345	124.22	0.00	0.00	326	A:ILE 341	85.92	0.00	0.00
327	B:GLN 346	67.27	0.00	0.00	327	A:THR 342	87.37	0.00	0.00
328	B:GLU 347	161.07	0.00	0.00	328	A:LYS 343	96.69	0.00	0.00

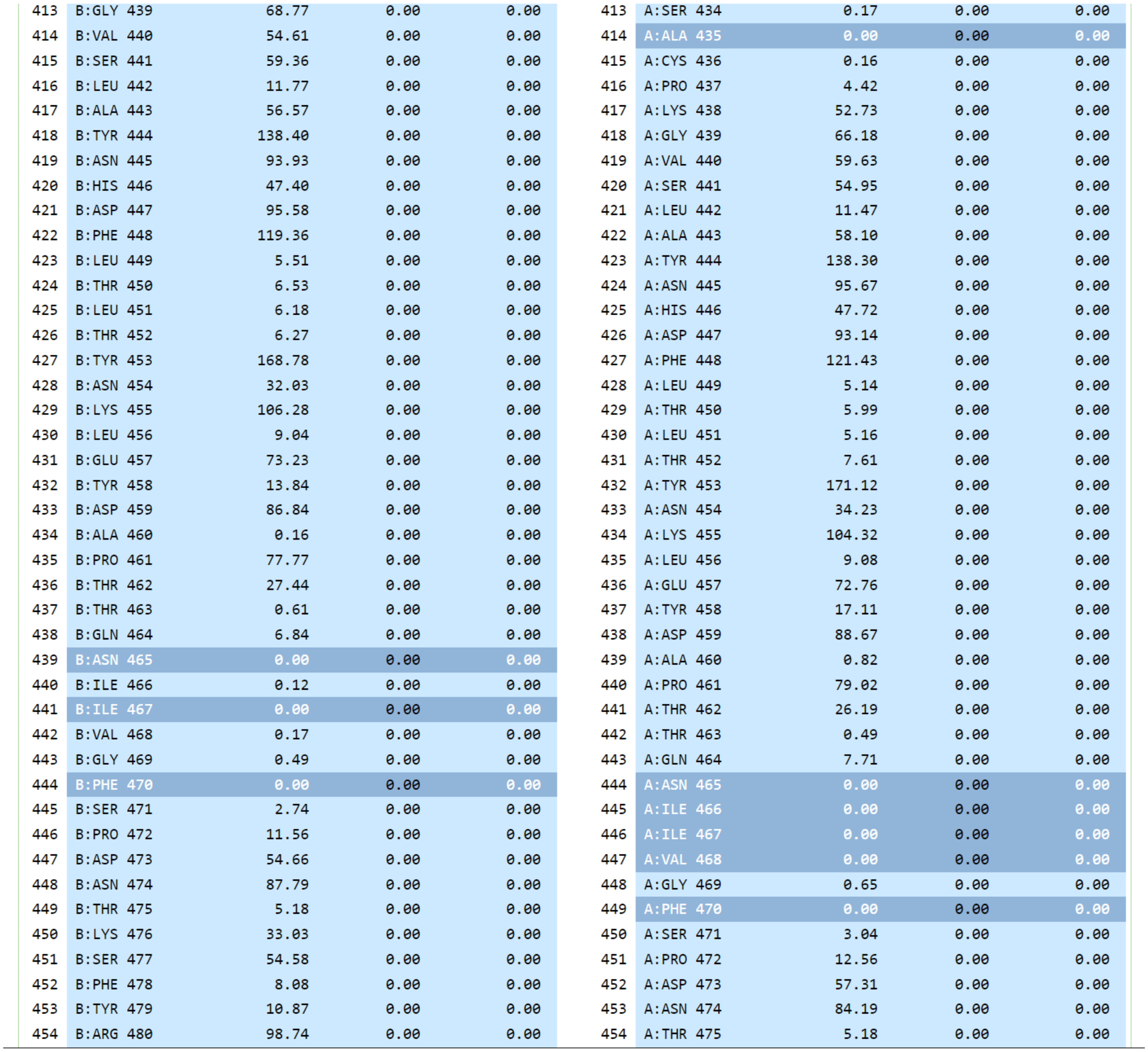


329	B:ILE 348	61.29	0.00	0.00	329	A:ILE 344	50.55	0.00	0.00
330	B:THR 349	136.30	0.00	0.00	330	A:LYS 345	123.21	0.00	0.00
331	B:TYR 350	97.32	0.00	0.00	331	A:GLN 346	66.66	0.00	0.00
332	B:ASN 351	135.37	0.00	0.00	332	A:GLU 347	161.12	0.00	0.00
333	B:ASP 352	147.55	0.00	0.00	333	A:ILE 348	61.30	0.00	0.00
334	B:LYS 353	231.50	0.00	0.00	334	A:THR 349	134.98	0.00	0.00
335	B:VAL 361	118.42	0.00	0.00	335	A:TYR 350	90.36	0.00	0.00
336	B:PRO 362	43.15	0.00	0.00	336	A:ASN 351	143.36	0.00	0.00
337	B:ALA 363	22.39	0.00	0.00	337	A:ASP 352	152.52	0.00	0.00
338	B:ALA 364	33.21	0.00	0.00	338	A:LYS 353	236.70	0.00	0.00
339	B:THR 365	24.24	0.00	0.00	339	A:ILE 360	156.14	0.00	0.00
340	B:ARG 366	43.27	0.00	0.00	340	A:VAL 361	66.14	0.00	0.00
341	B:ASN 367	11.45	0.00	0.00	341	A:PRO 362	40.22	0.00	0.00
342	B:GLU 368	29.23	0.00	0.00	342	A:ALA 363	20.96	0.00	0.00
343	B:ILE 369	36.25	0.00	0.00	343	A:ALA 364	34.21	0.00	0.00
344	B:LEU 370	17.39	0.00	0.00	344	A:THR 365	22.68	0.00	0.00
345	B:THR 371	61.71	0.00	0.00	345	A:ARG 366	44.48	0.00	0.00
346	B:ALA 372	4.51	0.00	0.00	346	A:ASN 367	9.57	0.00	0.00
347	B:THR 373	86.75	0.00	0.00	347	A:GLU 368	29.94	0.00	0.00
348	B:VAL 374	10.34	0.00	0.00	348	A:ILE 369	34.05	0.00	0.00
349	B:PRO 375	67.14	0.00	0.00	349	A:LEU 370	15.53	0.00	0.00
350	B:THR 376	134.80	0.00	0.00	350	A:THR 371	60.51	0.00	0.00
351	B:SER 377	67.30	0.00	0.00	351	A:ALA 372	4.89	0.00	0.00
352	B:ALA 378	94.85	0.00	0.00	352	A:THR 373	88.92	0.00	0.00
353	B:ASP 379	46.34	0.00	0.00	353	A:VAL 374	10.30	0.00	0.00
354	B:PRO 380	11.23	0.00	0.00	354	A:PRO 375	64.09	0.00	0.00
355	B:PHE 381	4.84	0.00	0.00	355	A:THR 376	135.05	0.00	0.00
356	B:PHE 382	8.20	0.00	0.00	356	A:SER 377	69.00	0.00	0.00
357	B:LYS 383	73.31	0.00	0.00	357	A:ALA 378	91.19	0.00	0.00
358	B:THR 384	36.31	0.00	0.00	358	A:ASP 379	49.51	0.00	0.00
359	B:ALA 385	3.32	0.00	0.00	359	A:PRO 380	11.53	0.00	0.00
360	B:ASP 386	25.74	0.00	0.00	360	A:PHE 381	5.84	0.00	0.00
361	B:ILE 387	0.17	0.00	0.00	361	A:PHE 382	6.84	0.00	0.00
362	B:ASN 388	7.05	0.00	0.00	362	A:LYS 383	70.05	0.00	0.00
363	B:TRP 389	52.03	0.00	0.00	363	A:THR 384	41.54	0.00	0.00
364	B:LYS 390	43.86	0.00	0.00	364	A:ALA 385	2.60	0.00	0.00
365	B:TYR 391	118.09	0.00	0.00	365	A:ASP 386	25.60	0.00	0.00
366	B:PHE 392	37.74	0.00	0.00	366	A:ILE 387	0.00	0.00	0.00
367	B:SER 393	95.46	0.00	0.00	367	A:ASN 388	8.60	0.00	0.00
368	B:PRO 394	121.73	0.00	0.00	368	A:TRP 389	51.18	0.00	0.00
369	B:GLY 395	6.71	0.00	0.00	369	A:LYS 390	30.98	0.00	0.00
370	B:LEU 396	70.76	0.00	0.00	370	A:TYR 391	92.35	0.00	0.00

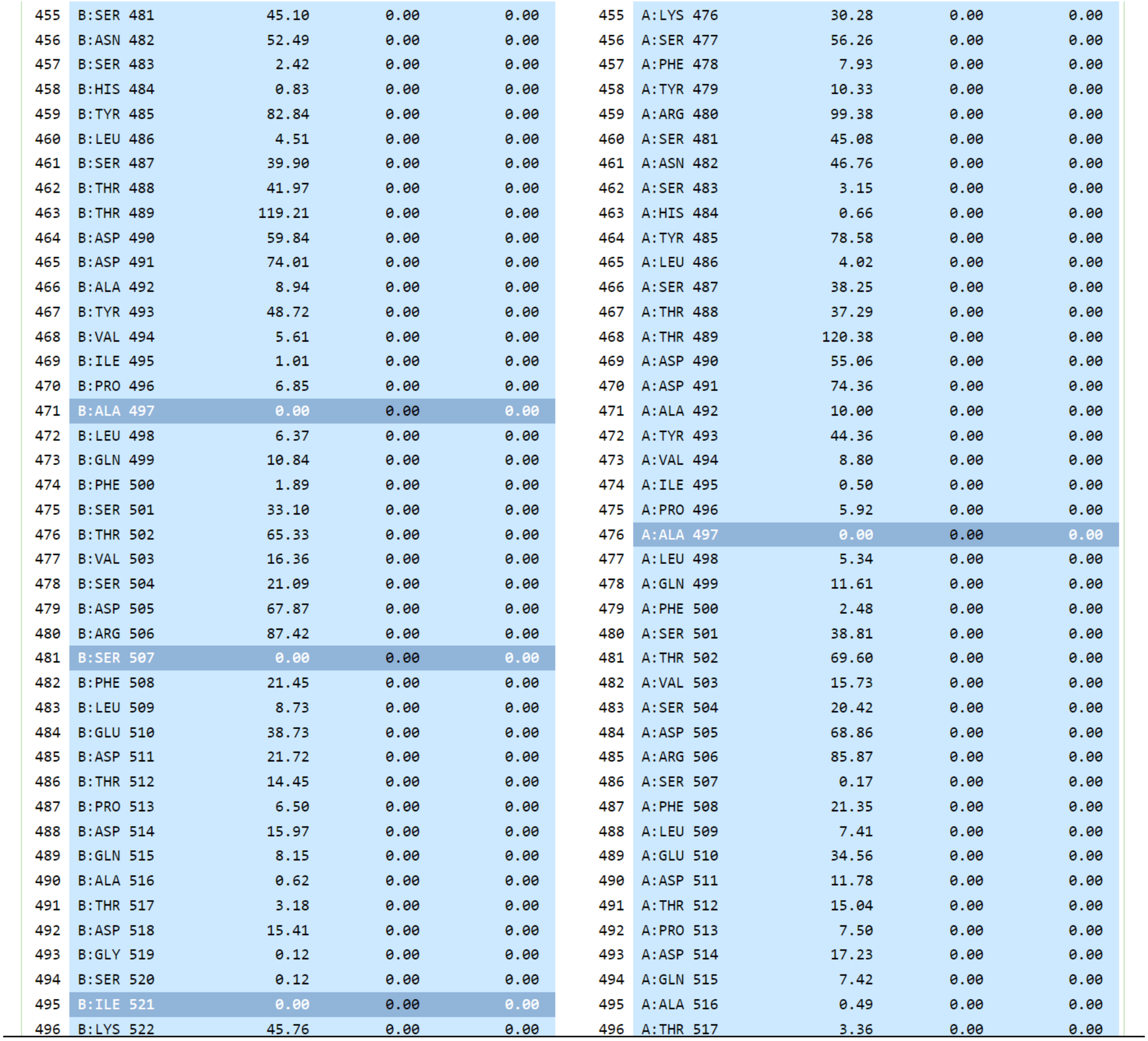


371	B:TYR 397	57.81	0.00	0.00
372	B:SER 398	9.43	0.00	0.00
373	B:GLY 399	0.00	0.00	0.00
374	B:TRP 400	0.00	0.00	0.00
375	B:ASN 401	25.96	0.00	0.00
376	B:ILE 402	0.00	0.00	0.00
377	B:LYS 403	61.71	0.00	0.00
378	B:PHE 404	1.49	0.00	0.00
379	B:ASP 405	43.15	0.00	0.00
380	B:ASP 406	135.96	0.00	0.00
381	B:THR 407	89.18	0.00	0.00
382	B:VAL 408	85.21	0.00	0.00
383	B:THR 409	27.38	0.00	0.00
384	B:LEU 410	8.63	0.00	0.00
385	B:LYS 411	130.13	0.00	0.00
386	B:SER 412	9.06	0.00	0.00
387	B:ARG 413	90.04	0.00	0.00
388	B:VAL 414	7.24	0.00	0.00
389	B:PRO 415	75.94	0.00	0.00
390	B:SER 416	45.16	0.00	0.00
391	B:ILE 417	126.10	0.00	0.00
392	B:ILE 418	9.94	0.00	0.00
393	B:PRO 419	14.46	0.00	0.00
394	B:SER 420	37.99	0.00	0.00
395	B:ASN 421	5.25	0.00	0.00
396	B:ILE 422	50.76	0.00	0.00
397	B:LEU 423	0.17	0.00	0.00
398	B:LYS 424	100.34	0.00	0.00
399	B:TYR 425	36.28	0.00	0.00
400	B:ASP 426	72.43	0.00	0.00
401	B:ASP 427	88.04	0.00	0.00
402	B:TYR 428	24.62	0.00	0.00
403	B:TYR 429	0.16	0.00	0.00
404	B:ILE 430	3.48	0.00	0.00
405	B:ARG 431	3.98	0.00	0.00
406	B:ALA 432	0.50	0.00	0.00
407	B:VAL 433	0.00	0.00	0.00
408	B:SER 434	0.31	0.00	0.00
409	B:ALA 435	0.00	0.00	0.00
410	B:CYS 436	0.16	0.00	0.00
411	B:PRO 437	4.30	0.00	0.00
412	B:LYS 438	54.20	0.00	0.00

371	A:PHE 392	41.89	0.00	0.00
372	A:SER 393	97.55	0.00	0.00
373	A:PRO 394	112.53	0.00	0.00
374	A:GLY 395	8.09	0.00	0.00
375	A:LEU 396	70.09	0.00	0.00
376	A:TYR 397	60.97	0.00	0.00
377	A:SER 398	10.30	0.00	0.00
378	A:GLY 399	0.00	0.00	0.00
379	A:TRP 400	0.00	0.00	0.00
380	A:ASN 401	27.67	0.00	0.00
381	A:ILE 402	0.16	0.00	0.00
382	A:LYS 403	65.97	0.00	0.00
383	A:PHE 404	1.50	0.00	0.00
384	A:ASP 405	42.34	0.00	0.00
385	A:ASP 406	133.79	0.00	0.00
386	A:THR 407	87.33	0.00	0.00
387	A:VAL 408	92.28	0.00	0.00
388	A:THR 409	27.77	0.00	0.00
389	A:LEU 410	9.00	0.00	0.00
390	A:LYS 411	132.61	0.00	0.00
391	A:SER 412	12.38	0.00	0.00
392	A:ARG 413	87.60	0.00	0.00
393	A:VAL 414	6.13	0.00	0.00
394	A:PRO 415	81.44	0.00	0.00
395	A:SER 416	43.34	0.00	0.00
396	A:ILE 417	126.33	0.00	0.00
397	A:ILE 418	9.27	0.00	0.00
398	A:PRO 419	20.13	0.00	0.00
399	A:SER 420	39.87	0.00	0.00
400	A:ASN 421	5.83	0.00	0.00
401	A:ILE 422	55.78	0.00	0.00
402	A:LEU 423	0.33	0.00	0.00
403	A:LYS 424	97.52	0.00	0.00
404	A:TYR 425	34.74	0.00	0.00
405	A:ASP 426	68.36	0.00	0.00
406	A:ASP 427	84.43	0.00	0.00
407	A:TYR 428	25.58	0.00	0.00
408	A:TYR 429	0.16	0.00	0.00
409	A:ILE 430	3.06	0.00	0.00
410	A:ARG 431	3.79	0.00	0.00
411	A:ALA 432	0.50	0.00	0.00
412	A:VAL 433	0.00	0.00	0.00

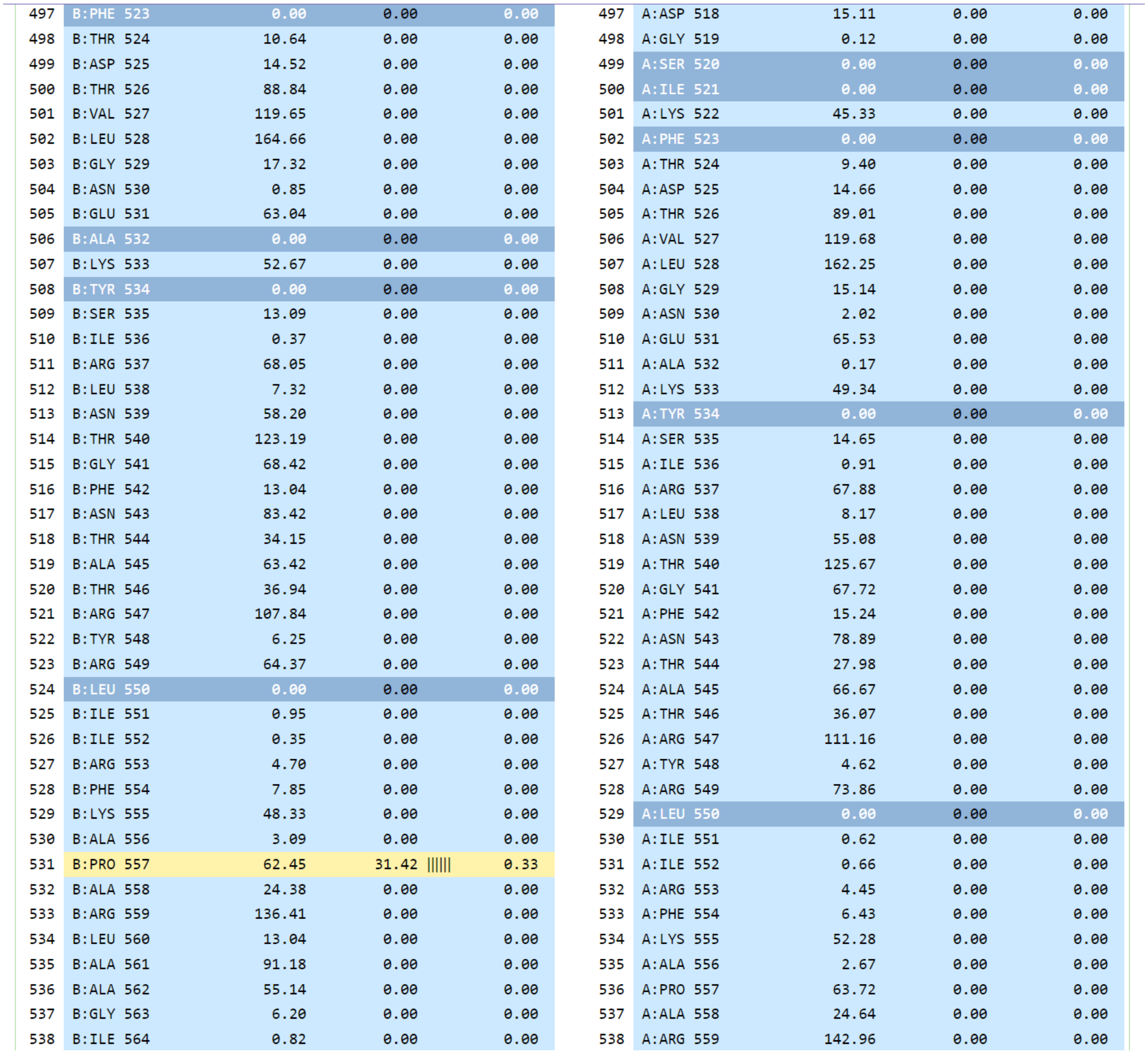


413	B:GLY 439	68.77	0.00	0.00	413	A:SER 434	0.17	0.00	0.00
414	B:VAL 440	54.61	0.00	0.00	414	A:ALA 435	0.00	0.00	0.00
415	B:SER 441	59.36	0.00	0.00	415	A:CYS 436	0.16	0.00	0.00
416	B:LEU 442	11.77	0.00	0.00	416	A:PRO 437	4.42	0.00	0.00
417	B:ALA 443	56.57	0.00	0.00	417	A:LYS 438	52.73	0.00	0.00
418	B:TYR 444	138.40	0.00	0.00	418	A:GLY 439	66.18	0.00	0.00
419	B:ASN 445	93.93	0.00	0.00	419	A:VAL 440	59.63	0.00	0.00
420	B:HIS 446	47.40	0.00	0.00	420	A:SER 441	54.95	0.00	0.00
421	B:ASP 447	95.58	0.00	0.00	421	A:LEU 442	11.47	0.00	0.00
422	B:PHE 448	119.36	0.00	0.00	422	A:ALA 443	58.10	0.00	0.00
423	B:LEU 449	5.51	0.00	0.00	423	A:TYR 444	138.30	0.00	0.00
424	B:THR 450	6.53	0.00	0.00	424	A:ASN 445	95.67	0.00	0.00
425	B:LEU 451	6.18	0.00	0.00	425	A:HIS 446	47.72	0.00	0.00
426	B:THR 452	6.27	0.00	0.00	426	A:ASP 447	93.14	0.00	0.00
427	B:TYR 453	168.78	0.00	0.00	427	A:PHE 448	121.43	0.00	0.00
428	B:ASN 454	32.03	0.00	0.00	428	A:LEU 449	5.14	0.00	0.00
429	B:LYS 455	106.28	0.00	0.00	429	A:THR 450	5.99	0.00	0.00
430	B:LEU 456	9.04	0.00	0.00	430	A:LEU 451	5.16	0.00	0.00
431	B:GLU 457	73.23	0.00	0.00	431	A:THR 452	7.61	0.00	0.00
432	B:TYR 458	13.84	0.00	0.00	432	A:TYR 453	171.12	0.00	0.00
433	B:ASP 459	86.84	0.00	0.00	433	A:ASN 454	34.23	0.00	0.00
434	B:ALA 460	0.16	0.00	0.00	434	A:LYS 455	104.32	0.00	0.00
435	B:PRO 461	77.77	0.00	0.00	435	A:LEU 456	9.08	0.00	0.00
436	B:THR 462	27.44	0.00	0.00	436	A:GLU 457	72.76	0.00	0.00
437	B:THR 463	0.61	0.00	0.00	437	A:TYR 458	17.11	0.00	0.00
438	B:GLN 464	6.84	0.00	0.00	438	A:ASP 459	88.67	0.00	0.00
439	B:ASN 465	0.00	0.00	0.00	439	A:ALA 460	0.82	0.00	0.00
440	B:ILE 466	0.12	0.00	0.00	440	A:PRO 461	79.02	0.00	0.00
441	B:ILE 467	0.00	0.00	0.00	441	A:THR 462	26.19	0.00	0.00
442	B:VAL 468	0.17	0.00	0.00	442	A:THR 463	0.49	0.00	0.00
443	B:GLY 469	0.49	0.00	0.00	443	A:GLN 464	7.71	0.00	0.00
444	B:PHE 470	0.00	0.00	0.00	444	A:ASN 465	0.00	0.00	0.00
445	B:SER 471	2.74	0.00	0.00	445	A:ILE 466	0.00	0.00	0.00
446	B:PRO 472	11.56	0.00	0.00	446	A:ILE 467	0.00	0.00	0.00
447	B:ASP 473	54.66	0.00	0.00	447	A:VAL 468	0.00	0.00	0.00
448	B:ASN 474	87.79	0.00	0.00	448	A:GLY 469	0.65	0.00	0.00
449	B:THR 475	5.18	0.00	0.00	449	A:PHE 470	0.00	0.00	0.00
450	B:LYS 476	33.03	0.00	0.00	450	A:SER 471	3.04	0.00	0.00
451	B:SER 477	54.58	0.00	0.00	451	A:PRO 472	12.56	0.00	0.00
452	B:PHE 478	8.08	0.00	0.00	452	A:ASP 473	57.31	0.00	0.00
453	B:TYR 479	10.87	0.00	0.00	453	A:ASN 474	84.19	0.00	0.00
454	B:ARG 480	98.74	0.00	0.00	454	A:THR 475	5.18	0.00	0.00



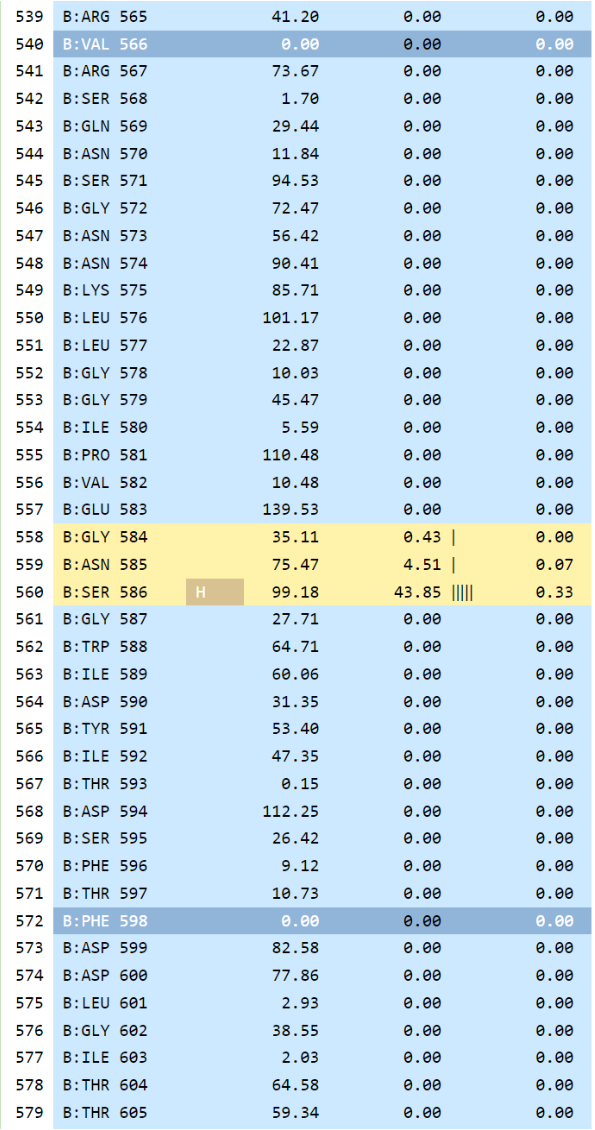
455	B:SER 481	45.10	0.00	0.00
456	B:ASN 482	52.49	0.00	0.00
457	B:SER 483	2.42	0.00	0.00
458	B:HIS 484	0.83	0.00	0.00
459	B:TYR 485	82.84	0.00	0.00
460	B:LEU 486	4.51	0.00	0.00
461	B:SER 487	39.90	0.00	0.00
462	B:THR 488	41.97	0.00	0.00
463	B:THR 489	119.21	0.00	0.00
464	B:ASP 490	59.84	0.00	0.00
465	B:ASP 491	74.01	0.00	0.00
466	B:ALA 492	8.94	0.00	0.00
467	B:TYR 493	48.72	0.00	0.00
468	B:VAL 494	5.61	0.00	0.00
469	B:ILE 495	1.01	0.00	0.00
470	B:PRO 496	6.85	0.00	0.00
471	B:ALA 497	0.00	0.00	0.00
472	B:LEU 498	6.37	0.00	0.00
473	B:GLN 499	10.84	0.00	0.00
474	B:PHE 500	1.89	0.00	0.00
475	B:SER 501	33.10	0.00	0.00
476	B:THR 502	65.33	0.00	0.00
477	B:VAL 503	16.36	0.00	0.00
478	B:SER 504	21.09	0.00	0.00
479	B:ASP 505	67.87	0.00	0.00
480	B:ARG 506	87.42	0.00	0.00
481	B:SER 507	0.00	0.00	0.00
482	B:PHE 508	21.45	0.00	0.00
483	B:LEU 509	8.73	0.00	0.00
484	B:GLU 510	38.73	0.00	0.00
485	B:ASP 511	21.72	0.00	0.00
486	B:THR 512	14.45	0.00	0.00
487	B:PRO 513	6.50	0.00	0.00
488	B:ASP 514	15.97	0.00	0.00
489	B:GLN 515	8.15	0.00	0.00
490	B:ALA 516	0.62	0.00	0.00
491	B:THR 517	3.18	0.00	0.00
492	B:ASP 518	15.41	0.00	0.00
493	B:GLY 519	0.12	0.00	0.00
494	B:SER 520	0.12	0.00	0.00
495	B:ILE 521	0.00	0.00	0.00
496	B:LYS 522	45.76	0.00	0.00

455	A:LYS 476	30.28	0.00	0.00
456	A:SER 477	56.26	0.00	0.00
457	A:PHE 478	7.93	0.00	0.00
458	A:TYR 479	10.33	0.00	0.00
459	A:ARG 480	99.38	0.00	0.00
460	A:SER 481	45.08	0.00	0.00
461	A:ASN 482	46.76	0.00	0.00
462	A:SER 483	3.15	0.00	0.00
463	A:HIS 484	0.66	0.00	0.00
464	A:TYR 485	78.58	0.00	0.00
465	A:LEU 486	4.02	0.00	0.00
466	A:SER 487	38.25	0.00	0.00
467	A:THR 488	37.29	0.00	0.00
468	A:THR 489	120.38	0.00	0.00
469	A:ASP 490	55.06	0.00	0.00
470	A:ASP 491	74.36	0.00	0.00
471	A:ALA 492	10.00	0.00	0.00
472	A:TYR 493	44.36	0.00	0.00
473	A:VAL 494	8.80	0.00	0.00
474	A:ILE 495	0.50	0.00	0.00
475	A:PRO 496	5.92	0.00	0.00
476	A:ALA 497	0.00	0.00	0.00
477	A:LEU 498	5.34	0.00	0.00
478	A:GLN 499	11.61	0.00	0.00
479	A:PHE 500	2.48	0.00	0.00
480	A:SER 501	38.81	0.00	0.00
481	A:THR 502	69.60	0.00	0.00
482	A:VAL 503	15.73	0.00	0.00
483	A:SER 504	20.42	0.00	0.00
484	A:ASP 505	68.86	0.00	0.00
485	A:ARG 506	85.87	0.00	0.00
486	A:SER 507	0.17	0.00	0.00
487	A:PHE 508	21.35	0.00	0.00
488	A:LEU 509	7.41	0.00	0.00
489	A:GLU 510	34.56	0.00	0.00
490	A:ASP 511	11.78	0.00	0.00
491	A:THR 512	15.04	0.00	0.00
492	A:PRO 513	7.50	0.00	0.00
493	A:ASP 514	17.23	0.00	0.00
494	A:GLN 515	7.42	0.00	0.00
495	A:ALA 516	0.49	0.00	0.00
496	A:THR 517	3.36	0.00	0.00

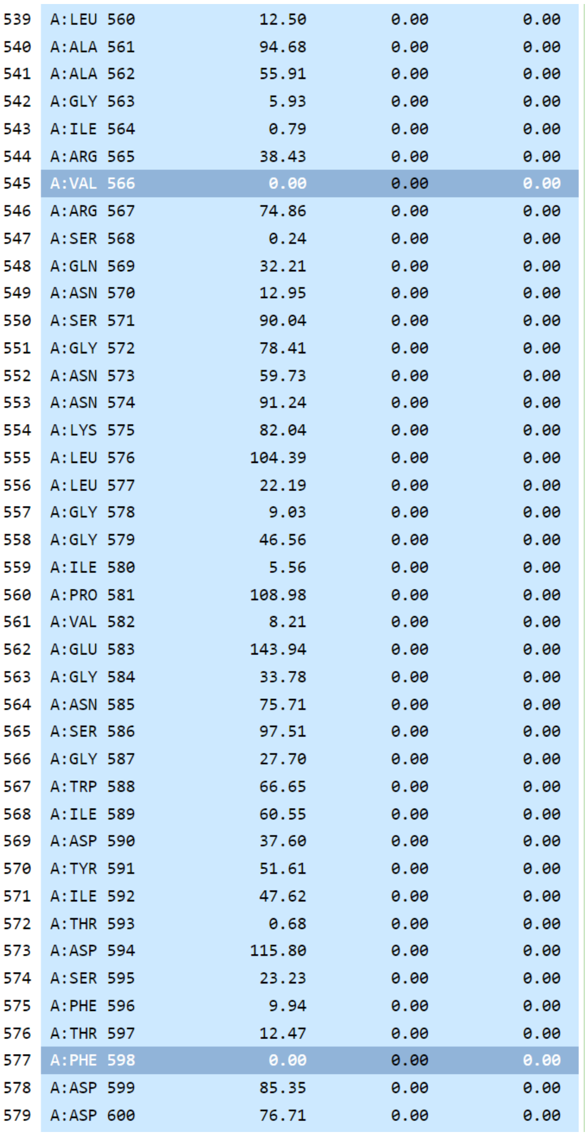


497	B:PHE 523	0.00	0.00	0.00
498	B:THR 524	10.64	0.00	0.00
499	B:ASP 525	14.52	0.00	0.00
500	B:THR 526	88.84	0.00	0.00
501	B:VAL 527	119.65	0.00	0.00
502	B:LEU 528	164.66	0.00	0.00
503	B:GLY 529	17.32	0.00	0.00
504	B:ASN 530	0.85	0.00	0.00
505	B:GLU 531	63.04	0.00	0.00
506	B:ALA 532	0.00	0.00	0.00
507	B:LYS 533	52.67	0.00	0.00
508	B:TYR 534	0.00	0.00	0.00
509	B:SER 535	13.09	0.00	0.00
510	B:ILE 536	0.37	0.00	0.00
511	B:ARG 537	68.05	0.00	0.00
512	B:LEU 538	7.32	0.00	0.00
513	B:ASN 539	58.20	0.00	0.00
514	B:THR 540	123.19	0.00	0.00
515	B:GLY 541	68.42	0.00	0.00
516	B:PHE 542	13.04	0.00	0.00
517	B:ASN 543	83.42	0.00	0.00
518	B:THR 544	34.15	0.00	0.00
519	B:ALA 545	63.42	0.00	0.00
520	B:THR 546	36.94	0.00	0.00
521	B:ARG 547	107.84	0.00	0.00
522	B:TYR 548	6.25	0.00	0.00
523	B:ARG 549	64.37	0.00	0.00
524	B:LEU 550	0.00	0.00	0.00
525	B:ILE 551	0.95	0.00	0.00
526	B:ILE 552	0.35	0.00	0.00
527	B:ARG 553	4.70	0.00	0.00
528	B:PHE 554	7.85	0.00	0.00
529	B:LYS 555	48.33	0.00	0.00
530	B:ALA 556	3.09	0.00	0.00
531	B:PRO 557	62.45	31.42 |||||	0.33
532	B:ALA 558	24.38	0.00	0.00
533	B:ARG 559	136.41	0.00	0.00
534	B:LEU 560	13.04	0.00	0.00
535	B:ALA 561	91.18	0.00	0.00
536	B:ALA 562	55.14	0.00	0.00
537	B:GLY 563	6.20	0.00	0.00
538	B:ILE 564	0.82	0.00	0.00

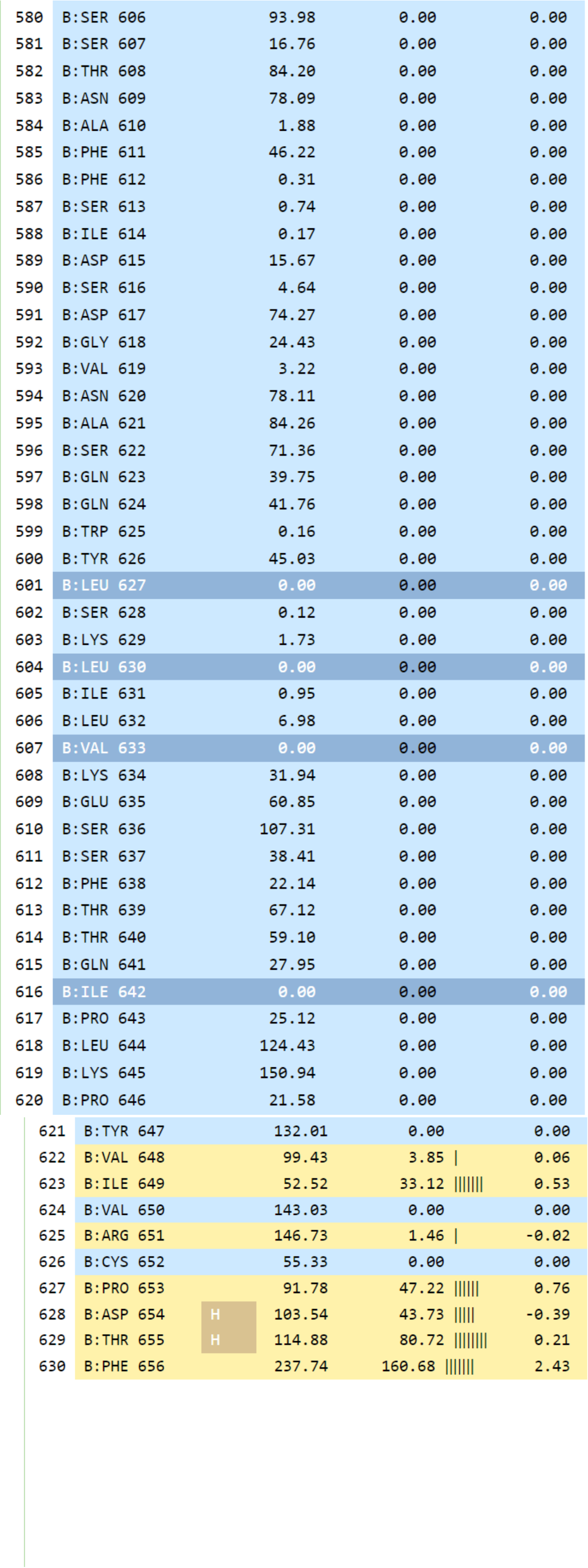
497	A:ASP 518	15.11	0.00	0.00
498	A:GLY 519	0.12	0.00	0.00
499	A:SER 520	0.00	0.00	0.00
500	A:ILE 521	0.00	0.00	0.00
501	A:LYS 522	45.33	0.00	0.00
502	A:PHE 523	0.00	0.00	0.00
503	A:THR 524	9.40	0.00	0.00
504	A:ASP 525	14.66	0.00	0.00
505	A:THR 526	89.01	0.00	0.00
506	A:VAL 527	119.68	0.00	0.00
507	A:LEU 528	162.25	0.00	0.00
508	A:GLY 529	15.14	0.00	0.00
509	A:ASN 530	2.02	0.00	0.00
510	A:GLU 531	65.53	0.00	0.00
511	A:ALA 532	0.17	0.00	0.00
512	A:LYS 533	49.34	0.00	0.00
513	A:TYR 534	0.00	0.00	0.00
514	A:SER 535	14.65	0.00	0.00
515	A:ILE 536	0.91	0.00	0.00
516	A:ARG 537	67.88	0.00	0.00
517	A:LEU 538	8.17	0.00	0.00
518	A:ASN 539	55.08	0.00	0.00
519	A:THR 540	125.67	0.00	0.00
520	A:GLY 541	67.72	0.00	0.00
521	A:PHE 542	15.24	0.00	0.00
522	A:ASN 543	78.89	0.00	0.00
523	A:THR 544	27.98	0.00	0.00
524	A:ALA 545	66.67	0.00	0.00
525	A:THR 546	36.07	0.00	0.00
526	A:ARG 547	111.16	0.00	0.00
527	A:TYR 548	4.62	0.00	0.00
528	A:ARG 549	73.86	0.00	0.00
529	A:LEU 550	0.00	0.00	0.00
530	A:ILE 551	0.62	0.00	0.00
531	A:ILE 552	0.66	0.00	0.00
532	A:ARG 553	4.45	0.00	0.00
533	A:PHE 554	6.43	0.00	0.00
534	A:LYS 555	52.28	0.00	0.00
535	A:ALA 556	2.67	0.00	0.00
536	A:PRO 557	63.72	0.00	0.00
537	A:ALA 558	24.64	0.00	0.00
538	A:ARG 559	142.96	0.00	0.00



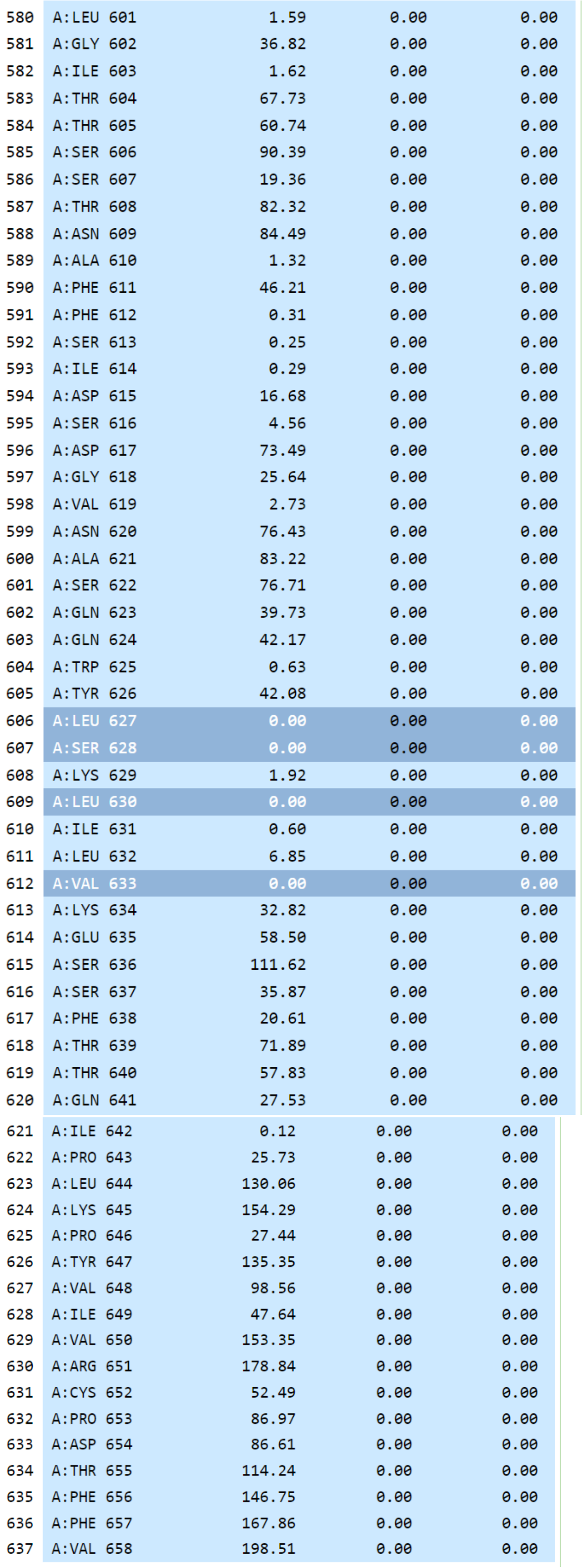
539	B:ARG 565		41.20	0.00	0.00
540	B:VAL 566		0.00	0.00	0.00
541	B:ARG 567		73.67	0.00	0.00
542	B:SER 568		1.70	0.00	0.00
543	B:GLN 569		29.44	0.00	0.00
544	B:ASN 570		11.84	0.00	0.00
545	B:SER 571		94.53	0.00	0.00
546	B:GLY 572		72.47	0.00	0.00
547	B:ASN 573		56.42	0.00	0.00
548	B:ASN 574		90.41	0.00	0.00
549	B:LYS 575		85.71	0.00	0.00
550	B:LEU 576		101.17	0.00	0.00
551	B:LEU 577		22.87	0.00	0.00
552	B:GLY 578		10.03	0.00	0.00
553	B:GLY 579		45.47	0.00	0.00
554	B:ILE 580		5.59	0.00	0.00
555	B:PRO 581		110.48	0.00	0.00
556	B:VAL 582		10.48	0.00	0.00
557	B:GLU 583		139.53	0.00	0.00
558	B:GLY 584		35.11	0.43 |	0.00
559	B:ASN 585		75.47	4.51 |	0.07
560	B:SER 586	H	99.18	43.85 |||||	0.33
561	B:GLY 587		27.71	0.00	0.00
562	B:TRP 588		64.71	0.00	0.00
563	B:ILE 589		60.06	0.00	0.00
564	B:ASP 590		31.35	0.00	0.00
565	B:TYR 591		53.40	0.00	0.00
566	B:ILE 592		47.35	0.00	0.00
567	B:THR 593		0.15	0.00	0.00
568	B:ASP 594		112.25	0.00	0.00
569	B:SER 595		26.42	0.00	0.00
570	B:PHE 596		9.12	0.00	0.00
571	B:THR 597		10.73	0.00	0.00
572	B:PHE 598		0.00	0.00	0.00
573	B:ASP 599		82.58	0.00	0.00
574	B:ASP 600		77.86	0.00	0.00
575	B:LEU 601		2.93	0.00	0.00
576	B:GLY 602		38.55	0.00	0.00
577	B:ILE 603		2.03	0.00	0.00
578	B:THR 604		64.58	0.00	0.00
579	B:THR 605		59.34	0.00	0.00



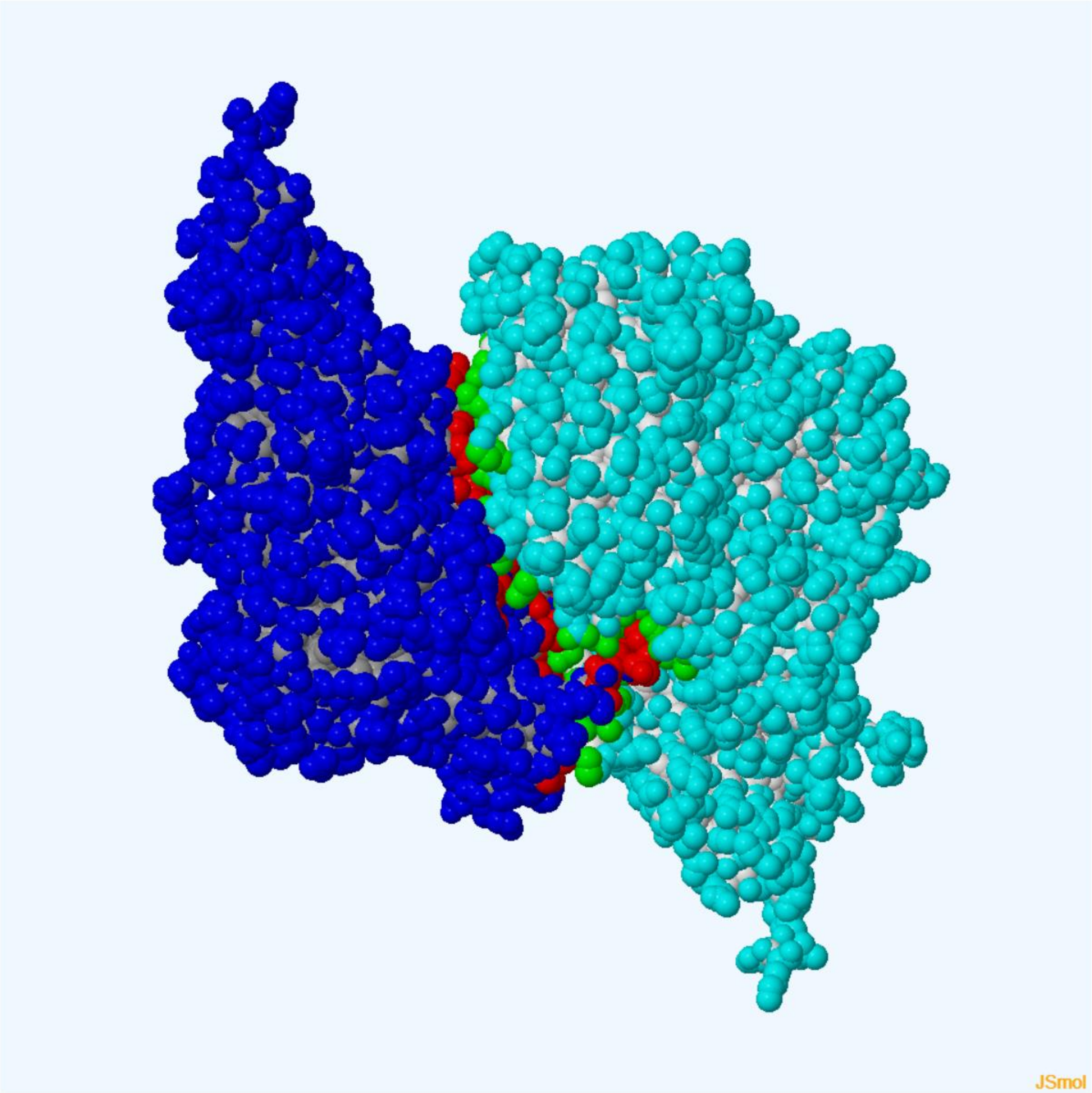
539	A:LEU 560	12.50	0.00	0.00
540	A:ALA 561	94.68	0.00	0.00
541	A:ALA 562	55.91	0.00	0.00
542	A:GLY 563	5.93	0.00	0.00
543	A:ILE 564	0.79	0.00	0.00
544	A:ARG 565	38.43	0.00	0.00
545	A:VAL 566	0.00	0.00	0.00
546	A:ARG 567	74.86	0.00	0.00
547	A:SER 568	0.24	0.00	0.00
548	A:GLN 569	32.21	0.00	0.00
549	A:ASN 570	12.95	0.00	0.00
550	A:SER 571	90.04	0.00	0.00
551	A:GLY 572	78.41	0.00	0.00
552	A:ASN 573	59.73	0.00	0.00
553	A:ASN 574	91.24	0.00	0.00
554	A:LYS 575	82.04	0.00	0.00
555	A:LEU 576	104.39	0.00	0.00
556	A:LEU 577	22.19	0.00	0.00
557	A:GLY 578	9.03	0.00	0.00
558	A:GLY 579	46.56	0.00	0.00
559	A:ILE 580	5.56	0.00	0.00
560	A:PRO 581	108.98	0.00	0.00
561	A:VAL 582	8.21	0.00	0.00
562	A:GLU 583	143.94	0.00	0.00
563	A:GLY 584	33.78	0.00	0.00
564	A:ASN 585	75.71	0.00	0.00
565	A:SER 586	97.51	0.00	0.00
566	A:GLY 587	27.70	0.00	0.00
567	A:TRP 588	66.65	0.00	0.00
568	A:ILE 589	60.55	0.00	0.00
569	A:ASP 590	37.60	0.00	0.00
570	A:TYR 591	51.61	0.00	0.00
571	A:ILE 592	47.62	0.00	0.00
572	A:THR 593	0.68	0.00	0.00
573	A:ASP 594	115.80	0.00	0.00
574	A:SER 595	23.23	0.00	0.00
575	A:PHE 596	9.94	0.00	0.00
576	A:THR 597	12.47	0.00	0.00
577	A:PHE 598	0.00	0.00	0.00
578	A:ASP 599	85.35	0.00	0.00
579	A:ASP 600	76.71	0.00	0.00



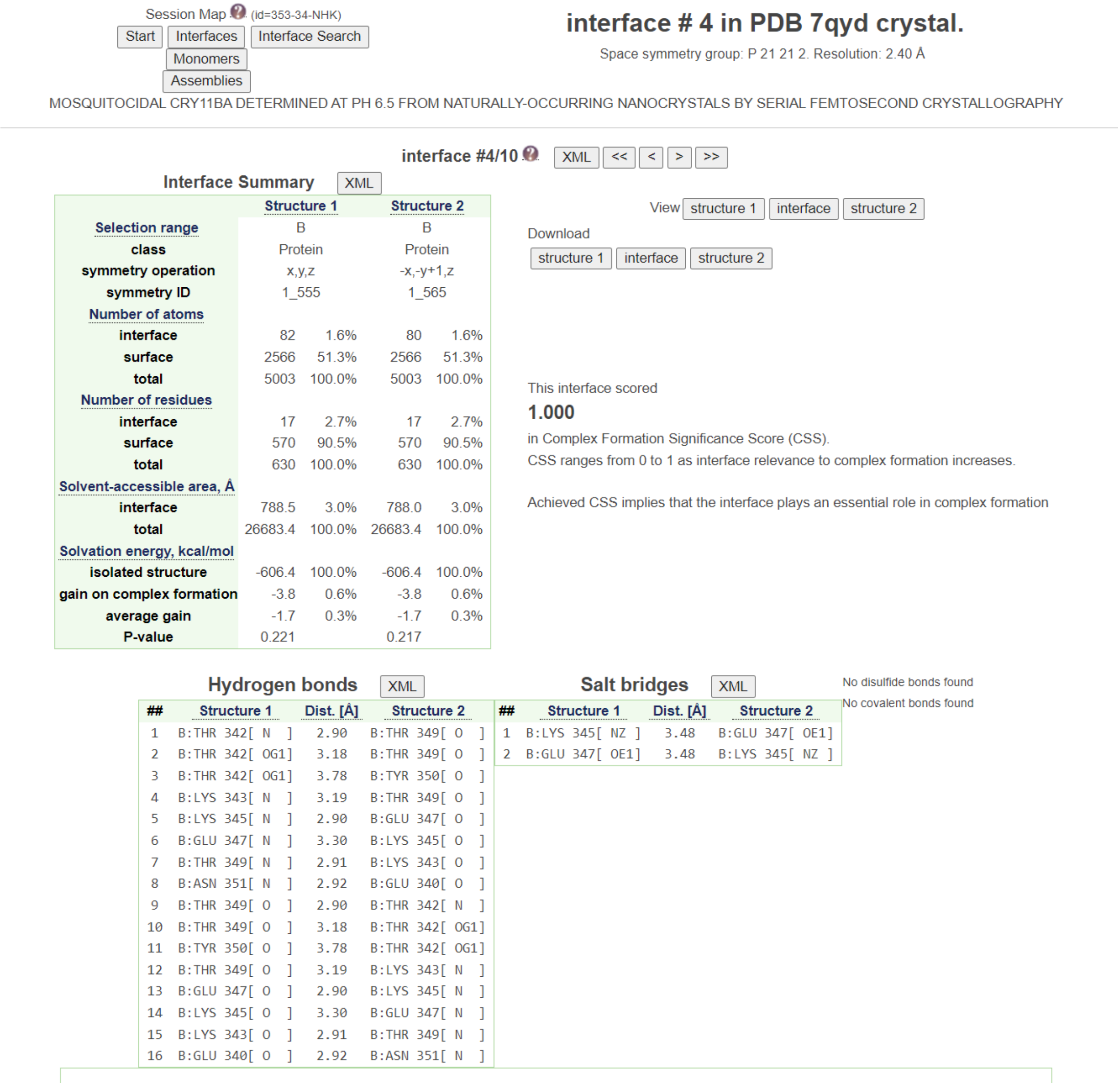
580	B:SER 606		93.98	0.00		0.00
581	B:SER 607		16.76	0.00		0.00
582	B:THR 608		84.20	0.00		0.00
583	B:ASN 609		78.09	0.00		0.00
584	B:ALA 610		1.88	0.00		0.00
585	B:PHE 611		46.22	0.00		0.00
586	B:PHE 612		0.31	0.00		0.00
587	B:SER 613		0.74	0.00		0.00
588	B:ILE 614		0.17	0.00		0.00
589	B:ASP 615		15.67	0.00		0.00
590	B:SER 616		4.64	0.00		0.00
591	B:ASP 617		74.27	0.00		0.00
592	B:GLY 618		24.43	0.00		0.00
593	B:VAL 619		3.22	0.00		0.00
594	B:ASN 620		78.11	0.00		0.00
595	B:ALA 621		84.26	0.00		0.00
596	B:SER 622		71.36	0.00		0.00
597	B:GLN 623		39.75	0.00		0.00
598	B:GLN 624		41.76	0.00		0.00
599	B:TRP 625		0.16	0.00		0.00
600	B:TYR 626		45.03	0.00		0.00
601	B:LEU 627		0.00	0.00		0.00
602	B:SER 628		0.12	0.00		0.00
603	B:LYS 629		1.73	0.00		0.00
604	B:LEU 630		0.00	0.00		0.00
605	B:ILE 631		0.95	0.00		0.00
606	B:LEU 632		6.98	0.00		0.00
607	B:VAL 633		0.00	0.00		0.00
608	B:LYS 634		31.94	0.00		0.00
609	B:GLU 635		60.85	0.00		0.00
610	B:SER 636		107.31	0.00		0.00
611	B:SER 637		38.41	0.00		0.00
612	B:PHE 638		22.14	0.00		0.00
613	B:THR 639		67.12	0.00		0.00
614	B:THR 640		59.10	0.00		0.00
615	B:GLN 641		27.95	0.00		0.00
616	B:ILE 642		0.00	0.00		0.00
617	B:PRO 643		25.12	0.00		0.00
618	B:LEU 644		124.43	0.00		0.00
619	B:LYS 645		150.94	0.00		0.00
620	B:PRO 646		21.58	0.00		0.00
621	B:TYR 647		132.01	0.00		0.00
622	B:VAL 648		99.43	3.85	|	0.06
623	B:ILE 649		52.52	33.12	|||||||	0.53
624	B:VAL 650		143.03	0.00		0.00
625	B:ARG 651		146.73	1.46	|	-0.02
626	B:CYS 652		55.33	0.00		0.00
627	B:PRO 653		91.78	47.22	||||||	0.76
628	B:ASP 654	H	103.54	43.73	|||||	-0.39
629	B:THR 655	H	114.88	80.72	||||||||	0.21
630	B:PHE 656		237.74	160.68	|||||||	2.43



580	A:LEU 601	1.59	0.00	0.00
581	A:GLY 602	36.82	0.00	0.00
582	A:ILE 603	1.62	0.00	0.00
583	A:THR 604	67.73	0.00	0.00
584	A:THR 605	60.74	0.00	0.00
585	A:SER 606	90.39	0.00	0.00
586	A:SER 607	19.36	0.00	0.00
587	A:THR 608	82.32	0.00	0.00
588	A:ASN 609	84.49	0.00	0.00
589	A:ALA 610	1.32	0.00	0.00
590	A:PHE 611	46.21	0.00	0.00
591	A:PHE 612	0.31	0.00	0.00
592	A:SER 613	0.25	0.00	0.00
593	A:ILE 614	0.29	0.00	0.00
594	A:ASP 615	16.68	0.00	0.00
595	A:SER 616	4.56	0.00	0.00
596	A:ASP 617	73.49	0.00	0.00
597	A:GLY 618	25.64	0.00	0.00
598	A:VAL 619	2.73	0.00	0.00
599	A:ASN 620	76.43	0.00	0.00
600	A:ALA 621	83.22	0.00	0.00
601	A:SER 622	76.71	0.00	0.00
602	A:GLN 623	39.73	0.00	0.00
603	A:GLN 624	42.17	0.00	0.00
604	A:TRP 625	0.63	0.00	0.00
605	A:TYR 626	42.08	0.00	0.00
606	A:LEU 627	0.00	0.00	0.00
607	A:SER 628	0.00	0.00	0.00
608	A:LYS 629	1.92	0.00	0.00
609	A:LEU 630	0.00	0.00	0.00
610	A:ILE 631	0.60	0.00	0.00
611	A:LEU 632	6.85	0.00	0.00
612	A:VAL 633	0.00	0.00	0.00
613	A:LYS 634	32.82	0.00	0.00
614	A:GLU 635	58.50	0.00	0.00
615	A:SER 636	111.62	0.00	0.00
616	A:SER 637	35.87	0.00	0.00
617	A:PHE 638	20.61	0.00	0.00
618	A:THR 639	71.89	0.00	0.00
619	A:THR 640	57.83	0.00	0.00
620	A:GLN 641	27.53	0.00	0.00
621	A:ILE 642	0.12	0.00	0.00
622	A:PRO 643	25.73	0.00	0.00
623	A:LEU 644	130.06	0.00	0.00
624	A:LYS 645	154.29	0.00	0.00
625	A:PRO 646	27.44	0.00	0.00
626	A:TYR 647	135.35	0.00	0.00
627	A:VAL 648	98.56	0.00	0.00
628	A:ILE 649	47.64	0.00	0.00
629	A:VAL 650	153.35	0.00	0.00
630	A:ARG 651	178.84	0.00	0.00
631	A:CYS 652	52.49	0.00	0.00
632	A:PRO 653	86.97	0.00	0.00
633	A:ASP 654	86.61	0.00	0.00
634	A:THR 655	114.24	0.00	0.00
635	A:PHE 656	146.75	0.00	0.00
636	A:PHE 657	167.86	0.00	0.00
637	A:VAL 658	198.51	0.00	0.00



JSmol



interface # 4 in PDB 7qyd crystal.

Space symmetry group: P 21 21 2. Resolution: 2.40 Å

MOSQUITOCIDAL CRY11BA DETERMINED AT PH 6.5 FROM NATURALLY-OCCURRING NANOCRYSTALS BY SERIAL FEMTOSECOND CRYSTALLOGRAPHY

interface #4/10

Interface Summary

	Structure 1		Structure 2	
Selection range	B		B	
class	Protein		Protein	
symmetry operation	x,y,z		-x,-y+1,z	
symmetry ID	1_555		1_565	
Number of atoms				
interface	82	1.6%	80	1.6%
surface	2566	51.3%	2566	51.3%
total	5003	100.0%	5003	100.0%
Number of residues				
interface	17	2.7%	17	2.7%
surface	570	90.5%	570	90.5%
total	630	100.0%	630	100.0%
Solvent-accessible area, Å				
interface	788.5	3.0%	788.0	3.0%
total	26683.4	100.0%	26683.4	100.0%
Solvation energy, kcal/mol				
isolated structure	-606.4	100.0%	-606.4	100.0%
gain on complex formation	-3.8	0.6%	-3.8	0.6%
average gain	-1.7	0.3%	-1.7	0.3%
P-value	0.221		0.217	

This interface scored

1.000

in Complex Formation Significance Score (CSS).

CSS ranges from 0 to 1 as interface relevance to complex formation increases.

Achieved CSS implies that the interface plays an essential role in complex formation

Hydrogen bonds

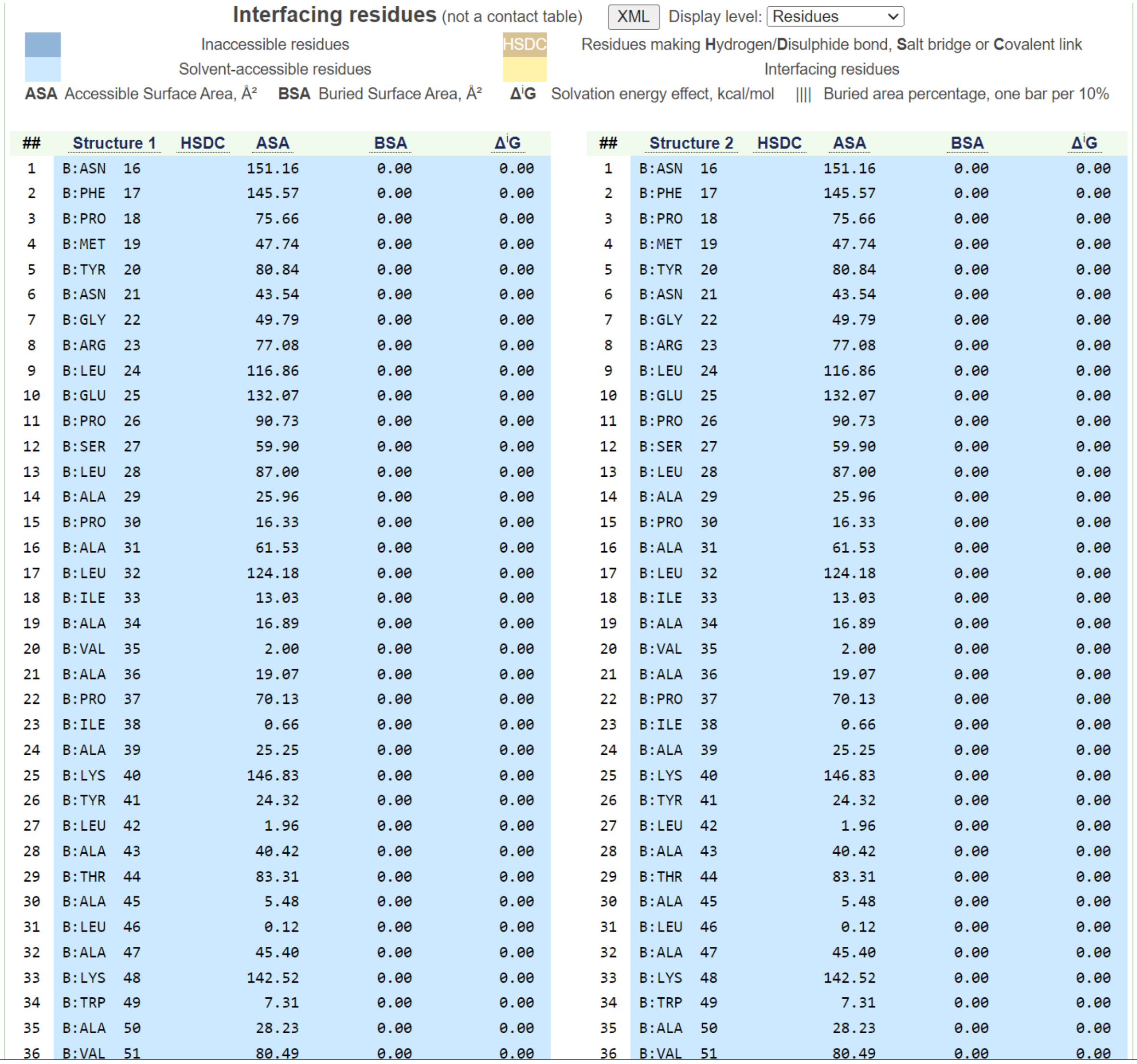
##	Structure 1	Dist. [Å]	Structure 2
1	B:THR 342[N]	2.90	B:THR 349[O]
2	B:THR 342[OG1]	3.18	B:THR 349[O]
3	B:THR 342[OG1]	3.78	B:TYR 350[O]
4	B:LYS 343[N]	3.19	B:THR 349[O]
5	B:LYS 345[N]	2.90	B:GLU 347[O]
6	B:GLU 347[N]	3.30	B:LYS 345[O]
7	B:THR 349[N]	2.91	B:LYS 343[O]
8	B:ASN 351[N]	2.92	B:GLU 340[O]
9	B:THR 349[O]	2.90	B:THR 342[N]
10	B:THR 349[O]	3.18	B:THR 342[OG1]
11	B:TYR 350[O]	3.78	B:THR 342[OG1]
12	B:THR 349[O]	3.19	B:LYS 343[N]
13	B:GLU 347[O]	2.90	B:LYS 345[N]
14	B:LYS 345[O]	3.30	B:GLU 347[N]
15	B:LYS 343[O]	2.91	B:THR 349[N]
16	B:GLU 340[O]	2.92	B:ASN 351[N]

Salt bridges

##	Structure 1	Dist. [Å]	Structure 2
1	B:LYS 345[NZ]	3.48	B:GLU 347[OE1]
2	B:GLU 347[OE1]	3.48	B:LYS 345[NZ]

No disulfide bonds found

No covalent bonds found



Interfacing residues (not a contact table)

▨ Inaccessible residues

□ Solvent-accessible residues

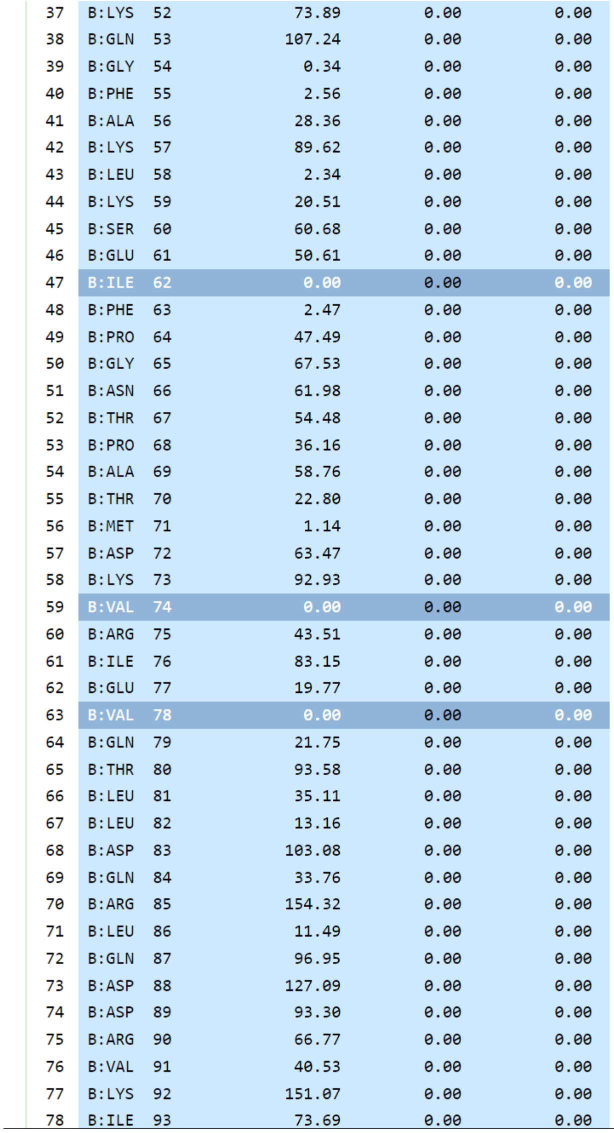
HSDC Residues making **H**ydrogen/**D**isulphide bond, **S**alt bridge or **C**ovalent link

□ Interfacing residues

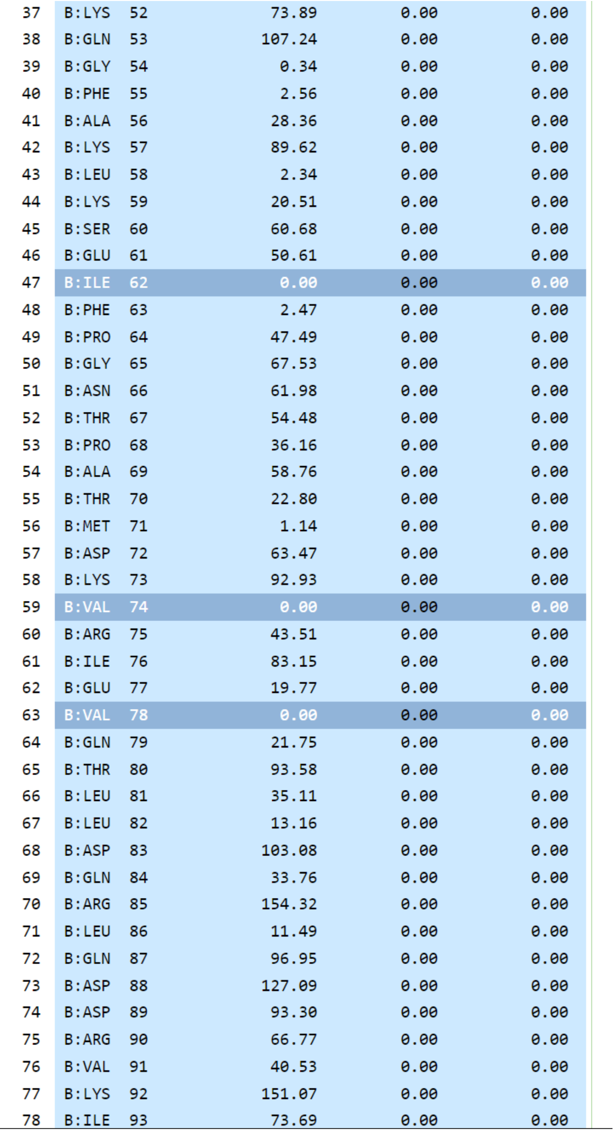
ASA Accessible Surface Area, Å^2 **BSA** Buried Surface Area, Å^2 **Δ^iG** Solvation energy effect, kcal/mol |||| Buried area percentage, one bar per 10%

##	Structure 1	HSDC	ASA	BSA	Δ^iG
1	B:ASN 16		151.16	0.00	0.00
2	B:PHE 17		145.57	0.00	0.00
3	B:PRO 18		75.66	0.00	0.00
4	B:MET 19		47.74	0.00	0.00
5	B:TYR 20		80.84	0.00	0.00
6	B:ASN 21		43.54	0.00	0.00
7	B:GLY 22		49.79	0.00	0.00
8	B:ARG 23		77.08	0.00	0.00
9	B:LEU 24		116.86	0.00	0.00
10	B:GLU 25		132.07	0.00	0.00
11	B:PRO 26		90.73	0.00	0.00
12	B:SER 27		59.90	0.00	0.00
13	B:LEU 28		87.00	0.00	0.00
14	B:ALA 29		25.96	0.00	0.00
15	B:PRO 30		16.33	0.00	0.00
16	B:ALA 31		61.53	0.00	0.00
17	B:LEU 32		124.18	0.00	0.00
18	B:ILE 33		13.03	0.00	0.00
19	B:ALA 34		16.89	0.00	0.00
20	B:VAL 35		2.00	0.00	0.00
21	B:ALA 36		19.07	0.00	0.00
22	B:PRO 37		70.13	0.00	0.00
23	B:ILE 38		0.66	0.00	0.00
24	B:ALA 39		25.25	0.00	0.00
25	B:LYS 40		146.83	0.00	0.00
26	B:TYR 41		24.32	0.00	0.00
27	B:LEU 42		1.96	0.00	0.00
28	B:ALA 43		40.42	0.00	0.00
29	B:THR 44		83.31	0.00	0.00
30	B:ALA 45		5.48	0.00	0.00
31	B:LEU 46		0.12	0.00	0.00
32	B:ALA 47		45.40	0.00	0.00
33	B:LYS 48		142.52	0.00	0.00
34	B:TRP 49		7.31	0.00	0.00
35	B:ALA 50		28.23	0.00	0.00
36	B:VAL 51		80.49	0.00	0.00

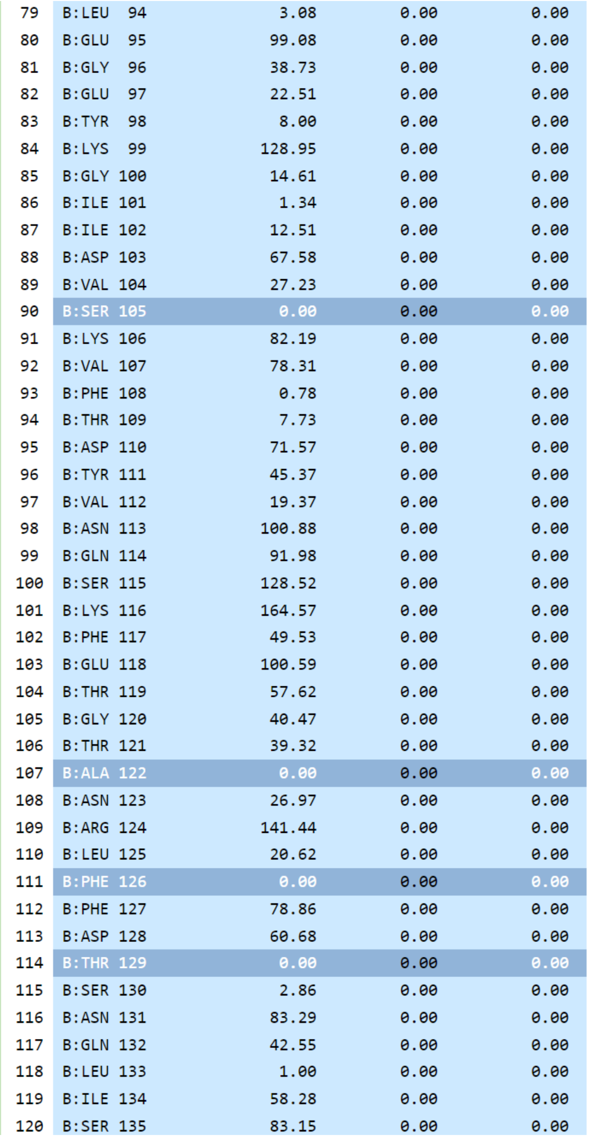
##	Structure 2	HSDC	ASA	BSA	Δ^iG
1	B:ASN 16		151.16	0.00	0.00
2	B:PHE 17		145.57	0.00	0.00
3	B:PRO 18		75.66	0.00	0.00
4	B:MET 19		47.74	0.00	0.00
5	B:TYR 20		80.84	0.00	0.00
6	B:ASN 21		43.54	0.00	0.00
7	B:GLY 22		49.79	0.00	0.00
8	B:ARG 23		77.08	0.00	0.00
9	B:LEU 24		116.86	0.00	0.00
10	B:GLU 25		132.07	0.00	0.00
11	B:PRO 26		90.73	0.00	0.00
12	B:SER 27		59.90	0.00	0.00
13	B:LEU 28		87.00	0.00	0.00
14	B:ALA 29		25.96	0.00	0.00
15	B:PRO 30		16.33	0.00	0.00
16	B:ALA 31		61.53	0.00	0.00
17	B:LEU 32		124.18	0.00	0.00
18	B:ILE 33		13.03	0.00	0.00
19	B:ALA 34		16.89	0.00	0.00
20	B:VAL 35		2.00	0.00	0.00
21	B:ALA 36		19.07	0.00	0.00
22	B:PRO 37		70.13	0.00	0.00
23	B:ILE 38		0.66	0.00	0.00
24	B:ALA 39		25.25	0.00	0.00
25	B:LYS 40		146.83	0.00	0.00
26	B:TYR 41		24.32	0.00	0.00
27	B:LEU 42		1.96	0.00	0.00
28	B:ALA 43		40.42	0.00	0.00
29	B:THR 44		83.31	0.00	0.00
30	B:ALA 45		5.48	0.00	0.00
31	B:LEU 46		0.12	0.00	0.00
32	B:ALA 47		45.40	0.00	0.00
33	B:LYS 48		142.52	0.00	0.00
34	B:TRP 49		7.31	0.00	0.00
35	B:ALA 50		28.23	0.00	0.00
36	B:VAL 51		80.49	0.00	0.00



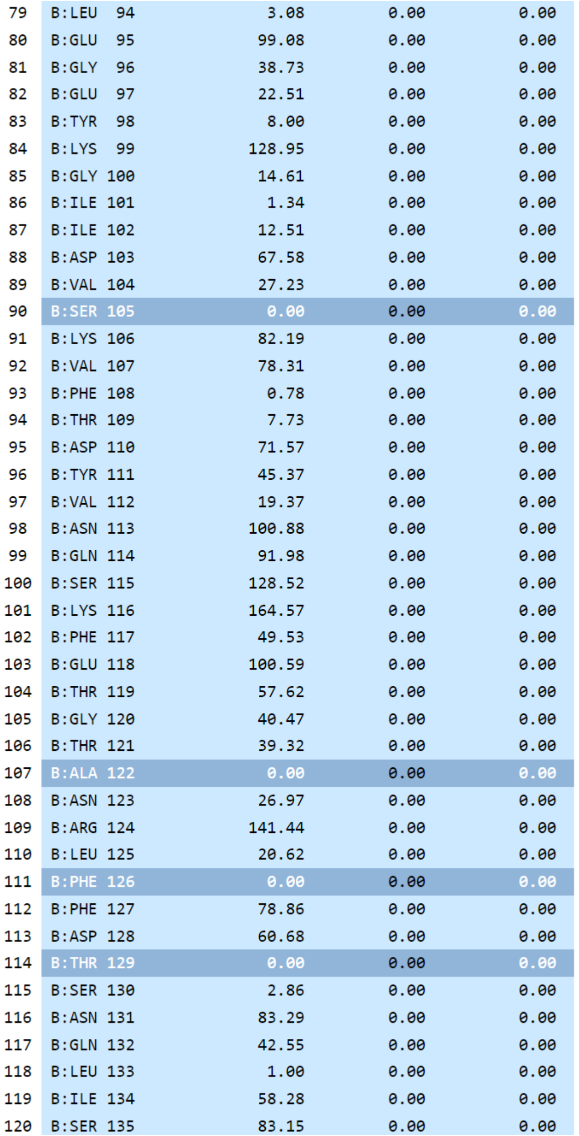
37	B:LYS	52	73.89	0.00	0.00
38	B:GLN	53	107.24	0.00	0.00
39	B:GLY	54	0.34	0.00	0.00
40	B:PHE	55	2.56	0.00	0.00
41	B:ALA	56	28.36	0.00	0.00
42	B:LYS	57	89.62	0.00	0.00
43	B:LEU	58	2.34	0.00	0.00
44	B:LYS	59	20.51	0.00	0.00
45	B:SER	60	60.68	0.00	0.00
46	B:GLU	61	50.61	0.00	0.00
47	B:ILE	62	0.00	0.00	0.00
48	B:PHE	63	2.47	0.00	0.00
49	B:PRO	64	47.49	0.00	0.00
50	B:GLY	65	67.53	0.00	0.00
51	B:ASN	66	61.98	0.00	0.00
52	B:THR	67	54.48	0.00	0.00
53	B:PRO	68	36.16	0.00	0.00
54	B:ALA	69	58.76	0.00	0.00
55	B:THR	70	22.80	0.00	0.00
56	B:MET	71	1.14	0.00	0.00
57	B:ASP	72	63.47	0.00	0.00
58	B:LYS	73	92.93	0.00	0.00
59	B:VAL	74	0.00	0.00	0.00
60	B:ARG	75	43.51	0.00	0.00
61	B:ILE	76	83.15	0.00	0.00
62	B:GLU	77	19.77	0.00	0.00
63	B:VAL	78	0.00	0.00	0.00
64	B:GLN	79	21.75	0.00	0.00
65	B:THR	80	93.58	0.00	0.00
66	B:LEU	81	35.11	0.00	0.00
67	B:LEU	82	13.16	0.00	0.00
68	B:ASP	83	103.08	0.00	0.00
69	B:GLN	84	33.76	0.00	0.00
70	B:ARG	85	154.32	0.00	0.00
71	B:LEU	86	11.49	0.00	0.00
72	B:GLN	87	96.95	0.00	0.00
73	B:ASP	88	127.09	0.00	0.00
74	B:ASP	89	93.30	0.00	0.00
75	B:ARG	90	66.77	0.00	0.00
76	B:VAL	91	40.53	0.00	0.00
77	B:LYS	92	151.07	0.00	0.00
78	B:ILE	93	73.69	0.00	0.00



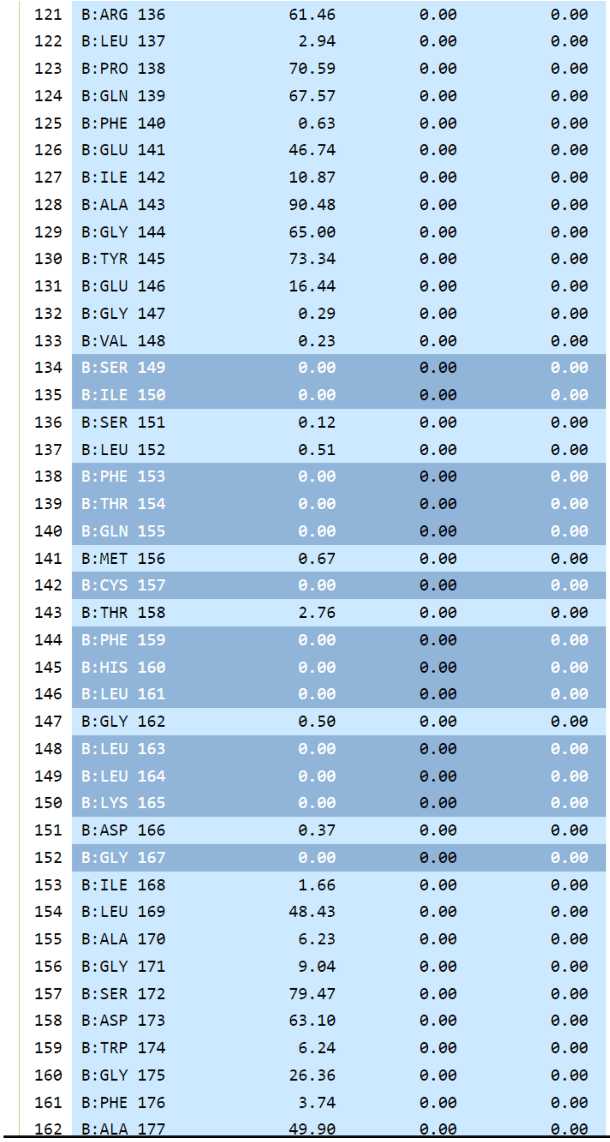
37	B:LYS	52	73.89	0.00	0.00
38	B:GLN	53	107.24	0.00	0.00
39	B:GLY	54	0.34	0.00	0.00
40	B:PHE	55	2.56	0.00	0.00
41	B:ALA	56	28.36	0.00	0.00
42	B:LYS	57	89.62	0.00	0.00
43	B:LEU	58	2.34	0.00	0.00
44	B:LYS	59	20.51	0.00	0.00
45	B:SER	60	60.68	0.00	0.00
46	B:GLU	61	50.61	0.00	0.00
47	B:ILE	62	0.00	0.00	0.00
48	B:PHE	63	2.47	0.00	0.00
49	B:PRO	64	47.49	0.00	0.00
50	B:GLY	65	67.53	0.00	0.00
51	B:ASN	66	61.98	0.00	0.00
52	B:THR	67	54.48	0.00	0.00
53	B:PRO	68	36.16	0.00	0.00
54	B:ALA	69	58.76	0.00	0.00
55	B:THR	70	22.80	0.00	0.00
56	B:MET	71	1.14	0.00	0.00
57	B:ASP	72	63.47	0.00	0.00
58	B:LYS	73	92.93	0.00	0.00
59	B:VAL	74	0.00	0.00	0.00
60	B:ARG	75	43.51	0.00	0.00
61	B:ILE	76	83.15	0.00	0.00
62	B:GLU	77	19.77	0.00	0.00
63	B:VAL	78	0.00	0.00	0.00
64	B:GLN	79	21.75	0.00	0.00
65	B:THR	80	93.58	0.00	0.00
66	B:LEU	81	35.11	0.00	0.00
67	B:LEU	82	13.16	0.00	0.00
68	B:ASP	83	103.08	0.00	0.00
69	B:GLN	84	33.76	0.00	0.00
70	B:ARG	85	154.32	0.00	0.00
71	B:LEU	86	11.49	0.00	0.00
72	B:GLN	87	96.95	0.00	0.00
73	B:ASP	88	127.09	0.00	0.00
74	B:ASP	89	93.30	0.00	0.00
75	B:ARG	90	66.77	0.00	0.00
76	B:VAL	91	40.53	0.00	0.00
77	B:LYS	92	151.07	0.00	0.00
78	B:ILE	93	73.69	0.00	0.00



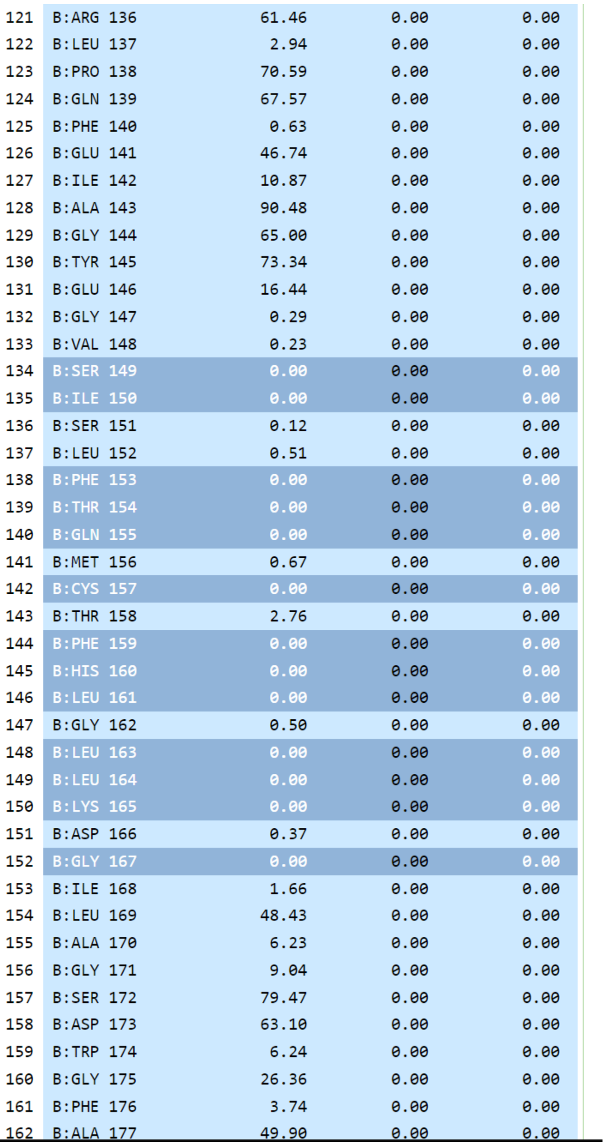
79	B:LEU 94	3.08	0.00	0.00
80	B:GLU 95	99.08	0.00	0.00
81	B:GLY 96	38.73	0.00	0.00
82	B:GLU 97	22.51	0.00	0.00
83	B:TYR 98	8.00	0.00	0.00
84	B:LYS 99	128.95	0.00	0.00
85	B:GLY 100	14.61	0.00	0.00
86	B:ILE 101	1.34	0.00	0.00
87	B:ILE 102	12.51	0.00	0.00
88	B:ASP 103	67.58	0.00	0.00
89	B:VAL 104	27.23	0.00	0.00
90	B:SER 105	0.00	0.00	0.00
91	B:LYS 106	82.19	0.00	0.00
92	B:VAL 107	78.31	0.00	0.00
93	B:PHE 108	0.78	0.00	0.00
94	B:THR 109	7.73	0.00	0.00
95	B:ASP 110	71.57	0.00	0.00
96	B:TYR 111	45.37	0.00	0.00
97	B:VAL 112	19.37	0.00	0.00
98	B:ASN 113	100.88	0.00	0.00
99	B:GLN 114	91.98	0.00	0.00
100	B:SER 115	128.52	0.00	0.00
101	B:LYS 116	164.57	0.00	0.00
102	B:PHE 117	49.53	0.00	0.00
103	B:GLU 118	100.59	0.00	0.00
104	B:THR 119	57.62	0.00	0.00
105	B:GLY 120	40.47	0.00	0.00
106	B:THR 121	39.32	0.00	0.00
107	B:ALA 122	0.00	0.00	0.00
108	B:ASN 123	26.97	0.00	0.00
109	B:ARG 124	141.44	0.00	0.00
110	B:LEU 125	20.62	0.00	0.00
111	B:PHE 126	0.00	0.00	0.00
112	B:PHE 127	78.86	0.00	0.00
113	B:ASP 128	60.68	0.00	0.00
114	B:THR 129	0.00	0.00	0.00
115	B:SER 130	2.86	0.00	0.00
116	B:ASN 131	83.29	0.00	0.00
117	B:GLN 132	42.55	0.00	0.00
118	B:LEU 133	1.00	0.00	0.00
119	B:ILE 134	58.28	0.00	0.00
120	B:SER 135	83.15	0.00	0.00



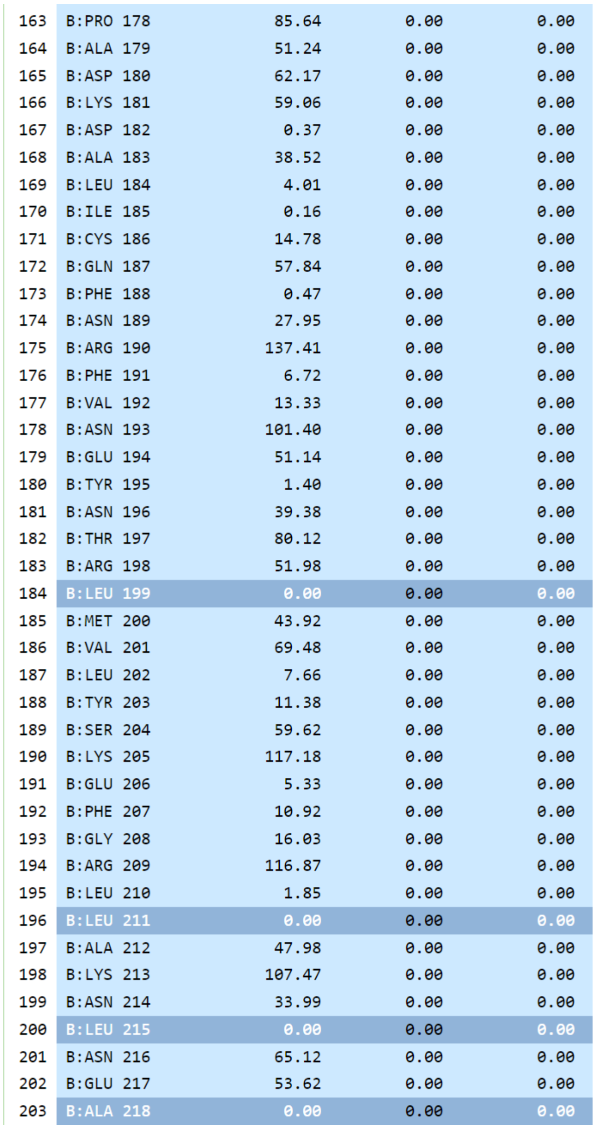
79	B:LEU 94	3.08	0.00	0.00
80	B:GLU 95	99.08	0.00	0.00
81	B:GLY 96	38.73	0.00	0.00
82	B:GLU 97	22.51	0.00	0.00
83	B:TYR 98	8.00	0.00	0.00
84	B:LYS 99	128.95	0.00	0.00
85	B:GLY 100	14.61	0.00	0.00
86	B:ILE 101	1.34	0.00	0.00
87	B:ILE 102	12.51	0.00	0.00
88	B:ASP 103	67.58	0.00	0.00
89	B:VAL 104	27.23	0.00	0.00
90	B:SER 105	0.00	0.00	0.00
91	B:LYS 106	82.19	0.00	0.00
92	B:VAL 107	78.31	0.00	0.00
93	B:PHE 108	0.78	0.00	0.00
94	B:THR 109	7.73	0.00	0.00
95	B:ASP 110	71.57	0.00	0.00
96	B:TYR 111	45.37	0.00	0.00
97	B:VAL 112	19.37	0.00	0.00
98	B:ASN 113	100.88	0.00	0.00
99	B:GLN 114	91.98	0.00	0.00
100	B:SER 115	128.52	0.00	0.00
101	B:LYS 116	164.57	0.00	0.00
102	B:PHE 117	49.53	0.00	0.00
103	B:GLU 118	100.59	0.00	0.00
104	B:THR 119	57.62	0.00	0.00
105	B:GLY 120	40.47	0.00	0.00
106	B:THR 121	39.32	0.00	0.00
107	B:ALA 122	0.00	0.00	0.00
108	B:ASN 123	26.97	0.00	0.00
109	B:ARG 124	141.44	0.00	0.00
110	B:LEU 125	20.62	0.00	0.00
111	B:PHE 126	0.00	0.00	0.00
112	B:PHE 127	78.86	0.00	0.00
113	B:ASP 128	60.68	0.00	0.00
114	B:THR 129	0.00	0.00	0.00
115	B:SER 130	2.86	0.00	0.00
116	B:ASN 131	83.29	0.00	0.00
117	B:GLN 132	42.55	0.00	0.00
118	B:LEU 133	1.00	0.00	0.00
119	B:ILE 134	58.28	0.00	0.00
120	B:SER 135	83.15	0.00	0.00



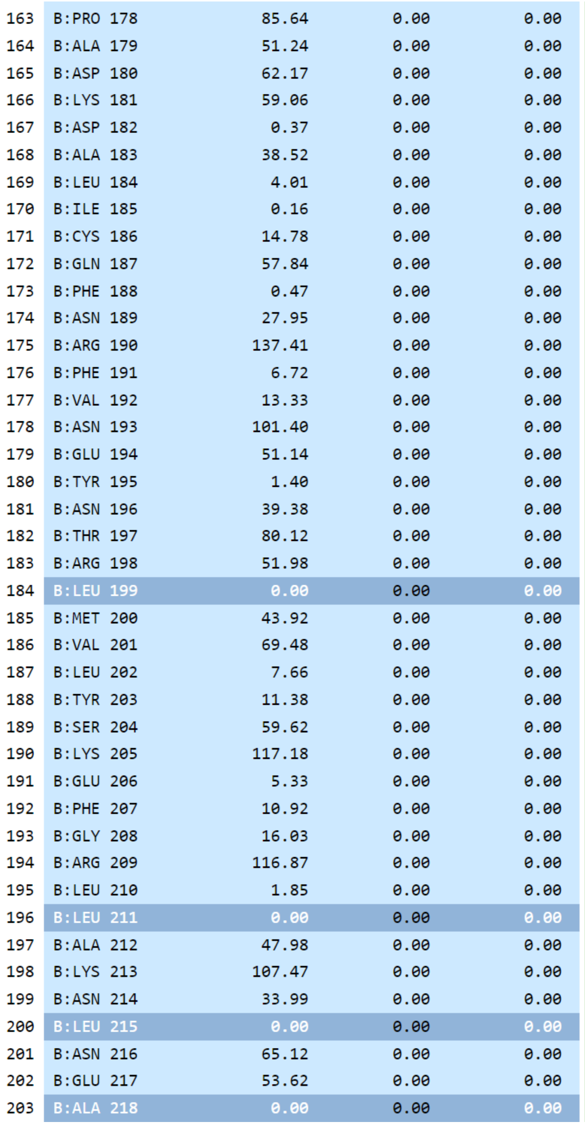
121	B:ARG 136	61.46	0.00	0.00
122	B:LEU 137	2.94	0.00	0.00
123	B:PRO 138	70.59	0.00	0.00
124	B:GLN 139	67.57	0.00	0.00
125	B:PHE 140	0.63	0.00	0.00
126	B:GLU 141	46.74	0.00	0.00
127	B:ILE 142	10.87	0.00	0.00
128	B:ALA 143	90.48	0.00	0.00
129	B:GLY 144	65.00	0.00	0.00
130	B:TYR 145	73.34	0.00	0.00
131	B:GLU 146	16.44	0.00	0.00
132	B:GLY 147	0.29	0.00	0.00
133	B:VAL 148	0.23	0.00	0.00
134	B:SER 149	0.00	0.00	0.00
135	B:ILE 150	0.00	0.00	0.00
136	B:SER 151	0.12	0.00	0.00
137	B:LEU 152	0.51	0.00	0.00
138	B:PHE 153	0.00	0.00	0.00
139	B:THR 154	0.00	0.00	0.00
140	B:GLN 155	0.00	0.00	0.00
141	B:MET 156	0.67	0.00	0.00
142	B:CYS 157	0.00	0.00	0.00
143	B:THR 158	2.76	0.00	0.00
144	B:PHE 159	0.00	0.00	0.00
145	B:HIS 160	0.00	0.00	0.00
146	B:LEU 161	0.00	0.00	0.00
147	B:GLY 162	0.50	0.00	0.00
148	B:LEU 163	0.00	0.00	0.00
149	B:LEU 164	0.00	0.00	0.00
150	B:LYS 165	0.00	0.00	0.00
151	B:ASP 166	0.37	0.00	0.00
152	B:GLY 167	0.00	0.00	0.00
153	B:ILE 168	1.66	0.00	0.00
154	B:LEU 169	48.43	0.00	0.00
155	B:ALA 170	6.23	0.00	0.00
156	B:GLY 171	9.04	0.00	0.00
157	B:SER 172	79.47	0.00	0.00
158	B:ASP 173	63.10	0.00	0.00
159	B:TRP 174	6.24	0.00	0.00
160	B:GLY 175	26.36	0.00	0.00
161	B:PHE 176	3.74	0.00	0.00
162	B:ALA 177	49.90	0.00	0.00



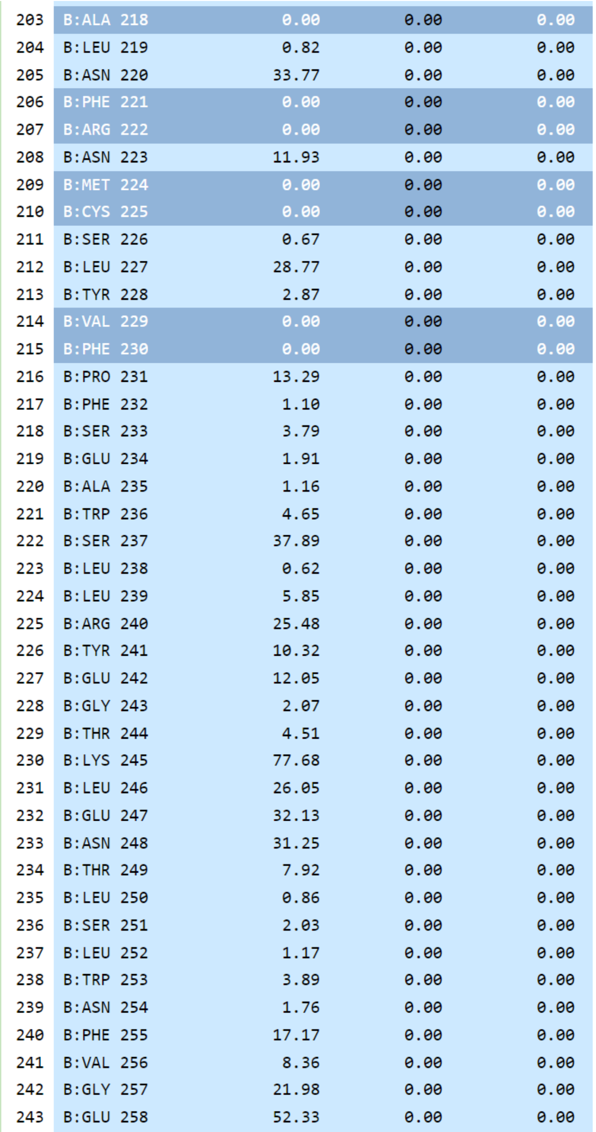
121	B:ARG 136	61.46	0.00	0.00
122	B:LEU 137	2.94	0.00	0.00
123	B:PRO 138	70.59	0.00	0.00
124	B:GLN 139	67.57	0.00	0.00
125	B:PHE 140	0.63	0.00	0.00
126	B:GLU 141	46.74	0.00	0.00
127	B:ILE 142	10.87	0.00	0.00
128	B:ALA 143	90.48	0.00	0.00
129	B:GLY 144	65.00	0.00	0.00
130	B:TYR 145	73.34	0.00	0.00
131	B:GLU 146	16.44	0.00	0.00
132	B:GLY 147	0.29	0.00	0.00
133	B:VAL 148	0.23	0.00	0.00
134	B:SER 149	0.00	0.00	0.00
135	B:ILE 150	0.00	0.00	0.00
136	B:SER 151	0.12	0.00	0.00
137	B:LEU 152	0.51	0.00	0.00
138	B:PHE 153	0.00	0.00	0.00
139	B:THR 154	0.00	0.00	0.00
140	B:GLN 155	0.00	0.00	0.00
141	B:MET 156	0.67	0.00	0.00
142	B:CYS 157	0.00	0.00	0.00
143	B:THR 158	2.76	0.00	0.00
144	B:PHE 159	0.00	0.00	0.00
145	B:HIS 160	0.00	0.00	0.00
146	B:LEU 161	0.00	0.00	0.00
147	B:GLY 162	0.50	0.00	0.00
148	B:LEU 163	0.00	0.00	0.00
149	B:LEU 164	0.00	0.00	0.00
150	B:LYS 165	0.00	0.00	0.00
151	B:ASP 166	0.37	0.00	0.00
152	B:GLY 167	0.00	0.00	0.00
153	B:ILE 168	1.66	0.00	0.00
154	B:LEU 169	48.43	0.00	0.00
155	B:ALA 170	6.23	0.00	0.00
156	B:GLY 171	9.04	0.00	0.00
157	B:SER 172	79.47	0.00	0.00
158	B:ASP 173	63.10	0.00	0.00
159	B:TRP 174	6.24	0.00	0.00
160	B:GLY 175	26.36	0.00	0.00
161	B:PHE 176	3.74	0.00	0.00
162	B:ALA 177	49.90	0.00	0.00



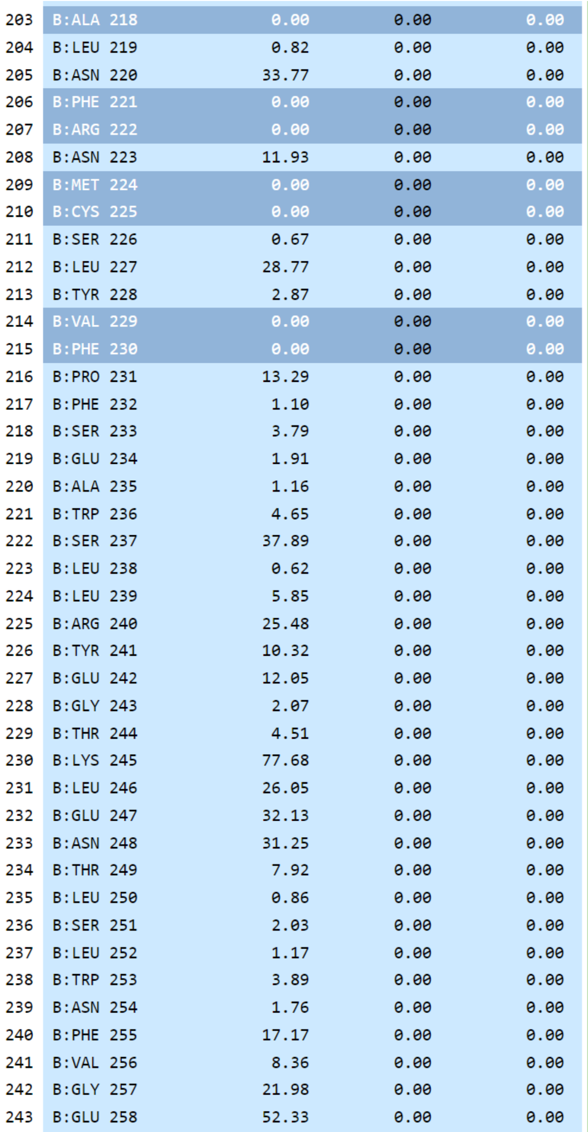
163	B:PRO 178	85.64	0.00	0.00
164	B:ALA 179	51.24	0.00	0.00
165	B:ASP 180	62.17	0.00	0.00
166	B:LYS 181	59.06	0.00	0.00
167	B:ASP 182	0.37	0.00	0.00
168	B:ALA 183	38.52	0.00	0.00
169	B:LEU 184	4.01	0.00	0.00
170	B:ILE 185	0.16	0.00	0.00
171	B:CYS 186	14.78	0.00	0.00
172	B:GLN 187	57.84	0.00	0.00
173	B:PHE 188	0.47	0.00	0.00
174	B:ASN 189	27.95	0.00	0.00
175	B:ARG 190	137.41	0.00	0.00
176	B:PHE 191	6.72	0.00	0.00
177	B:VAL 192	13.33	0.00	0.00
178	B:ASN 193	101.40	0.00	0.00
179	B:GLU 194	51.14	0.00	0.00
180	B:TYR 195	1.40	0.00	0.00
181	B:ASN 196	39.38	0.00	0.00
182	B:THR 197	80.12	0.00	0.00
183	B:ARG 198	51.98	0.00	0.00
184	B:LEU 199	0.00	0.00	0.00
185	B:MET 200	43.92	0.00	0.00
186	B:VAL 201	69.48	0.00	0.00
187	B:LEU 202	7.66	0.00	0.00
188	B:TYR 203	11.38	0.00	0.00
189	B:SER 204	59.62	0.00	0.00
190	B:LYS 205	117.18	0.00	0.00
191	B:GLU 206	5.33	0.00	0.00
192	B:PHE 207	10.92	0.00	0.00
193	B:GLY 208	16.03	0.00	0.00
194	B:ARG 209	116.87	0.00	0.00
195	B:LEU 210	1.85	0.00	0.00
196	B:LEU 211	0.00	0.00	0.00
197	B:ALA 212	47.98	0.00	0.00
198	B:LYS 213	107.47	0.00	0.00
199	B:ASN 214	33.99	0.00	0.00
200	B:LEU 215	0.00	0.00	0.00
201	B:ASN 216	65.12	0.00	0.00
202	B:GLU 217	53.62	0.00	0.00
203	B:ALA 218	0.00	0.00	0.00



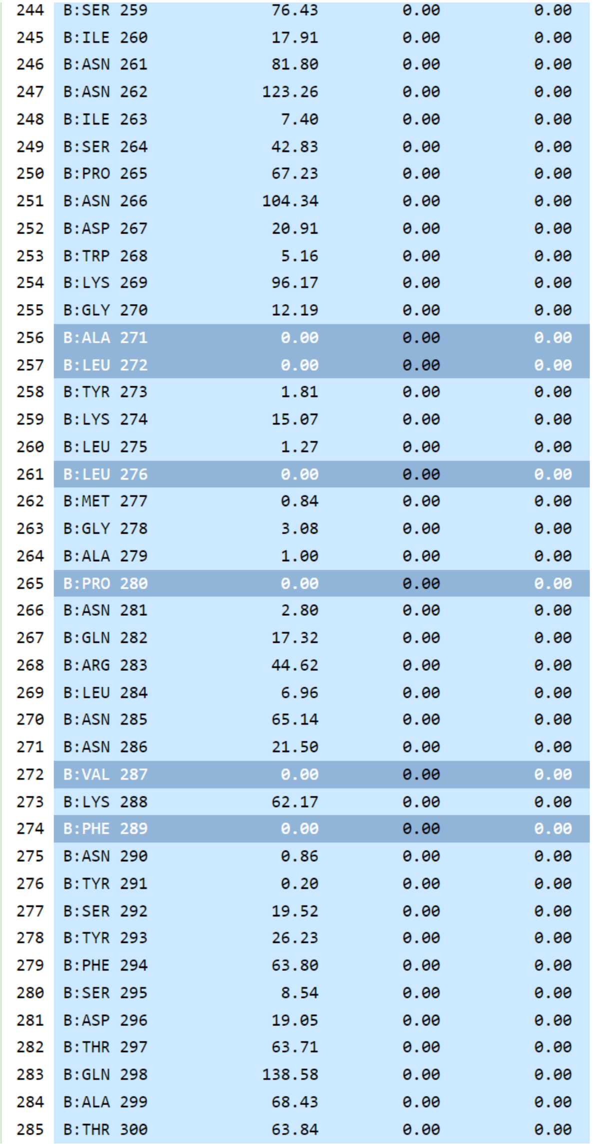
163	B:PRO 178	85.64	0.00	0.00
164	B:ALA 179	51.24	0.00	0.00
165	B:ASP 180	62.17	0.00	0.00
166	B:LYS 181	59.06	0.00	0.00
167	B:ASP 182	0.37	0.00	0.00
168	B:ALA 183	38.52	0.00	0.00
169	B:LEU 184	4.01	0.00	0.00
170	B:ILE 185	0.16	0.00	0.00
171	B:CYS 186	14.78	0.00	0.00
172	B:GLN 187	57.84	0.00	0.00
173	B:PHE 188	0.47	0.00	0.00
174	B:ASN 189	27.95	0.00	0.00
175	B:ARG 190	137.41	0.00	0.00
176	B:PHE 191	6.72	0.00	0.00
177	B:VAL 192	13.33	0.00	0.00
178	B:ASN 193	101.40	0.00	0.00
179	B:GLU 194	51.14	0.00	0.00
180	B:TYR 195	1.40	0.00	0.00
181	B:ASN 196	39.38	0.00	0.00
182	B:THR 197	80.12	0.00	0.00
183	B:ARG 198	51.98	0.00	0.00
184	B:LEU 199	0.00	0.00	0.00
185	B:MET 200	43.92	0.00	0.00
186	B:VAL 201	69.48	0.00	0.00
187	B:LEU 202	7.66	0.00	0.00
188	B:TYR 203	11.38	0.00	0.00
189	B:SER 204	59.62	0.00	0.00
190	B:LYS 205	117.18	0.00	0.00
191	B:GLU 206	5.33	0.00	0.00
192	B:PHE 207	10.92	0.00	0.00
193	B:GLY 208	16.03	0.00	0.00
194	B:ARG 209	116.87	0.00	0.00
195	B:LEU 210	1.85	0.00	0.00
196	B:LEU 211	0.00	0.00	0.00
197	B:ALA 212	47.98	0.00	0.00
198	B:LYS 213	107.47	0.00	0.00
199	B:ASN 214	33.99	0.00	0.00
200	B:LEU 215	0.00	0.00	0.00
201	B:ASN 216	65.12	0.00	0.00
202	B:GLU 217	53.62	0.00	0.00
203	B:ALA 218	0.00	0.00	0.00



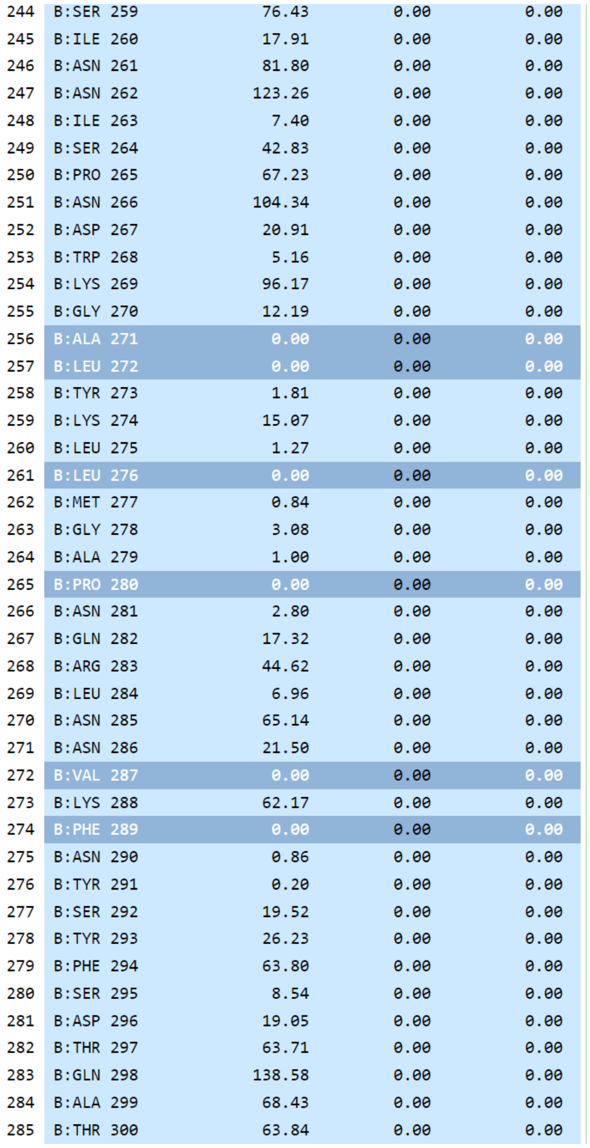
203	B:ALA 218	0.00	0.00	0.00
204	B:LEU 219	0.82	0.00	0.00
205	B:ASN 220	33.77	0.00	0.00
206	B:PHE 221	0.00	0.00	0.00
207	B:ARG 222	0.00	0.00	0.00
208	B:ASN 223	11.93	0.00	0.00
209	B:MET 224	0.00	0.00	0.00
210	B:CYS 225	0.00	0.00	0.00
211	B:SER 226	0.67	0.00	0.00
212	B:LEU 227	28.77	0.00	0.00
213	B:TYR 228	2.87	0.00	0.00
214	B:VAL 229	0.00	0.00	0.00
215	B:PHE 230	0.00	0.00	0.00
216	B:PRO 231	13.29	0.00	0.00
217	B:PHE 232	1.10	0.00	0.00
218	B:SER 233	3.79	0.00	0.00
219	B:GLU 234	1.91	0.00	0.00
220	B:ALA 235	1.16	0.00	0.00
221	B:TRP 236	4.65	0.00	0.00
222	B:SER 237	37.89	0.00	0.00
223	B:LEU 238	0.62	0.00	0.00
224	B:LEU 239	5.85	0.00	0.00
225	B:ARG 240	25.48	0.00	0.00
226	B:TYR 241	10.32	0.00	0.00
227	B:GLU 242	12.05	0.00	0.00
228	B:GLY 243	2.07	0.00	0.00
229	B:THR 244	4.51	0.00	0.00
230	B:LYS 245	77.68	0.00	0.00
231	B:LEU 246	26.05	0.00	0.00
232	B:GLU 247	32.13	0.00	0.00
233	B:ASN 248	31.25	0.00	0.00
234	B:THR 249	7.92	0.00	0.00
235	B:LEU 250	0.86	0.00	0.00
236	B:SER 251	2.03	0.00	0.00
237	B:LEU 252	1.17	0.00	0.00
238	B:TRP 253	3.89	0.00	0.00
239	B:ASN 254	1.76	0.00	0.00
240	B:PHE 255	17.17	0.00	0.00
241	B:VAL 256	8.36	0.00	0.00
242	B:GLY 257	21.98	0.00	0.00
243	B:GLU 258	52.33	0.00	0.00



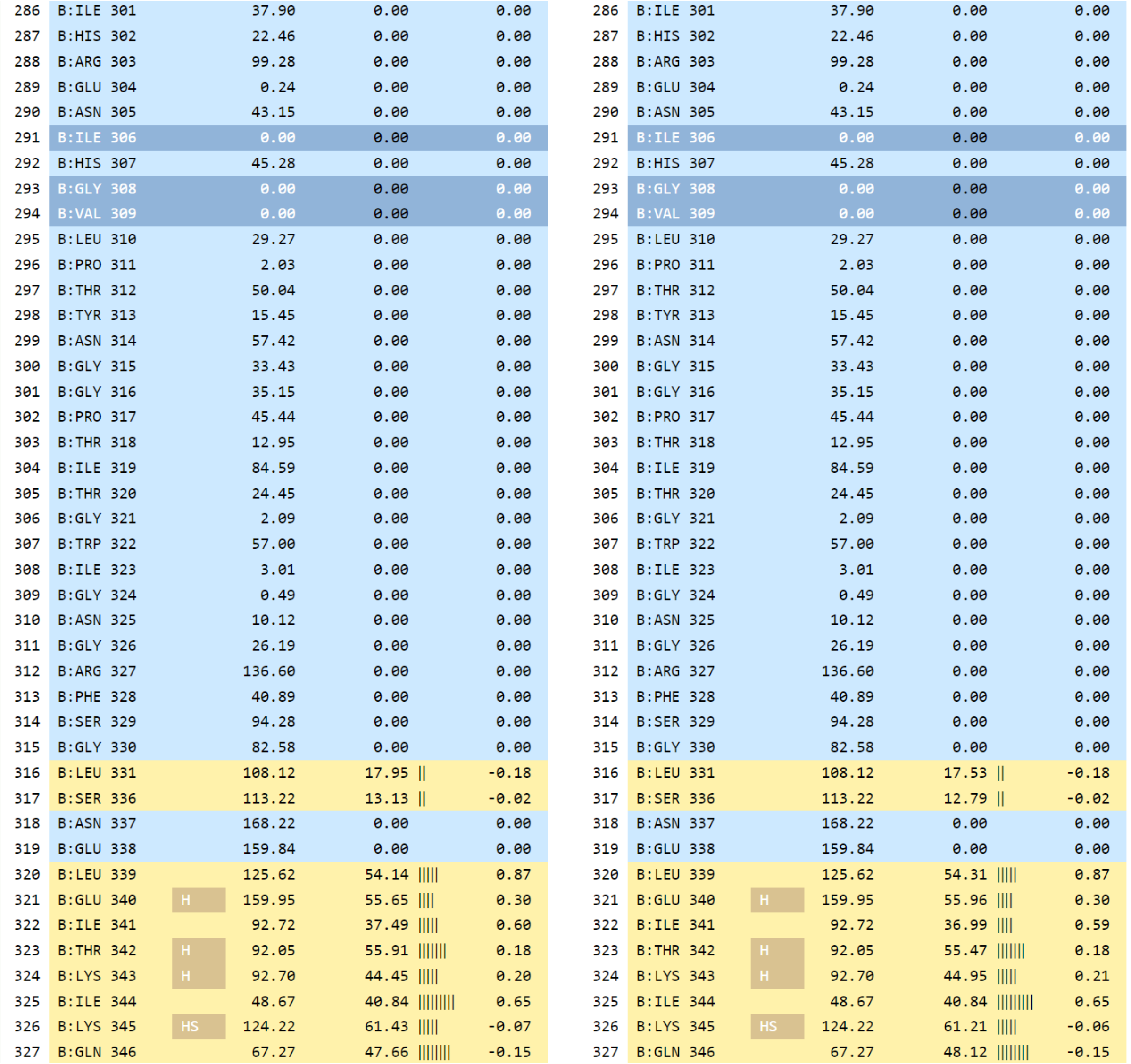
203	B:ALA 218	0.00	0.00	0.00
204	B:LEU 219	0.82	0.00	0.00
205	B:ASN 220	33.77	0.00	0.00
206	B:PHE 221	0.00	0.00	0.00
207	B:ARG 222	0.00	0.00	0.00
208	B:ASN 223	11.93	0.00	0.00
209	B:MET 224	0.00	0.00	0.00
210	B:CYS 225	0.00	0.00	0.00
211	B:SER 226	0.67	0.00	0.00
212	B:LEU 227	28.77	0.00	0.00
213	B:TYR 228	2.87	0.00	0.00
214	B:VAL 229	0.00	0.00	0.00
215	B:PHE 230	0.00	0.00	0.00
216	B:PRO 231	13.29	0.00	0.00
217	B:PHE 232	1.10	0.00	0.00
218	B:SER 233	3.79	0.00	0.00
219	B:GLU 234	1.91	0.00	0.00
220	B:ALA 235	1.16	0.00	0.00
221	B:TRP 236	4.65	0.00	0.00
222	B:SER 237	37.89	0.00	0.00
223	B:LEU 238	0.62	0.00	0.00
224	B:LEU 239	5.85	0.00	0.00
225	B:ARG 240	25.48	0.00	0.00
226	B:TYR 241	10.32	0.00	0.00
227	B:GLU 242	12.05	0.00	0.00
228	B:GLY 243	2.07	0.00	0.00
229	B:THR 244	4.51	0.00	0.00
230	B:LYS 245	77.68	0.00	0.00
231	B:LEU 246	26.05	0.00	0.00
232	B:GLU 247	32.13	0.00	0.00
233	B:ASN 248	31.25	0.00	0.00
234	B:THR 249	7.92	0.00	0.00
235	B:LEU 250	0.86	0.00	0.00
236	B:SER 251	2.03	0.00	0.00
237	B:LEU 252	1.17	0.00	0.00
238	B:TRP 253	3.89	0.00	0.00
239	B:ASN 254	1.76	0.00	0.00
240	B:PHE 255	17.17	0.00	0.00
241	B:VAL 256	8.36	0.00	0.00
242	B:GLY 257	21.98	0.00	0.00
243	B:GLU 258	52.33	0.00	0.00



244	B:SER 259	76.43	0.00	0.00
245	B:ILE 260	17.91	0.00	0.00
246	B:ASN 261	81.80	0.00	0.00
247	B:ASN 262	123.26	0.00	0.00
248	B:ILE 263	7.40	0.00	0.00
249	B:SER 264	42.83	0.00	0.00
250	B:PRO 265	67.23	0.00	0.00
251	B:ASN 266	104.34	0.00	0.00
252	B:ASP 267	20.91	0.00	0.00
253	B:TRP 268	5.16	0.00	0.00
254	B:LYS 269	96.17	0.00	0.00
255	B:GLY 270	12.19	0.00	0.00
256	B:ALA 271	0.00	0.00	0.00
257	B:LEU 272	0.00	0.00	0.00
258	B:TYR 273	1.81	0.00	0.00
259	B:LYS 274	15.07	0.00	0.00
260	B:LEU 275	1.27	0.00	0.00
261	B:LEU 276	0.00	0.00	0.00
262	B:MET 277	0.84	0.00	0.00
263	B:GLY 278	3.08	0.00	0.00
264	B:ALA 279	1.00	0.00	0.00
265	B:PRO 280	0.00	0.00	0.00
266	B:ASN 281	2.80	0.00	0.00
267	B:GLN 282	17.32	0.00	0.00
268	B:ARG 283	44.62	0.00	0.00
269	B:LEU 284	6.96	0.00	0.00
270	B:ASN 285	65.14	0.00	0.00
271	B:ASN 286	21.50	0.00	0.00
272	B:VAL 287	0.00	0.00	0.00
273	B:LYS 288	62.17	0.00	0.00
274	B:PHE 289	0.00	0.00	0.00
275	B:ASN 290	0.86	0.00	0.00
276	B:TYR 291	0.20	0.00	0.00
277	B:SER 292	19.52	0.00	0.00
278	B:TYR 293	26.23	0.00	0.00
279	B:PHE 294	63.80	0.00	0.00
280	B:SER 295	8.54	0.00	0.00
281	B:ASP 296	19.05	0.00	0.00
282	B:THR 297	63.71	0.00	0.00
283	B:GLN 298	138.58	0.00	0.00
284	B:ALA 299	68.43	0.00	0.00
285	B:THR 300	63.84	0.00	0.00

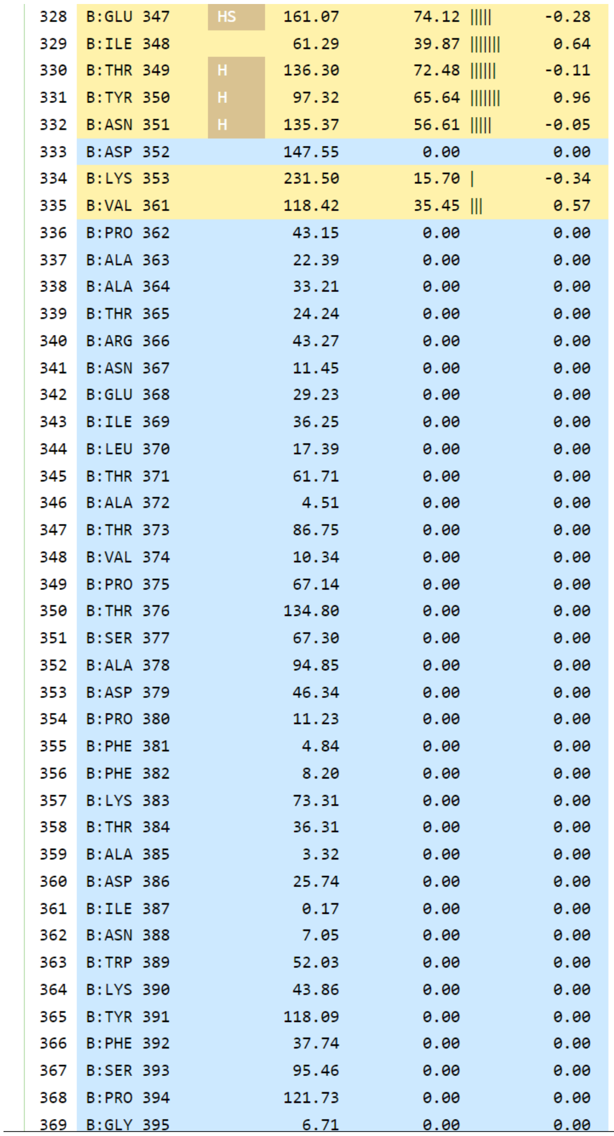


244	B:SER 259	76.43	0.00	0.00
245	B:ILE 260	17.91	0.00	0.00
246	B:ASN 261	81.80	0.00	0.00
247	B:ASN 262	123.26	0.00	0.00
248	B:ILE 263	7.40	0.00	0.00
249	B:SER 264	42.83	0.00	0.00
250	B:PRO 265	67.23	0.00	0.00
251	B:ASN 266	104.34	0.00	0.00
252	B:ASP 267	20.91	0.00	0.00
253	B:TRP 268	5.16	0.00	0.00
254	B:LYS 269	96.17	0.00	0.00
255	B:GLY 270	12.19	0.00	0.00
256	B:ALA 271	0.00	0.00	0.00
257	B:LEU 272	0.00	0.00	0.00
258	B:TYR 273	1.81	0.00	0.00
259	B:LYS 274	15.07	0.00	0.00
260	B:LEU 275	1.27	0.00	0.00
261	B:LEU 276	0.00	0.00	0.00
262	B:MET 277	0.84	0.00	0.00
263	B:GLY 278	3.08	0.00	0.00
264	B:ALA 279	1.00	0.00	0.00
265	B:PRO 280	0.00	0.00	0.00
266	B:ASN 281	2.80	0.00	0.00
267	B:GLN 282	17.32	0.00	0.00
268	B:ARG 283	44.62	0.00	0.00
269	B:LEU 284	6.96	0.00	0.00
270	B:ASN 285	65.14	0.00	0.00
271	B:ASN 286	21.50	0.00	0.00
272	B:VAL 287	0.00	0.00	0.00
273	B:LYS 288	62.17	0.00	0.00
274	B:PHE 289	0.00	0.00	0.00
275	B:ASN 290	0.86	0.00	0.00
276	B:TYR 291	0.20	0.00	0.00
277	B:SER 292	19.52	0.00	0.00
278	B:TYR 293	26.23	0.00	0.00
279	B:PHE 294	63.80	0.00	0.00
280	B:SER 295	8.54	0.00	0.00
281	B:ASP 296	19.05	0.00	0.00
282	B:THR 297	63.71	0.00	0.00
283	B:GLN 298	138.58	0.00	0.00
284	B:ALA 299	68.43	0.00	0.00
285	B:THR 300	63.84	0.00	0.00

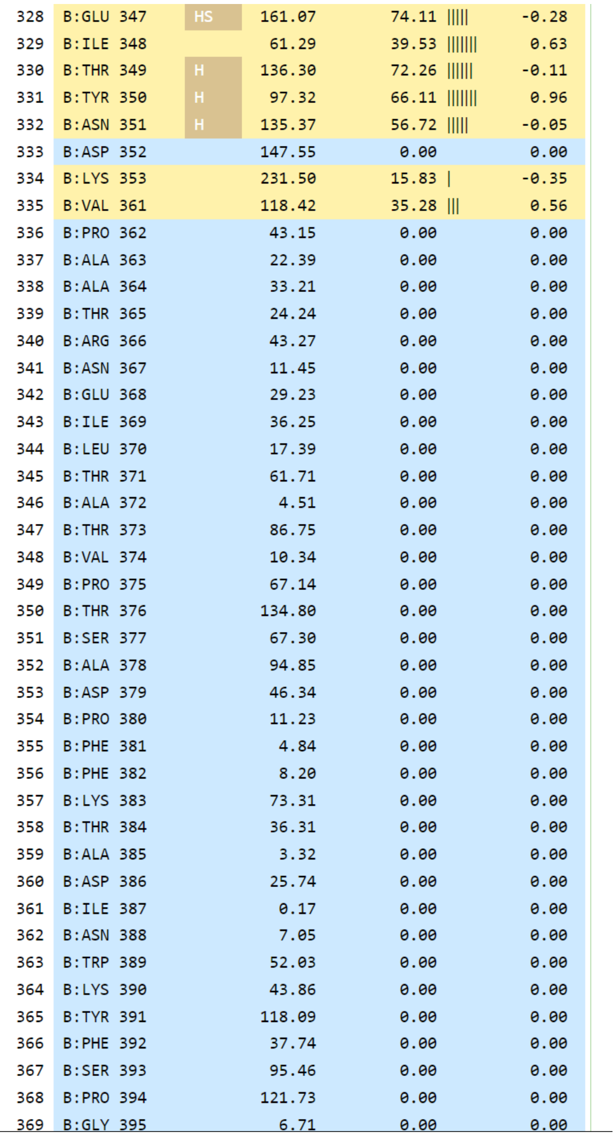


<table>
<tr><td>286</td><td>B:ILE 301</td><td></td><td>37.90</td><td>0.00</td><td>0.00</td></tr>
<tr><td>287</td><td>B:HIS 302</td><td></td><td>22.46</td><td>0.00</td><td>0.00</td></tr>
<tr><td>288</td><td>B:ARG 303</td><td></td><td>99.28</td><td>0.00</td><td>0.00</td></tr>
<tr><td>289</td><td>B:GLU 304</td><td></td><td>0.24</td><td>0.00</td><td>0.00</td></tr>
<tr><td>290</td><td>B:ASN 305</td><td></td><td>43.15</td><td>0.00</td><td>0.00</td></tr>
<tr><td>291</td><td>B:ILE 306</td><td></td><td>0.00</td><td>0.00</td><td>0.00</td></tr>
<tr><td>292</td><td>B:HIS 307</td><td></td><td>45.28</td><td>0.00</td><td>0.00</td></tr>
<tr><td>293</td><td>B:GLY 308</td><td></td><td>0.00</td><td>0.00</td><td>0.00</td></tr>
<tr><td>294</td><td>B:VAL 309</td><td></td><td>0.00</td><td>0.00</td><td>0.00</td></tr>
<tr><td>295</td><td>B:LEU 310</td><td></td><td>29.27</td><td>0.00</td><td>0.00</td></tr>
<tr><td>296</td><td>B:PRO 311</td><td></td><td>2.03</td><td>0.00</td><td>0.00</td></tr>
<tr><td>297</td><td>B:THR 312</td><td></td><td>50.04</td><td>0.00</td><td>0.00</td></tr>
<tr><td>298</td><td>B:TYR 313</td><td></td><td>15.45</td><td>0.00</td><td>0.00</td></tr>
<tr><td>299</td><td>B:ASN 314</td><td></td><td>57.42</td><td>0.00</td><td>0.00</td></tr>
<tr><td>300</td><td>B:GLY 315</td><td></td><td>33.43</td><td>0.00</td><td>0.00</td></tr>
<tr><td>301</td><td>B:GLY 316</td><td></td><td>35.15</td><td>0.00</td><td>0.00</td></tr>
<tr><td>302</td><td>B:PRO 317</td><td></td><td>45.44</td><td>0.00</td><td>0.00</td></tr>
<tr><td>303</td><td>B:THR 318</td><td></td><td>12.95</td><td>0.00</td><td>0.00</td></tr>
<tr><td>304</td><td>B:ILE 319</td><td></td><td>84.59</td><td>0.00</td><td>0.00</td></tr>
<tr><td>305</td><td>B:THR 320</td><td></td><td>24.45</td><td>0.00</td><td>0.00</td></tr>
<tr><td>306</td><td>B:GLY 321</td><td></td><td>2.09</td><td>0.00</td><td>0.00</td></tr>
<tr><td>307</td><td>B:TRP 322</td><td></td><td>57.00</td><td>0.00</td><td>0.00</td></tr>
<tr><td>308</td><td>B:ILE 323</td><td></td><td>3.01</td><td>0.00</td><td>0.00</td></tr>
<tr><td>309</td><td>B:GLY 324</td><td></td><td>0.49</td><td>0.00</td><td>0.00</td></tr>
<tr><td>310</td><td>B:ASN 325</td><td></td><td>10.12</td><td>0.00</td><td>0.00</td></tr>
<tr><td>311</td><td>B:GLY 326</td><td></td><td>26.19</td><td>0.00</td><td>0.00</td></tr>
<tr><td>312</td><td>B:ARG 327</td><td></td><td>136.60</td><td>0.00</td><td>0.00</td></tr>
<tr><td>313</td><td>B:PHE 328</td><td></td><td>40.89</td><td>0.00</td><td>0.00</td></tr>
<tr><td>314</td><td>B:SER 329</td><td></td><td>94.28</td><td>0.00</td><td>0.00</td></tr>
<tr><td>315</td><td>B:GLY 330</td><td></td><td>82.58</td><td>0.00</td><td>0.00</td></tr>
<tr><td>316</td><td>B:LEU 331</td><td></td><td>108.12</td><td>17.95 ||</td><td>-0.18</td></tr>
<tr><td>317</td><td>B:SER 336</td><td></td><td>113.22</td><td>13.13 ||</td><td>-0.02</td></tr>
<tr><td>318</td><td>B:ASN 337</td><td></td><td>168.22</td><td>0.00</td><td>0.00</td></tr>
<tr><td>319</td><td>B:GLU 338</td><td></td><td>159.84</td><td>0.00</td><td>0.00</td></tr>
<tr><td>320</td><td>B:LEU 339</td><td></td><td>125.62</td><td>54.14 |||||</td><td>0.87</td></tr>
<tr><td>321</td><td>B:GLU 340</td><td>H</td><td>159.95</td><td>55.65 ||||</td><td>0.30</td></tr>
<tr><td>322</td><td>B:ILE 341</td><td></td><td>92.72</td><td>37.49 |||||</td><td>0.60</td></tr>
<tr><td>323</td><td>B:THR 342</td><td>H</td><td>92.05</td><td>55.91 |||||||</td><td>0.18</td></tr>
<tr><td>324</td><td>B:LYS 343</td><td>H</td><td>92.70</td><td>44.45 |||||</td><td>0.20</td></tr>
<tr><td>325</td><td>B:ILE 344</td><td></td><td>48.67</td><td>40.84 |||||||||</td><td>0.65</td></tr>
<tr><td>326</td><td>B:LYS 345</td><td>HS</td><td>124.22</td><td>61.43 |||||</td><td>-0.07</td></tr>
<tr><td>327</td><td>B:GLN 346</td><td></td><td>67.27</td><td>47.66 ||||||||</td><td>-0.15</td></tr>
</table>

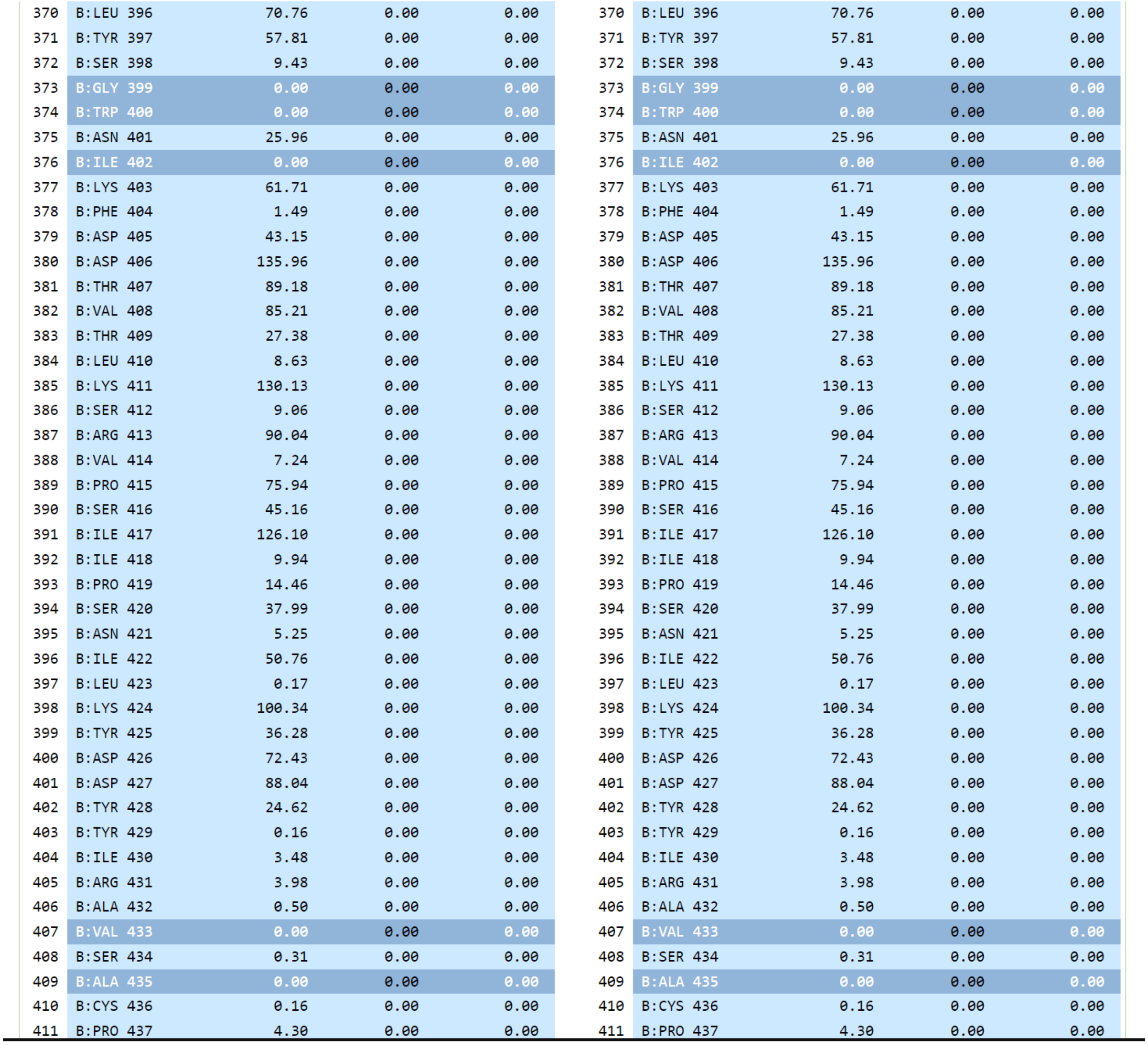
<table>
<tr><td>286</td><td>B:ILE 301</td><td></td><td>37.90</td><td>0.00</td><td>0.00</td></tr>
<tr><td>287</td><td>B:HIS 302</td><td></td><td>22.46</td><td>0.00</td><td>0.00</td></tr>
<tr><td>288</td><td>B:ARG 303</td><td></td><td>99.28</td><td>0.00</td><td>0.00</td></tr>
<tr><td>289</td><td>B:GLU 304</td><td></td><td>0.24</td><td>0.00</td><td>0.00</td></tr>
<tr><td>290</td><td>B:ASN 305</td><td></td><td>43.15</td><td>0.00</td><td>0.00</td></tr>
<tr><td>291</td><td>B:ILE 306</td><td></td><td>0.00</td><td>0.00</td><td>0.00</td></tr>
<tr><td>292</td><td>B:HIS 307</td><td></td><td>45.28</td><td>0.00</td><td>0.00</td></tr>
<tr><td>293</td><td>B:GLY 308</td><td></td><td>0.00</td><td>0.00</td><td>0.00</td></tr>
<tr><td>294</td><td>B:VAL 309</td><td></td><td>0.00</td><td>0.00</td><td>0.00</td></tr>
<tr><td>295</td><td>B:LEU 310</td><td></td><td>29.27</td><td>0.00</td><td>0.00</td></tr>
<tr><td>296</td><td>B:PRO 311</td><td></td><td>2.03</td><td>0.00</td><td>0.00</td></tr>
<tr><td>297</td><td>B:THR 312</td><td></td><td>50.04</td><td>0.00</td><td>0.00</td></tr>
<tr><td>298</td><td>B:TYR 313</td><td></td><td>15.45</td><td>0.00</td><td>0.00</td></tr>
<tr><td>299</td><td>B:ASN 314</td><td></td><td>57.42</td><td>0.00</td><td>0.00</td></tr>
<tr><td>300</td><td>B:GLY 315</td><td></td><td>33.43</td><td>0.00</td><td>0.00</td></tr>
<tr><td>301</td><td>B:GLY 316</td><td></td><td>35.15</td><td>0.00</td><td>0.00</td></tr>
<tr><td>302</td><td>B:PRO 317</td><td></td><td>45.44</td><td>0.00</td><td>0.00</td></tr>
<tr><td>303</td><td>B:THR 318</td><td></td><td>12.95</td><td>0.00</td><td>0.00</td></tr>
<tr><td>304</td><td>B:ILE 319</td><td></td><td>84.59</td><td>0.00</td><td>0.00</td></tr>
<tr><td>305</td><td>B:THR 320</td><td></td><td>24.45</td><td>0.00</td><td>0.00</td></tr>
<tr><td>306</td><td>B:GLY 321</td><td></td><td>2.09</td><td>0.00</td><td>0.00</td></tr>
<tr><td>307</td><td>B:TRP 322</td><td></td><td>57.00</td><td>0.00</td><td>0.00</td></tr>
<tr><td>308</td><td>B:ILE 323</td><td></td><td>3.01</td><td>0.00</td><td>0.00</td></tr>
<tr><td>309</td><td>B:GLY 324</td><td></td><td>0.49</td><td>0.00</td><td>0.00</td></tr>
<tr><td>310</td><td>B:ASN 325</td><td></td><td>10.12</td><td>0.00</td><td>0.00</td></tr>
<tr><td>311</td><td>B:GLY 326</td><td></td><td>26.19</td><td>0.00</td><td>0.00</td></tr>
<tr><td>312</td><td>B:ARG 327</td><td></td><td>136.60</td><td>0.00</td><td>0.00</td></tr>
<tr><td>313</td><td>B:PHE 328</td><td></td><td>40.89</td><td>0.00</td><td>0.00</td></tr>
<tr><td>314</td><td>B:SER 329</td><td></td><td>94.28</td><td>0.00</td><td>0.00</td></tr>
<tr><td>315</td><td>B:GLY 330</td><td></td><td>82.58</td><td>0.00</td><td>0.00</td></tr>
<tr><td>316</td><td>B:LEU 331</td><td></td><td>108.12</td><td>17.53 ||</td><td>-0.18</td></tr>
<tr><td>317</td><td>B:SER 336</td><td></td><td>113.22</td><td>12.79 ||</td><td>-0.02</td></tr>
<tr><td>318</td><td>B:ASN 337</td><td></td><td>168.22</td><td>0.00</td><td>0.00</td></tr>
<tr><td>319</td><td>B:GLU 338</td><td></td><td>159.84</td><td>0.00</td><td>0.00</td></tr>
<tr><td>320</td><td>B:LEU 339</td><td></td><td>125.62</td><td>54.31 |||||</td><td>0.87</td></tr>
<tr><td>321</td><td>B:GLU 340</td><td>H</td><td>159.95</td><td>55.96 ||||</td><td>0.30</td></tr>
<tr><td>322</td><td>B:ILE 341</td><td></td><td>92.72</td><td>36.99 ||||</td><td>0.59</td></tr>
<tr><td>323</td><td>B:THR 342</td><td>H</td><td>92.05</td><td>55.47 |||||||</td><td>0.18</td></tr>
<tr><td>324</td><td>B:LYS 343</td><td>H</td><td>92.70</td><td>44.95 |||||</td><td>0.21</td></tr>
<tr><td>325</td><td>B:ILE 344</td><td></td><td>48.67</td><td>40.84 |||||||||</td><td>0.65</td></tr>
<tr><td>326</td><td>B:LYS 345</td><td>HS</td><td>124.22</td><td>61.21 |||||</td><td>-0.06</td></tr>
<tr><td>327</td><td>B:GLN 346</td><td></td><td>67.27</td><td>48.12 ||||||||</td><td>-0.15</td></tr>
</table>



328	B:GLU 347	HS	161.07	74.12	|||||	-0.28
329	B:ILE 348		61.29	39.87	|||||||	0.64
330	B:THR 349	H	136.30	72.48	||||||	-0.11
331	B:TYR 350	H	97.32	65.64	|||||||	0.96
332	B:ASN 351	H	135.37	56.61	|||||	-0.05
333	B:ASP 352		147.55	0.00		0.00
334	B:LYS 353		231.50	15.70	|	-0.34
335	B:VAL 361		118.42	35.45	|||	0.57
336	B:PRO 362		43.15	0.00		0.00
337	B:ALA 363		22.39	0.00		0.00
338	B:ALA 364		33.21	0.00		0.00
339	B:THR 365		24.24	0.00		0.00
340	B:ARG 366		43.27	0.00		0.00
341	B:ASN 367		11.45	0.00		0.00
342	B:GLU 368		29.23	0.00		0.00
343	B:ILE 369		36.25	0.00		0.00
344	B:LEU 370		17.39	0.00		0.00
345	B:THR 371		61.71	0.00		0.00
346	B:ALA 372		4.51	0.00		0.00
347	B:THR 373		86.75	0.00		0.00
348	B:VAL 374		10.34	0.00		0.00
349	B:PRO 375		67.14	0.00		0.00
350	B:THR 376		134.80	0.00		0.00
351	B:SER 377		67.30	0.00		0.00
352	B:ALA 378		94.85	0.00		0.00
353	B:ASP 379		46.34	0.00		0.00
354	B:PRO 380		11.23	0.00		0.00
355	B:PHE 381		4.84	0.00		0.00
356	B:PHE 382		8.20	0.00		0.00
357	B:LYS 383		73.31	0.00		0.00
358	B:THR 384		36.31	0.00		0.00
359	B:ALA 385		3.32	0.00		0.00
360	B:ASP 386		25.74	0.00		0.00
361	B:ILE 387		0.17	0.00		0.00
362	B:ASN 388		7.05	0.00		0.00
363	B:TRP 389		52.03	0.00		0.00
364	B:LYS 390		43.86	0.00		0.00
365	B:TYR 391		118.09	0.00		0.00
366	B:PHE 392		37.74	0.00		0.00
367	B:SER 393		95.46	0.00		0.00
368	B:PRO 394		121.73	0.00		0.00
369	B:GLY 395		6.71	0.00		0.00

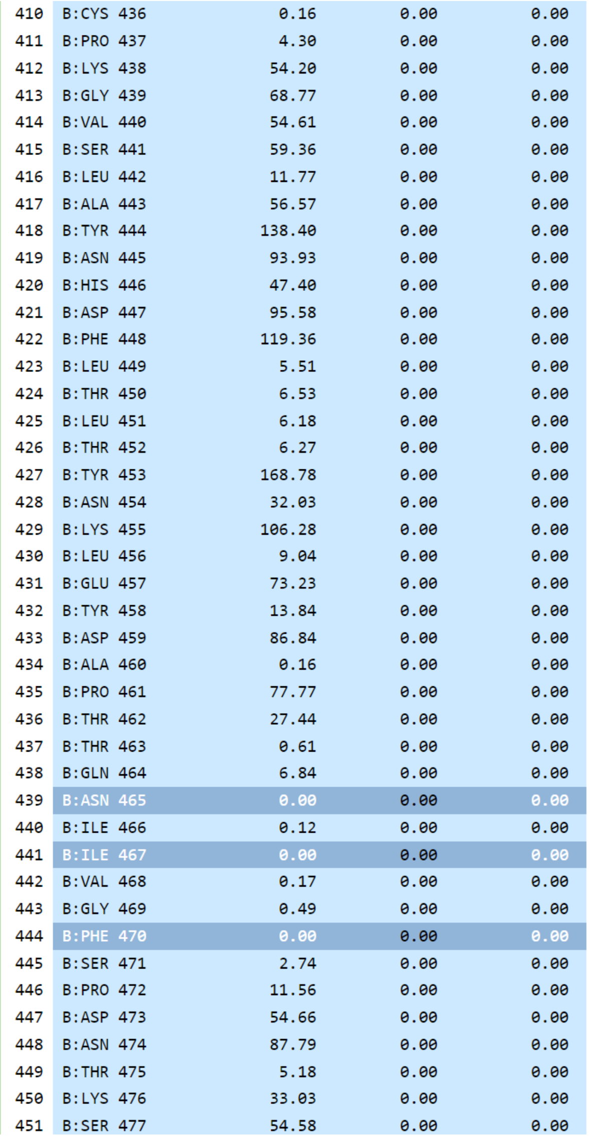


328	B:GLU 347	HS	161.07	74.11	|||||	-0.28
329	B:ILE 348		61.29	39.53	|||||||	0.63
330	B:THR 349	H	136.30	72.26	||||||	-0.11
331	B:TYR 350	H	97.32	66.11	|||||||	0.96
332	B:ASN 351	H	135.37	56.72	|||||	-0.05
333	B:ASP 352		147.55	0.00		0.00
334	B:LYS 353		231.50	15.83	|	-0.35
335	B:VAL 361		118.42	35.28	|||	0.56
336	B:PRO 362		43.15	0.00		0.00
337	B:ALA 363		22.39	0.00		0.00
338	B:ALA 364		33.21	0.00		0.00
339	B:THR 365		24.24	0.00		0.00
340	B:ARG 366		43.27	0.00		0.00
341	B:ASN 367		11.45	0.00		0.00
342	B:GLU 368		29.23	0.00		0.00
343	B:ILE 369		36.25	0.00		0.00
344	B:LEU 370		17.39	0.00		0.00
345	B:THR 371		61.71	0.00		0.00
346	B:ALA 372		4.51	0.00		0.00
347	B:THR 373		86.75	0.00		0.00
348	B:VAL 374		10.34	0.00		0.00
349	B:PRO 375		67.14	0.00		0.00
350	B:THR 376		134.80	0.00		0.00
351	B:SER 377		67.30	0.00		0.00
352	B:ALA 378		94.85	0.00		0.00
353	B:ASP 379		46.34	0.00		0.00
354	B:PRO 380		11.23	0.00		0.00
355	B:PHE 381		4.84	0.00		0.00
356	B:PHE 382		8.20	0.00		0.00
357	B:LYS 383		73.31	0.00		0.00
358	B:THR 384		36.31	0.00		0.00
359	B:ALA 385		3.32	0.00		0.00
360	B:ASP 386		25.74	0.00		0.00
361	B:ILE 387		0.17	0.00		0.00
362	B:ASN 388		7.05	0.00		0.00
363	B:TRP 389		52.03	0.00		0.00
364	B:LYS 390		43.86	0.00		0.00
365	B:TYR 391		118.09	0.00		0.00
366	B:PHE 392		37.74	0.00		0.00
367	B:SER 393		95.46	0.00		0.00
368	B:PRO 394		121.73	0.00		0.00
369	B:GLY 395		6.71	0.00		0.00

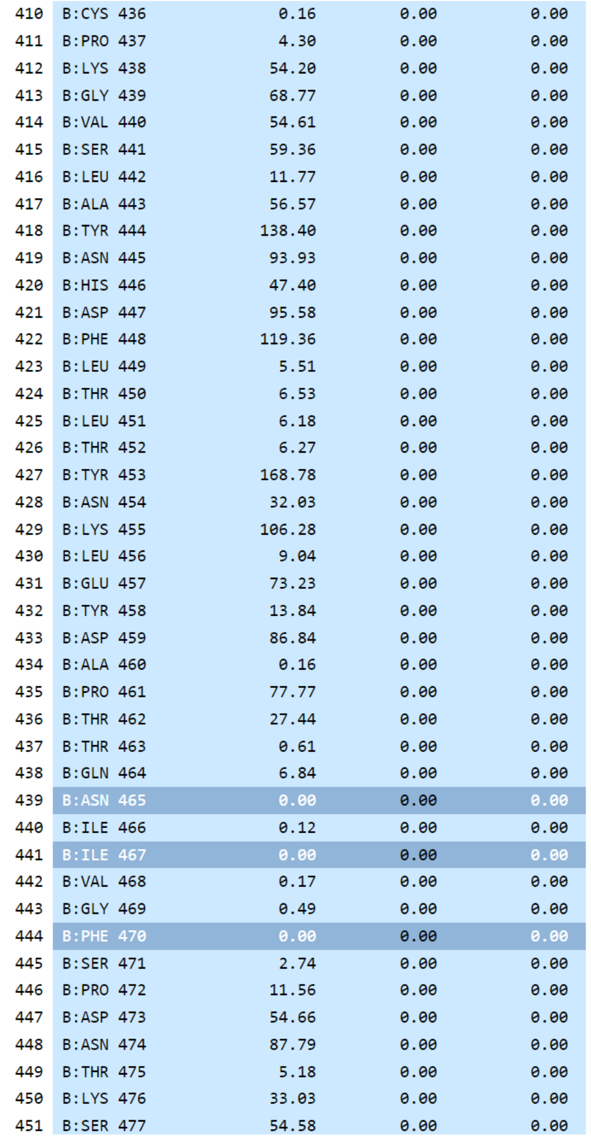


370	B:LEU 396	70.76	0.00	0.00
371	B:TYR 397	57.81	0.00	0.00
372	B:SER 398	9.43	0.00	0.00
373	B:GLY 399	0.00	0.00	0.00
374	B:TRP 400	0.00	0.00	0.00
375	B:ASN 401	25.96	0.00	0.00
376	B:ILE 402	0.00	0.00	0.00
377	B:LYS 403	61.71	0.00	0.00
378	B:PHE 404	1.49	0.00	0.00
379	B:ASP 405	43.15	0.00	0.00
380	B:ASP 406	135.96	0.00	0.00
381	B:THR 407	89.18	0.00	0.00
382	B:VAL 408	85.21	0.00	0.00
383	B:THR 409	27.38	0.00	0.00
384	B:LEU 410	8.63	0.00	0.00
385	B:LYS 411	130.13	0.00	0.00
386	B:SER 412	9.06	0.00	0.00
387	B:ARG 413	90.04	0.00	0.00
388	B:VAL 414	7.24	0.00	0.00
389	B:PRO 415	75.94	0.00	0.00
390	B:SER 416	45.16	0.00	0.00
391	B:ILE 417	126.10	0.00	0.00
392	B:ILE 418	9.94	0.00	0.00
393	B:PRO 419	14.46	0.00	0.00
394	B:SER 420	37.99	0.00	0.00
395	B:ASN 421	5.25	0.00	0.00
396	B:ILE 422	50.76	0.00	0.00
397	B:LEU 423	0.17	0.00	0.00
398	B:LYS 424	100.34	0.00	0.00
399	B:TYR 425	36.28	0.00	0.00
400	B:ASP 426	72.43	0.00	0.00
401	B:ASP 427	88.04	0.00	0.00
402	B:TYR 428	24.62	0.00	0.00
403	B:TYR 429	0.16	0.00	0.00
404	B:ILE 430	3.48	0.00	0.00
405	B:ARG 431	3.98	0.00	0.00
406	B:ALA 432	0.50	0.00	0.00
407	B:VAL 433	0.00	0.00	0.00
408	B:SER 434	0.31	0.00	0.00
409	B:ALA 435	0.00	0.00	0.00
410	B:CYS 436	0.16	0.00	0.00
411	B:PRO 437	4.30	0.00	0.00

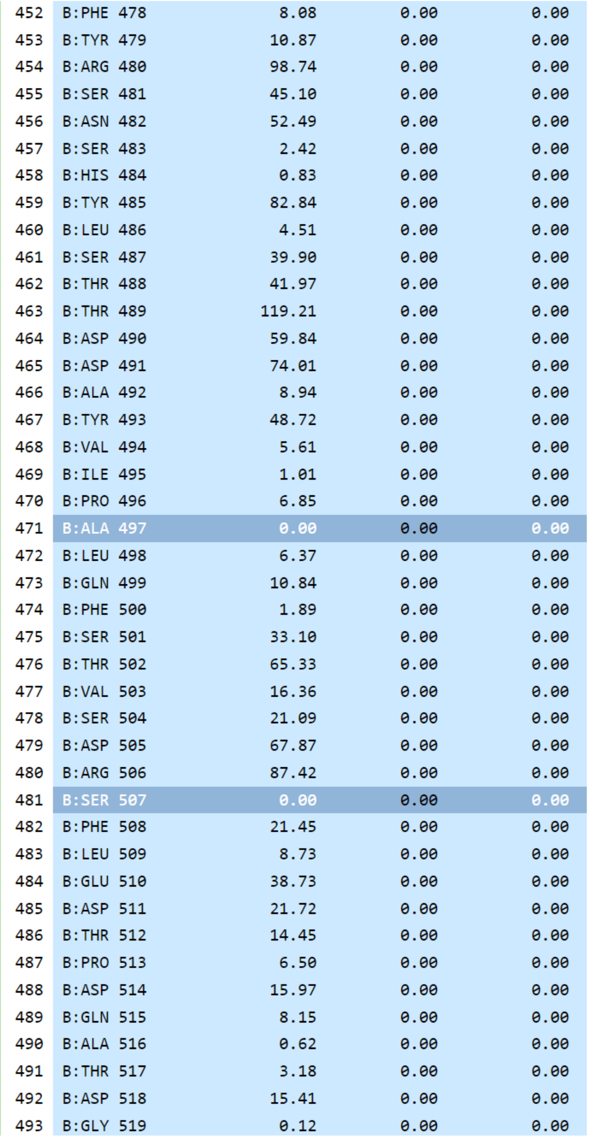
370	B:LEU 396	70.76	0.00	0.00
371	B:TYR 397	57.81	0.00	0.00
372	B:SER 398	9.43	0.00	0.00
373	B:GLY 399	0.00	0.00	0.00
374	B:TRP 400	0.00	0.00	0.00
375	B:ASN 401	25.96	0.00	0.00
376	B:ILE 402	0.00	0.00	0.00
377	B:LYS 403	61.71	0.00	0.00
378	B:PHE 404	1.49	0.00	0.00
379	B:ASP 405	43.15	0.00	0.00
380	B:ASP 406	135.96	0.00	0.00
381	B:THR 407	89.18	0.00	0.00
382	B:VAL 408	85.21	0.00	0.00
383	B:THR 409	27.38	0.00	0.00
384	B:LEU 410	8.63	0.00	0.00
385	B:LYS 411	130.13	0.00	0.00
386	B:SER 412	9.06	0.00	0.00
387	B:ARG 413	90.04	0.00	0.00
388	B:VAL 414	7.24	0.00	0.00
389	B:PRO 415	75.94	0.00	0.00
390	B:SER 416	45.16	0.00	0.00
391	B:ILE 417	126.10	0.00	0.00
392	B:ILE 418	9.94	0.00	0.00
393	B:PRO 419	14.46	0.00	0.00
394	B:SER 420	37.99	0.00	0.00
395	B:ASN 421	5.25	0.00	0.00
396	B:ILE 422	50.76	0.00	0.00
397	B:LEU 423	0.17	0.00	0.00
398	B:LYS 424	100.34	0.00	0.00
399	B:TYR 425	36.28	0.00	0.00
400	B:ASP 426	72.43	0.00	0.00
401	B:ASP 427	88.04	0.00	0.00
402	B:TYR 428	24.62	0.00	0.00
403	B:TYR 429	0.16	0.00	0.00
404	B:ILE 430	3.48	0.00	0.00
405	B:ARG 431	3.98	0.00	0.00
406	B:ALA 432	0.50	0.00	0.00
407	B:VAL 433	0.00	0.00	0.00
408	B:SER 434	0.31	0.00	0.00
409	B:ALA 435	0.00	0.00	0.00
410	B:CYS 436	0.16	0.00	0.00
411	B:PRO 437	4.30	0.00	0.00



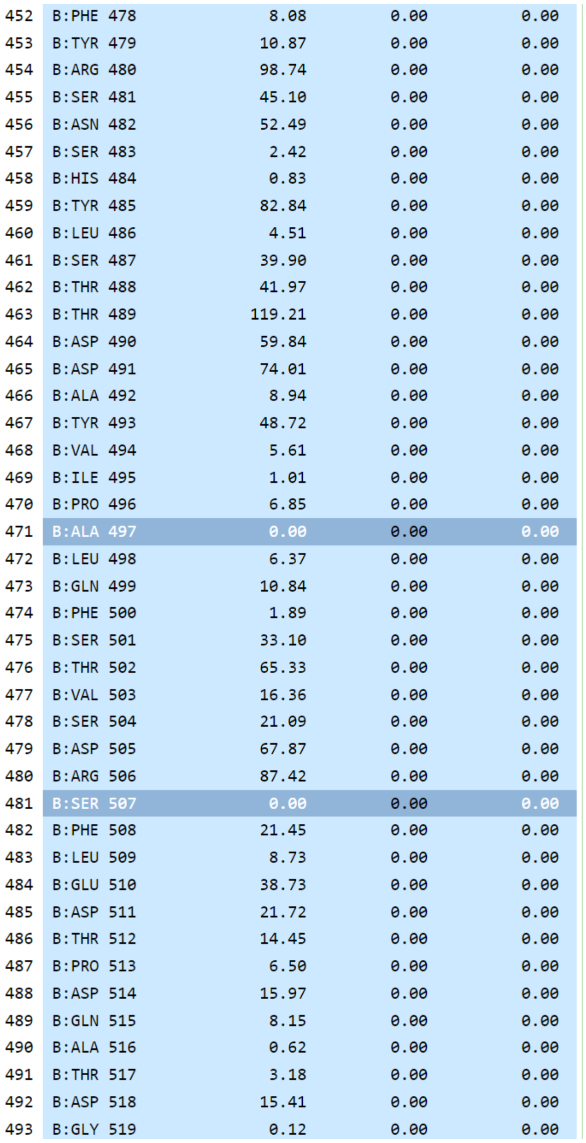
410	B:CYS 436	0.16	0.00	0.00
411	B:PRO 437	4.30	0.00	0.00
412	B:LYS 438	54.20	0.00	0.00
413	B:GLY 439	68.77	0.00	0.00
414	B:VAL 440	54.61	0.00	0.00
415	B:SER 441	59.36	0.00	0.00
416	B:LEU 442	11.77	0.00	0.00
417	B:ALA 443	56.57	0.00	0.00
418	B:TYR 444	138.40	0.00	0.00
419	B:ASN 445	93.93	0.00	0.00
420	B:HIS 446	47.40	0.00	0.00
421	B:ASP 447	95.58	0.00	0.00
422	B:PHE 448	119.36	0.00	0.00
423	B:LEU 449	5.51	0.00	0.00
424	B:THR 450	6.53	0.00	0.00
425	B:LEU 451	6.18	0.00	0.00
426	B:THR 452	6.27	0.00	0.00
427	B:TYR 453	168.78	0.00	0.00
428	B:ASN 454	32.03	0.00	0.00
429	B:LYS 455	106.28	0.00	0.00
430	B:LEU 456	9.04	0.00	0.00
431	B:GLU 457	73.23	0.00	0.00
432	B:TYR 458	13.84	0.00	0.00
433	B:ASP 459	86.84	0.00	0.00
434	B:ALA 460	0.16	0.00	0.00
435	B:PRO 461	77.77	0.00	0.00
436	B:THR 462	27.44	0.00	0.00
437	B:THR 463	0.61	0.00	0.00
438	B:GLN 464	6.84	0.00	0.00
439	B:ASN 465	0.00	0.00	0.00
440	B:ILE 466	0.12	0.00	0.00
441	B:ILE 467	0.00	0.00	0.00
442	B:VAL 468	0.17	0.00	0.00
443	B:GLY 469	0.49	0.00	0.00
444	B:PHE 470	0.00	0.00	0.00
445	B:SER 471	2.74	0.00	0.00
446	B:PRO 472	11.56	0.00	0.00
447	B:ASP 473	54.66	0.00	0.00
448	B:ASN 474	87.79	0.00	0.00
449	B:THR 475	5.18	0.00	0.00
450	B:LYS 476	33.03	0.00	0.00
451	B:SER 477	54.58	0.00	0.00



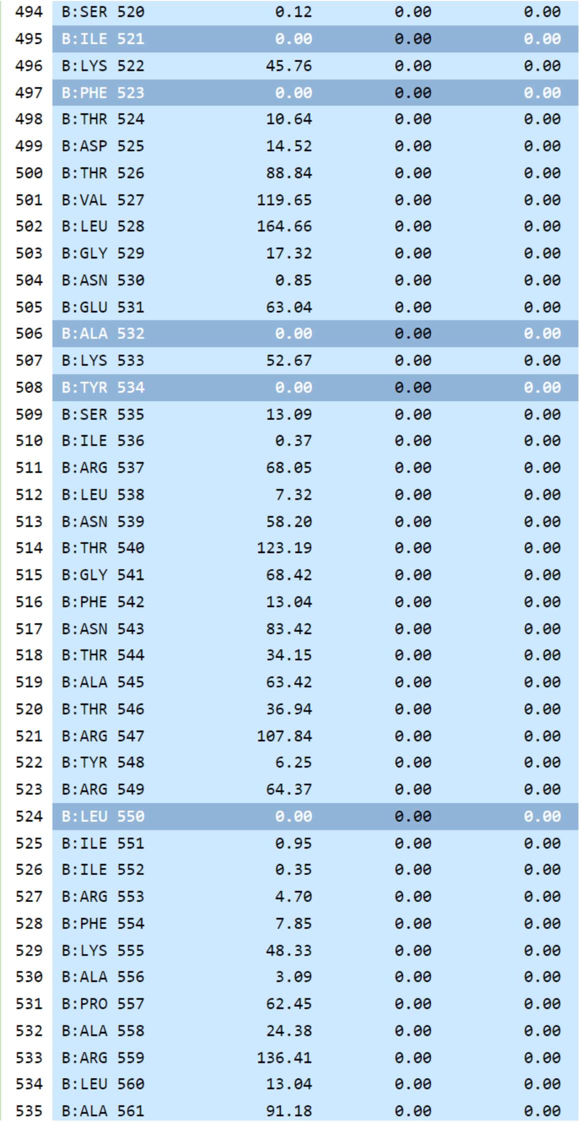
410	B:CYS 436	0.16	0.00	0.00
411	B:PRO 437	4.30	0.00	0.00
412	B:LYS 438	54.20	0.00	0.00
413	B:GLY 439	68.77	0.00	0.00
414	B:VAL 440	54.61	0.00	0.00
415	B:SER 441	59.36	0.00	0.00
416	B:LEU 442	11.77	0.00	0.00
417	B:ALA 443	56.57	0.00	0.00
418	B:TYR 444	138.40	0.00	0.00
419	B:ASN 445	93.93	0.00	0.00
420	B:HIS 446	47.40	0.00	0.00
421	B:ASP 447	95.58	0.00	0.00
422	B:PHE 448	119.36	0.00	0.00
423	B:LEU 449	5.51	0.00	0.00
424	B:THR 450	6.53	0.00	0.00
425	B:LEU 451	6.18	0.00	0.00
426	B:THR 452	6.27	0.00	0.00
427	B:TYR 453	168.78	0.00	0.00
428	B:ASN 454	32.03	0.00	0.00
429	B:LYS 455	106.28	0.00	0.00
430	B:LEU 456	9.04	0.00	0.00
431	B:GLU 457	73.23	0.00	0.00
432	B:TYR 458	13.84	0.00	0.00
433	B:ASP 459	86.84	0.00	0.00
434	B:ALA 460	0.16	0.00	0.00
435	B:PRO 461	77.77	0.00	0.00
436	B:THR 462	27.44	0.00	0.00
437	B:THR 463	0.61	0.00	0.00
438	B:GLN 464	6.84	0.00	0.00
439	B:ASN 465	0.00	0.00	0.00
440	B:ILE 466	0.12	0.00	0.00
441	B:ILE 467	0.00	0.00	0.00
442	B:VAL 468	0.17	0.00	0.00
443	B:GLY 469	0.49	0.00	0.00
444	B:PHE 470	0.00	0.00	0.00
445	B:SER 471	2.74	0.00	0.00
446	B:PRO 472	11.56	0.00	0.00
447	B:ASP 473	54.66	0.00	0.00
448	B:ASN 474	87.79	0.00	0.00
449	B:THR 475	5.18	0.00	0.00
450	B:LYS 476	33.03	0.00	0.00
451	B:SER 477	54.58	0.00	0.00



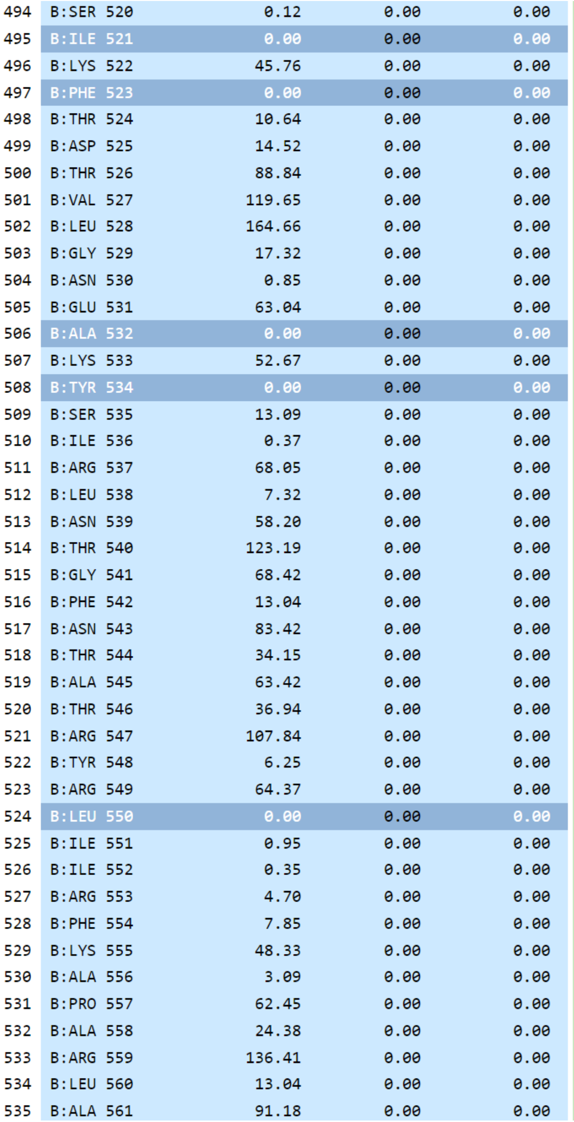
452	B:PHE 478	8.08	0.00	0.00
453	B:TYR 479	10.87	0.00	0.00
454	B:ARG 480	98.74	0.00	0.00
455	B:SER 481	45.10	0.00	0.00
456	B:ASN 482	52.49	0.00	0.00
457	B:SER 483	2.42	0.00	0.00
458	B:HIS 484	0.83	0.00	0.00
459	B:TYR 485	82.84	0.00	0.00
460	B:LEU 486	4.51	0.00	0.00
461	B:SER 487	39.90	0.00	0.00
462	B:THR 488	41.97	0.00	0.00
463	B:THR 489	119.21	0.00	0.00
464	B:ASP 490	59.84	0.00	0.00
465	B:ASP 491	74.01	0.00	0.00
466	B:ALA 492	8.94	0.00	0.00
467	B:TYR 493	48.72	0.00	0.00
468	B:VAL 494	5.61	0.00	0.00
469	B:ILE 495	1.01	0.00	0.00
470	B:PRO 496	6.85	0.00	0.00
471	B:ALA 497	0.00	0.00	0.00
472	B:LEU 498	6.37	0.00	0.00
473	B:GLN 499	10.84	0.00	0.00
474	B:PHE 500	1.89	0.00	0.00
475	B:SER 501	33.10	0.00	0.00
476	B:THR 502	65.33	0.00	0.00
477	B:VAL 503	16.36	0.00	0.00
478	B:SER 504	21.09	0.00	0.00
479	B:ASP 505	67.87	0.00	0.00
480	B:ARG 506	87.42	0.00	0.00
481	B:SER 507	0.00	0.00	0.00
482	B:PHE 508	21.45	0.00	0.00
483	B:LEU 509	8.73	0.00	0.00
484	B:GLU 510	38.73	0.00	0.00
485	B:ASP 511	21.72	0.00	0.00
486	B:THR 512	14.45	0.00	0.00
487	B:PRO 513	6.50	0.00	0.00
488	B:ASP 514	15.97	0.00	0.00
489	B:GLN 515	8.15	0.00	0.00
490	B:ALA 516	0.62	0.00	0.00
491	B:THR 517	3.18	0.00	0.00
492	B:ASP 518	15.41	0.00	0.00
493	B:GLY 519	0.12	0.00	0.00



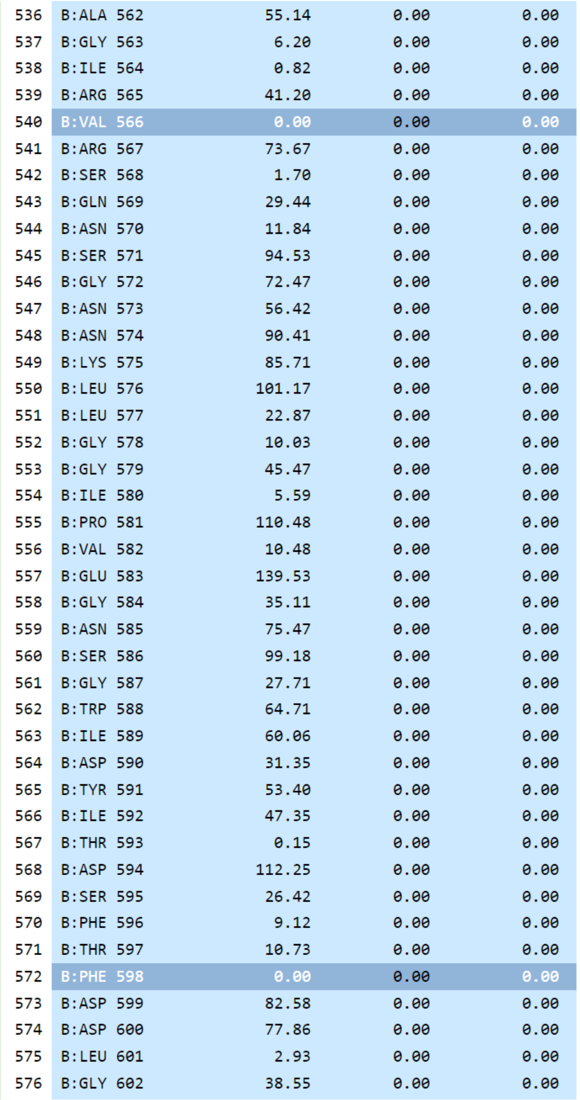
452	B:PHE 478	8.08	0.00	0.00
453	B:TYR 479	10.87	0.00	0.00
454	B:ARG 480	98.74	0.00	0.00
455	B:SER 481	45.10	0.00	0.00
456	B:ASN 482	52.49	0.00	0.00
457	B:SER 483	2.42	0.00	0.00
458	B:HIS 484	0.83	0.00	0.00
459	B:TYR 485	82.84	0.00	0.00
460	B:LEU 486	4.51	0.00	0.00
461	B:SER 487	39.90	0.00	0.00
462	B:THR 488	41.97	0.00	0.00
463	B:THR 489	119.21	0.00	0.00
464	B:ASP 490	59.84	0.00	0.00
465	B:ASP 491	74.01	0.00	0.00
466	B:ALA 492	8.94	0.00	0.00
467	B:TYR 493	48.72	0.00	0.00
468	B:VAL 494	5.61	0.00	0.00
469	B:ILE 495	1.01	0.00	0.00
470	B:PRO 496	6.85	0.00	0.00
471	B:ALA 497	0.00	0.00	0.00
472	B:LEU 498	6.37	0.00	0.00
473	B:GLN 499	10.84	0.00	0.00
474	B:PHE 500	1.89	0.00	0.00
475	B:SER 501	33.10	0.00	0.00
476	B:THR 502	65.33	0.00	0.00
477	B:VAL 503	16.36	0.00	0.00
478	B:SER 504	21.09	0.00	0.00
479	B:ASP 505	67.87	0.00	0.00
480	B:ARG 506	87.42	0.00	0.00
481	B:SER 507	0.00	0.00	0.00
482	B:PHE 508	21.45	0.00	0.00
483	B:LEU 509	8.73	0.00	0.00
484	B:GLU 510	38.73	0.00	0.00
485	B:ASP 511	21.72	0.00	0.00
486	B:THR 512	14.45	0.00	0.00
487	B:PRO 513	6.50	0.00	0.00
488	B:ASP 514	15.97	0.00	0.00
489	B:GLN 515	8.15	0.00	0.00
490	B:ALA 516	0.62	0.00	0.00
491	B:THR 517	3.18	0.00	0.00
492	B:ASP 518	15.41	0.00	0.00
493	B:GLY 519	0.12	0.00	0.00



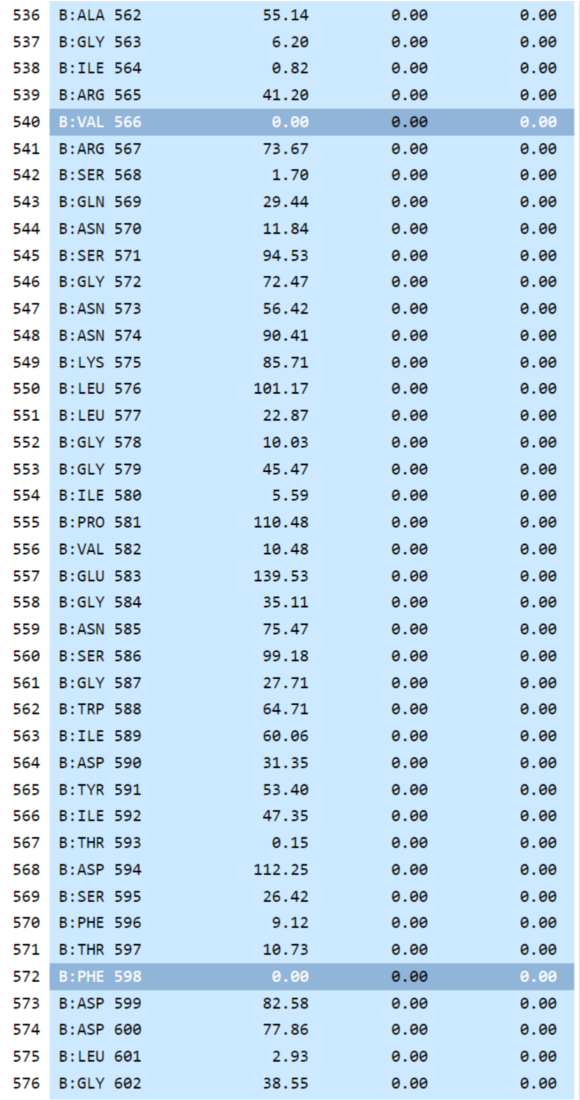
494	B:SER 520	0.12	0.00	0.00
495	B:ILE 521	0.00	0.00	0.00
496	B:LYS 522	45.76	0.00	0.00
497	B:PHE 523	0.00	0.00	0.00
498	B:THR 524	10.64	0.00	0.00
499	B:ASP 525	14.52	0.00	0.00
500	B:THR 526	88.84	0.00	0.00
501	B:VAL 527	119.65	0.00	0.00
502	B:LEU 528	164.66	0.00	0.00
503	B:GLY 529	17.32	0.00	0.00
504	B:ASN 530	0.85	0.00	0.00
505	B:GLU 531	63.04	0.00	0.00
506	B:ALA 532	0.00	0.00	0.00
507	B:LYS 533	52.67	0.00	0.00
508	B:TYR 534	0.00	0.00	0.00
509	B:SER 535	13.09	0.00	0.00
510	B:ILE 536	0.37	0.00	0.00
511	B:ARG 537	68.05	0.00	0.00
512	B:LEU 538	7.32	0.00	0.00
513	B:ASN 539	58.20	0.00	0.00
514	B:THR 540	123.19	0.00	0.00
515	B:GLY 541	68.42	0.00	0.00
516	B:PHE 542	13.04	0.00	0.00
517	B:ASN 543	83.42	0.00	0.00
518	B:THR 544	34.15	0.00	0.00
519	B:ALA 545	63.42	0.00	0.00
520	B:THR 546	36.94	0.00	0.00
521	B:ARG 547	107.84	0.00	0.00
522	B:TYR 548	6.25	0.00	0.00
523	B:ARG 549	64.37	0.00	0.00
524	B:LEU 550	0.00	0.00	0.00
525	B:ILE 551	0.95	0.00	0.00
526	B:ILE 552	0.35	0.00	0.00
527	B:ARG 553	4.70	0.00	0.00
528	B:PHE 554	7.85	0.00	0.00
529	B:LYS 555	48.33	0.00	0.00
530	B:ALA 556	3.09	0.00	0.00
531	B:PRO 557	62.45	0.00	0.00
532	B:ALA 558	24.38	0.00	0.00
533	B:ARG 559	136.41	0.00	0.00
534	B:LEU 560	13.04	0.00	0.00
535	B:ALA 561	91.18	0.00	0.00



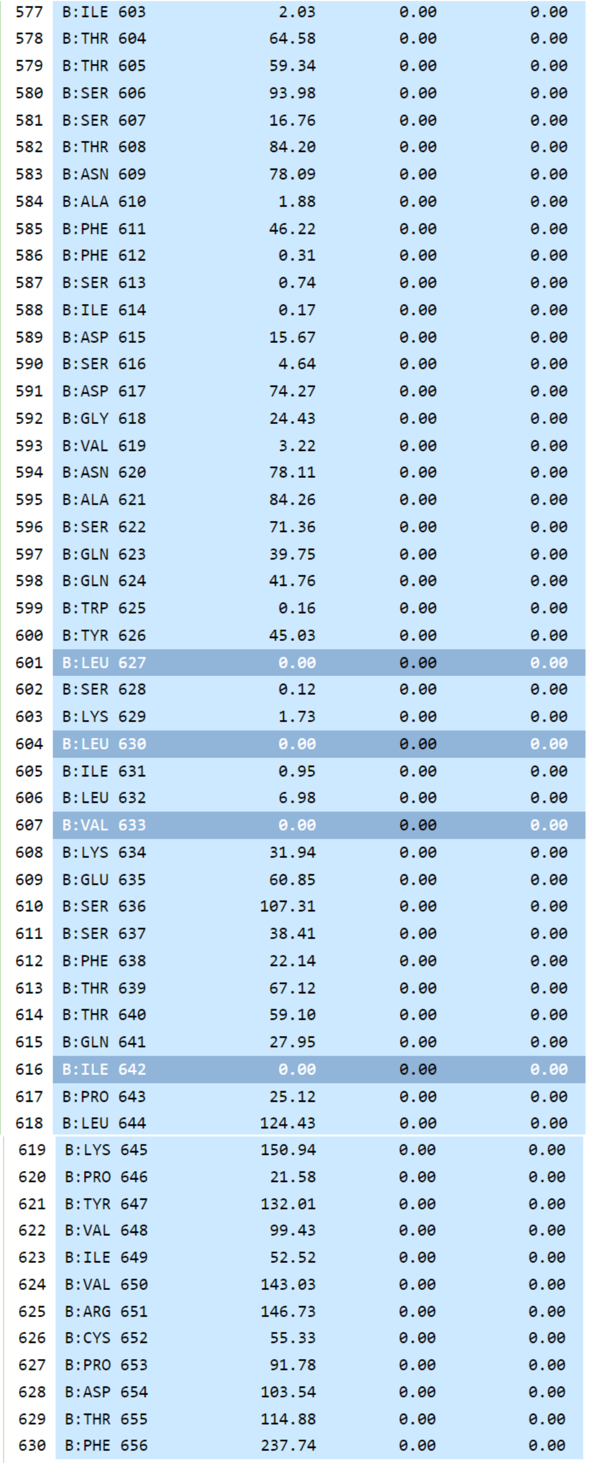
494	B:SER 520	0.12	0.00	0.00
495	B:ILE 521	0.00	0.00	0.00
496	B:LYS 522	45.76	0.00	0.00
497	B:PHE 523	0.00	0.00	0.00
498	B:THR 524	10.64	0.00	0.00
499	B:ASP 525	14.52	0.00	0.00
500	B:THR 526	88.84	0.00	0.00
501	B:VAL 527	119.65	0.00	0.00
502	B:LEU 528	164.66	0.00	0.00
503	B:GLY 529	17.32	0.00	0.00
504	B:ASN 530	0.85	0.00	0.00
505	B:GLU 531	63.04	0.00	0.00
506	B:ALA 532	0.00	0.00	0.00
507	B:LYS 533	52.67	0.00	0.00
508	B:TYR 534	0.00	0.00	0.00
509	B:SER 535	13.09	0.00	0.00
510	B:ILE 536	0.37	0.00	0.00
511	B:ARG 537	68.05	0.00	0.00
512	B:LEU 538	7.32	0.00	0.00
513	B:ASN 539	58.20	0.00	0.00
514	B:THR 540	123.19	0.00	0.00
515	B:GLY 541	68.42	0.00	0.00
516	B:PHE 542	13.04	0.00	0.00
517	B:ASN 543	83.42	0.00	0.00
518	B:THR 544	34.15	0.00	0.00
519	B:ALA 545	63.42	0.00	0.00
520	B:THR 546	36.94	0.00	0.00
521	B:ARG 547	107.84	0.00	0.00
522	B:TYR 548	6.25	0.00	0.00
523	B:ARG 549	64.37	0.00	0.00
524	B:LEU 550	0.00	0.00	0.00
525	B:ILE 551	0.95	0.00	0.00
526	B:ILE 552	0.35	0.00	0.00
527	B:ARG 553	4.70	0.00	0.00
528	B:PHE 554	7.85	0.00	0.00
529	B:LYS 555	48.33	0.00	0.00
530	B:ALA 556	3.09	0.00	0.00
531	B:PRO 557	62.45	0.00	0.00
532	B:ALA 558	24.38	0.00	0.00
533	B:ARG 559	136.41	0.00	0.00
534	B:LEU 560	13.04	0.00	0.00
535	B:ALA 561	91.18	0.00	0.00



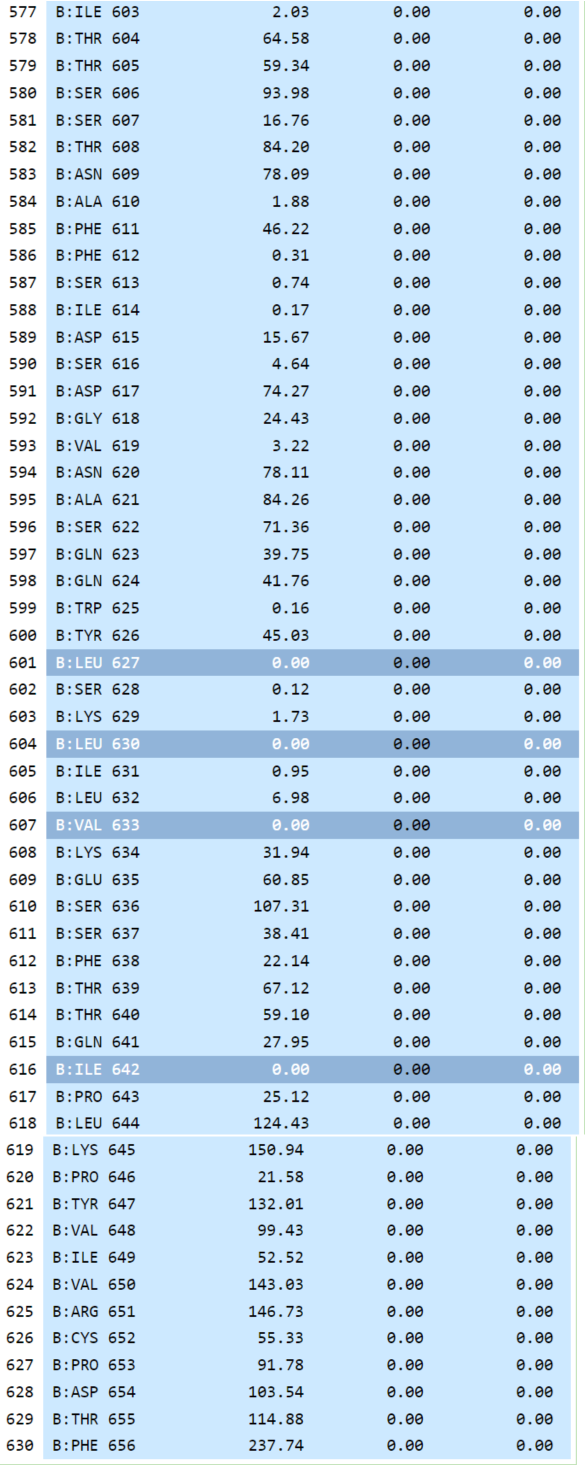
536	B:ALA 562	55.14	0.00	0.00
537	B:GLY 563	6.20	0.00	0.00
538	B:ILE 564	0.82	0.00	0.00
539	B:ARG 565	41.20	0.00	0.00
540	B:VAL 566	0.00	0.00	0.00
541	B:ARG 567	73.67	0.00	0.00
542	B:SER 568	1.70	0.00	0.00
543	B:GLN 569	29.44	0.00	0.00
544	B:ASN 570	11.84	0.00	0.00
545	B:SER 571	94.53	0.00	0.00
546	B:GLY 572	72.47	0.00	0.00
547	B:ASN 573	56.42	0.00	0.00
548	B:ASN 574	90.41	0.00	0.00
549	B:LYS 575	85.71	0.00	0.00
550	B:LEU 576	101.17	0.00	0.00
551	B:LEU 577	22.87	0.00	0.00
552	B:GLY 578	10.03	0.00	0.00
553	B:GLY 579	45.47	0.00	0.00
554	B:ILE 580	5.59	0.00	0.00
555	B:PRO 581	110.48	0.00	0.00
556	B:VAL 582	10.48	0.00	0.00
557	B:GLU 583	139.53	0.00	0.00
558	B:GLY 584	35.11	0.00	0.00
559	B:ASN 585	75.47	0.00	0.00
560	B:SER 586	99.18	0.00	0.00
561	B:GLY 587	27.71	0.00	0.00
562	B:TRP 588	64.71	0.00	0.00
563	B:ILE 589	60.06	0.00	0.00
564	B:ASP 590	31.35	0.00	0.00
565	B:TYR 591	53.40	0.00	0.00
566	B:ILE 592	47.35	0.00	0.00
567	B:THR 593	0.15	0.00	0.00
568	B:ASP 594	112.25	0.00	0.00
569	B:SER 595	26.42	0.00	0.00
570	B:PHE 596	9.12	0.00	0.00
571	B:THR 597	10.73	0.00	0.00
572	B:PHE 598	0.00	0.00	0.00
573	B:ASP 599	82.58	0.00	0.00
574	B:ASP 600	77.86	0.00	0.00
575	B:LEU 601	2.93	0.00	0.00
576	B:GLY 602	38.55	0.00	0.00



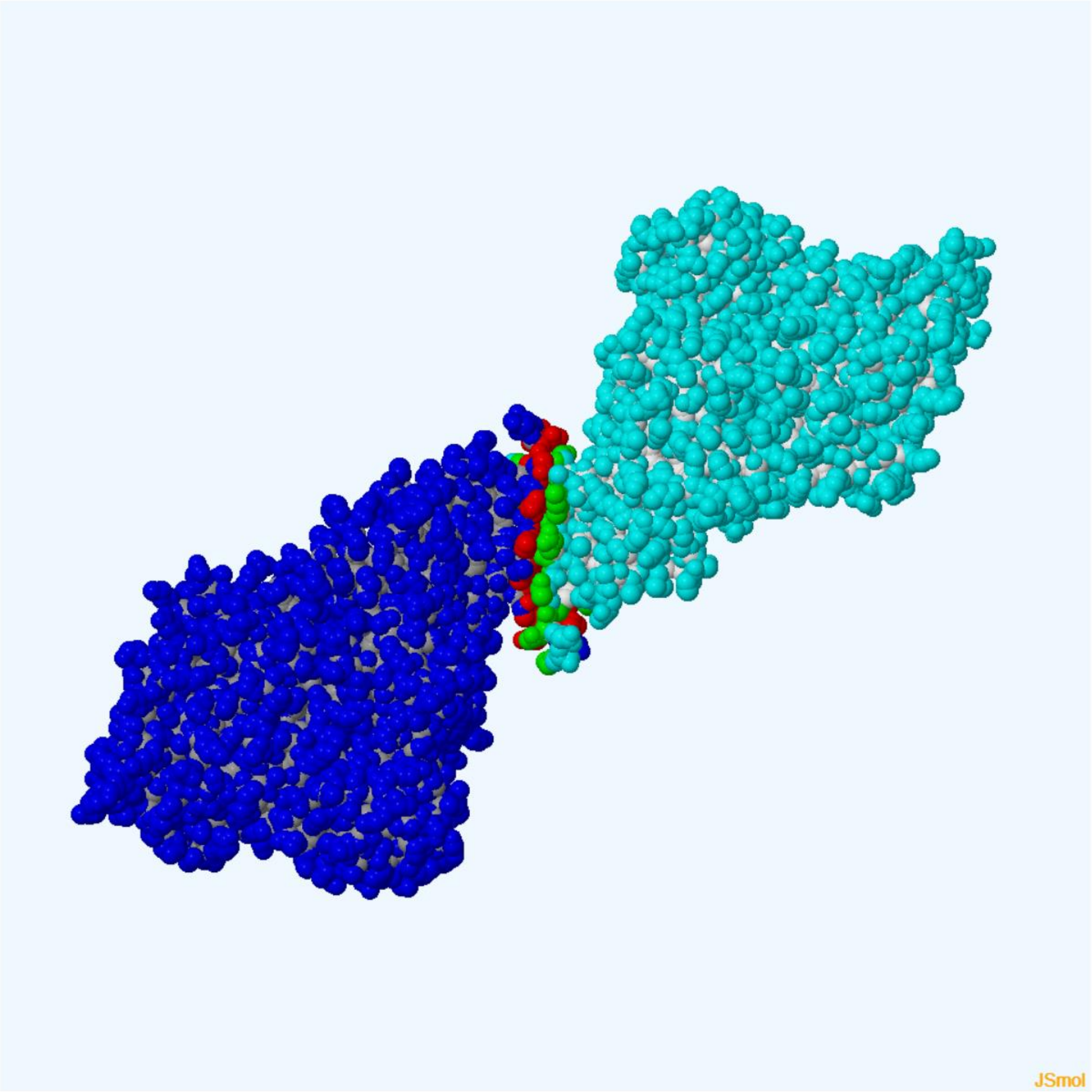
536	B:ALA 562	55.14	0.00	0.00
537	B:GLY 563	6.20	0.00	0.00
538	B:ILE 564	0.82	0.00	0.00
539	B:ARG 565	41.20	0.00	0.00
540	B:VAL 566	0.00	0.00	0.00
541	B:ARG 567	73.67	0.00	0.00
542	B:SER 568	1.70	0.00	0.00
543	B:GLN 569	29.44	0.00	0.00
544	B:ASN 570	11.84	0.00	0.00
545	B:SER 571	94.53	0.00	0.00
546	B:GLY 572	72.47	0.00	0.00
547	B:ASN 573	56.42	0.00	0.00
548	B:ASN 574	90.41	0.00	0.00
549	B:LYS 575	85.71	0.00	0.00
550	B:LEU 576	101.17	0.00	0.00
551	B:LEU 577	22.87	0.00	0.00
552	B:GLY 578	10.03	0.00	0.00
553	B:GLY 579	45.47	0.00	0.00
554	B:ILE 580	5.59	0.00	0.00
555	B:PRO 581	110.48	0.00	0.00
556	B:VAL 582	10.48	0.00	0.00
557	B:GLU 583	139.53	0.00	0.00
558	B:GLY 584	35.11	0.00	0.00
559	B:ASN 585	75.47	0.00	0.00
560	B:SER 586	99.18	0.00	0.00
561	B:GLY 587	27.71	0.00	0.00
562	B:TRP 588	64.71	0.00	0.00
563	B:ILE 589	60.06	0.00	0.00
564	B:ASP 590	31.35	0.00	0.00
565	B:TYR 591	53.40	0.00	0.00
566	B:ILE 592	47.35	0.00	0.00
567	B:THR 593	0.15	0.00	0.00
568	B:ASP 594	112.25	0.00	0.00
569	B:SER 595	26.42	0.00	0.00
570	B:PHE 596	9.12	0.00	0.00
571	B:THR 597	10.73	0.00	0.00
572	B:PHE 598	0.00	0.00	0.00
573	B:ASP 599	82.58	0.00	0.00
574	B:ASP 600	77.86	0.00	0.00
575	B:LEU 601	2.93	0.00	0.00
576	B:GLY 602	38.55	0.00	0.00



577	B:ILE 603	2.03	0.00	0.00
578	B:THR 604	64.58	0.00	0.00
579	B:THR 605	59.34	0.00	0.00
580	B:SER 606	93.98	0.00	0.00
581	B:SER 607	16.76	0.00	0.00
582	B:THR 608	84.20	0.00	0.00
583	B:ASN 609	78.09	0.00	0.00
584	B:ALA 610	1.88	0.00	0.00
585	B:PHE 611	46.22	0.00	0.00
586	B:PHE 612	0.31	0.00	0.00
587	B:SER 613	0.74	0.00	0.00
588	B:ILE 614	0.17	0.00	0.00
589	B:ASP 615	15.67	0.00	0.00
590	B:SER 616	4.64	0.00	0.00
591	B:ASP 617	74.27	0.00	0.00
592	B:GLY 618	24.43	0.00	0.00
593	B:VAL 619	3.22	0.00	0.00
594	B:ASN 620	78.11	0.00	0.00
595	B:ALA 621	84.26	0.00	0.00
596	B:SER 622	71.36	0.00	0.00
597	B:GLN 623	39.75	0.00	0.00
598	B:GLN 624	41.76	0.00	0.00
599	B:TRP 625	0.16	0.00	0.00
600	B:TYR 626	45.03	0.00	0.00
601	B:LEU 627	0.00	0.00	0.00
602	B:SER 628	0.12	0.00	0.00
603	B:LYS 629	1.73	0.00	0.00
604	B:LEU 630	0.00	0.00	0.00
605	B:ILE 631	0.95	0.00	0.00
606	B:LEU 632	6.98	0.00	0.00
607	B:VAL 633	0.00	0.00	0.00
608	B:LYS 634	31.94	0.00	0.00
609	B:GLU 635	60.85	0.00	0.00
610	B:SER 636	107.31	0.00	0.00
611	B:SER 637	38.41	0.00	0.00
612	B:PHE 638	22.14	0.00	0.00
613	B:THR 639	67.12	0.00	0.00
614	B:THR 640	59.10	0.00	0.00
615	B:GLN 641	27.95	0.00	0.00
616	B:ILE 642	0.00	0.00	0.00
617	B:PRO 643	25.12	0.00	0.00
618	B:LEU 644	124.43	0.00	0.00
619	B:LYS 645	150.94	0.00	0.00
620	B:PRO 646	21.58	0.00	0.00
621	B:TYR 647	132.01	0.00	0.00
622	B:VAL 648	99.43	0.00	0.00
623	B:ILE 649	52.52	0.00	0.00
624	B:VAL 650	143.03	0.00	0.00
625	B:ARG 651	146.73	0.00	0.00
626	B:CYS 652	55.33	0.00	0.00
627	B:PRO 653	91.78	0.00	0.00
628	B:ASP 654	103.54	0.00	0.00
629	B:THR 655	114.88	0.00	0.00
630	B:PHE 656	237.74	0.00	0.00



577	B:ILE 603	2.03	0.00	0.00
578	B:THR 604	64.58	0.00	0.00
579	B:THR 605	59.34	0.00	0.00
580	B:SER 606	93.98	0.00	0.00
581	B:SER 607	16.76	0.00	0.00
582	B:THR 608	84.20	0.00	0.00
583	B:ASN 609	78.09	0.00	0.00
584	B:ALA 610	1.88	0.00	0.00
585	B:PHE 611	46.22	0.00	0.00
586	B:PHE 612	0.31	0.00	0.00
587	B:SER 613	0.74	0.00	0.00
588	B:ILE 614	0.17	0.00	0.00
589	B:ASP 615	15.67	0.00	0.00
590	B:SER 616	4.64	0.00	0.00
591	B:ASP 617	74.27	0.00	0.00
592	B:GLY 618	24.43	0.00	0.00
593	B:VAL 619	3.22	0.00	0.00
594	B:ASN 620	78.11	0.00	0.00
595	B:ALA 621	84.26	0.00	0.00
596	B:SER 622	71.36	0.00	0.00
597	B:GLN 623	39.75	0.00	0.00
598	B:GLN 624	41.76	0.00	0.00
599	B:TRP 625	0.16	0.00	0.00
600	B:TYR 626	45.03	0.00	0.00
601	B:LEU 627	0.00	0.00	0.00
602	B:SER 628	0.12	0.00	0.00
603	B:LYS 629	1.73	0.00	0.00
604	B:LEU 630	0.00	0.00	0.00
605	B:ILE 631	0.95	0.00	0.00
606	B:LEU 632	6.98	0.00	0.00
607	B:VAL 633	0.00	0.00	0.00
608	B:LYS 634	31.94	0.00	0.00
609	B:GLU 635	60.85	0.00	0.00
610	B:SER 636	107.31	0.00	0.00
611	B:SER 637	38.41	0.00	0.00
612	B:PHE 638	22.14	0.00	0.00
613	B:THR 639	67.12	0.00	0.00
614	B:THR 640	59.10	0.00	0.00
615	B:GLN 641	27.95	0.00	0.00
616	B:ILE 642	0.00	0.00	0.00
617	B:PRO 643	25.12	0.00	0.00
618	B:LEU 644	124.43	0.00	0.00
619	B:LYS 645	150.94	0.00	0.00
620	B:PRO 646	21.58	0.00	0.00
621	B:TYR 647	132.01	0.00	0.00
622	B:VAL 648	99.43	0.00	0.00
623	B:ILE 649	52.52	0.00	0.00
624	B:VAL 650	143.03	0.00	0.00
625	B:ARG 651	146.73	0.00	0.00
626	B:CYS 652	55.33	0.00	0.00
627	B:PRO 653	91.78	0.00	0.00
628	B:ASP 654	103.54	0.00	0.00
629	B:THR 655	114.88	0.00	0.00
630	B:PHE 656	237.74	0.00	0.00


JSmol

Appendix C

Self-assembly of *Bacillus thuringiensis* subsp. *israelensis* during sporulation harboring Cry11Ba crystalline inclusions

C.1 Introduction

Alongside understanding Cry11Ba at the macromolecular and complex level, the self-assembly of the Cry11Ba crystals *in vivo* has been a fascinating topic to explore. When thinking about canonical crystallography techniques, the biophysical mechanics are understood, but there is still a significant gap in understanding which crystals will and will not produce well-packaged/organized crystals that are conducive for crystallographic structural studies.[1] While whole careers can be spent in attempts of crystallizing a single complex or protein, *Bacillus thuringiensis* (*Bt*) have evolved to a level of efficiency that scientists dream of when producing protein crystals. The *Bacillus thuringiensis* (*Bt*) system displays how it is possible to consistently produce protein crystals[2–4] and raises the questions: how is this possible, are there specific cellular structures that aid in crystallization, and how soon do crystals develop in the sporulation process (**Chapter 1, Fig. 1.2**). This ability to study the sporulation process *in vivo* with high resolution was not possible until the recent *resolution revolution* that occurred within the structural field and the evolution of its techniques, especially in electron microscopy (EM) where the 1 µm - 1 Å range are now filled with techniques that can produce useful high-resolution structural data. The biggest gap was around the 1 µm - 10 nm range, which is often where cellular structures' sizes range and are not observable from light microscopy alone.[5,6]

C.2 Results and Discussion

By utilizing transmission electron microscopy (TEM) and cyto-cryo electron tomography (cyto-cryoET), the Cry11Ba crystalline inclusions nucleation and maturation through sporulation can be studied *in vivo*. This is vital as this is the natural formation process of these Cry and Cyt toxins and removes the concern of a protein being altered *in vitro* that could affect its structure during crystallization. First, imaging the *Bt* throughout their sporulation process via TEM was crucial in mapping and timing the full process. From these studies, it was determined that sporulation occurs between 8-10 total hours from asymmetric division to fully lysed cells. There were also key points in the sporulation process that were observed that had not been previously reported. One was the condensing of the nucleoid for asymmetric division, where microcrystals observed, thus crystal nucleation happens earlier than originally believed in the sporulation process (**Figs. C.1a-c,C.2**). These crystals were also only observable via a TF200 EM, which utilizes a Field Emission Gun (FEG) which provides a more coherent, brighter beam of electrons at a higher energy level to penetrate thicker samples, such as these *Bt* cells during sporulation. The next stage of interest observed was the clear separation of the crystalline inclusion and spore still within the cell. During this stage, a thin membrane was observed within the *Bt* cell that had not previously been seen (**Fig. C.2**). This could be one of two possible cellular structures. The first being another layer of the spore cortex that is beginning to layer to the spore; however, it is formed throughout the full volume of the cell. The second is that this is a membrane that is aiding in coordination of the crystalline inclusion and spore and a concentrator, similar to a sitting or hanging crystallization drop, that increases the amount of Cry11Ba in an area to help increase collisions and thus crystal packing. While the first is possible, it is less probable as in order to fully engulf the spore evenly it would have to cover the crystal and would not be able to pass it without puncturing a hole in it. The second is more probable as this would answer how the crystal nucleation is possible in such a large volume compared to Cry11Ba monomers and the coordination would help define the poles of the *Bt* cells to regulate resources during the resource heavy process of sporulation correctly. Third, the spore is often stuck within the cell membrane upon sporulation (**Fig. C.3**). While the typical way to purify crystals incorporates sonication to mechanically separate spores and crystals from cell membranes and ensure disentanglement, that is not the same in nature. Thus, it is possible that crystal size and shape may be a factor in determining cell lysis and their ability to exit the cell more efficiently than the membrane. Through TEMs, the sporulation process has been recapitulated to other proposed schematics and observed some other events during sporulation whose features would not have been resolved without the magnification power of EM.

The sporulation state is observable via phase contrast confocal microscopy, which converts phase shifts in light that pass through a specimen's transparent parts and causes different amounts of light to pass through the sample. For *Bt* cells, this has been crucial in identifying the sporulation state as this allows for the quickest imaging of the cells in determining when collecting and determining whether the cells have reached the correct sporulation. The clearest sporulation state for *Bt* cells is the late sporulation state where a clear distinction between endospore and crystalline inclusion are observed (**Fig. C.4**). Phase contrast confocal microscopy is also used in determining and confirming the purity of fractions after sucrose gradients where the crystalline inclusions and endospores will be in separate phases or focal planes and aids in identification of each. By combining the power of phase contrast microscopy with fluorescence imaging, a more complete understanding can be reached for *in vivo* processes and imaging. Within *Bt* cells, it has been believed that there are other cellular structures or proteins that aid in the sporulation process, whether that be crystalline inclusion or endospore maturation. Bt152 is one such protein, where in absence of this, the parasporal body is not stable enough to form and degrades within the *Bt* cell; however, in the presence it is able to rescue and stabilize the parasporal body and for lamellar matrices' formation.[7] With this understanding of Bt152 being required to help stabilize and rescue the crystalline inclusion formation, it was tagged with Green Fluorescent Protein (GFP) in order to see which part of the parasporal body it coordinates to during sporulation. Time points of cultures with Bt152-GFP expressed within *Bt* cells harboring Cry11Ba crystals were collected over the time period when sporulation was observed and at the 12 hr time point, the fluorescent signal was high and diffused without the cell, with little to none outside the *Bt* cell membranes. Upon further growth to the 16 hr time point, the fluorescent signal displayed accumulation around the crystalline inclusion and not the endospore (**Fig. C.5a,b**). This coalescing around the Cry11Ba crystalline inclusion can be understood as Bt152 was able to rescue the crystalline inclusions formation within the parasporal body, thus it would be likely that it has some interactions with the crystallizing protein, ie. Cry11Ba, in some way helps coordinate the crystallization *in vivo* by way of acting as a concentrator and encapsulating the protein within a cellular structure to increase Cry11Ba collisions for effective and consistent crystal packing. The other possibility is that it interacts with the Cry11Ba monomers in such a way to stabilize them until crystal nucleation occurs and then decorates the crystalline inclusion as an artifact of residual interactions on the crystal surface. Bt152 was observed to have a similar sequence to that of Hemagglutinin-33 (HA-33), which is part of the neurotoxin associated proteins, which helped with activity and binding to the target cells while being resistant to proteases.[8] While different in their host organisms, it is interesting that these two proteins are both expressed in rod-shaped bacteria, interact with a toxin, and are responsible for the toxin's stability/activity. The conserved residues in both could be the key factors in being able to bind to the toxin and through additional analyses indicate within the elusive paratoxins, where the most active site is located by mapping their interactions.

New events about the crystalline inclusion formation have been identified; however, due to the lower energy of the TEM microscopes utilized and thickness of the samples are still limiting more structural information from being gained. To overcome this, a stronger electron source was utilized, vitrifying *Bti* cells, and increasing contrast with a Volta phase plate, the cellular structure details in the sporulation process can be elucidated. With *Bt* cells harboring the Cry11Ba crystalline inclusions at various time points vitrified to preserve the cellular structures, but halt the maturation process, a Focused Ion Beam (FIB) will mill these to a thickness that allows for imaging the most interior cellular structures. Preliminary screening and experiments have shown proof of concept, where *Bt* cells that were close to natural lysis, resulting in thin enough membranes to observe cellular content through thin vitrified ice (**Fig. C.6a-c**). In this initial tomogram, it was observed that there are many vacuoles, the endospore, the crystalline inclusion, and some residual bands that ran through the cell. The residual bands could be the result of the exterior membrane sitting on the grid in a way that folded the membrane that way or a new membrane or cellular structure that aids in helping coordinate and form the Cry11Ba crystalline inclusion. Additional *Bt* cells at earlier time points have been vitrified for future imaging to reach high enough resolution to observe any additional physical cell machinery that has not been observable previously.

C.3 Figures

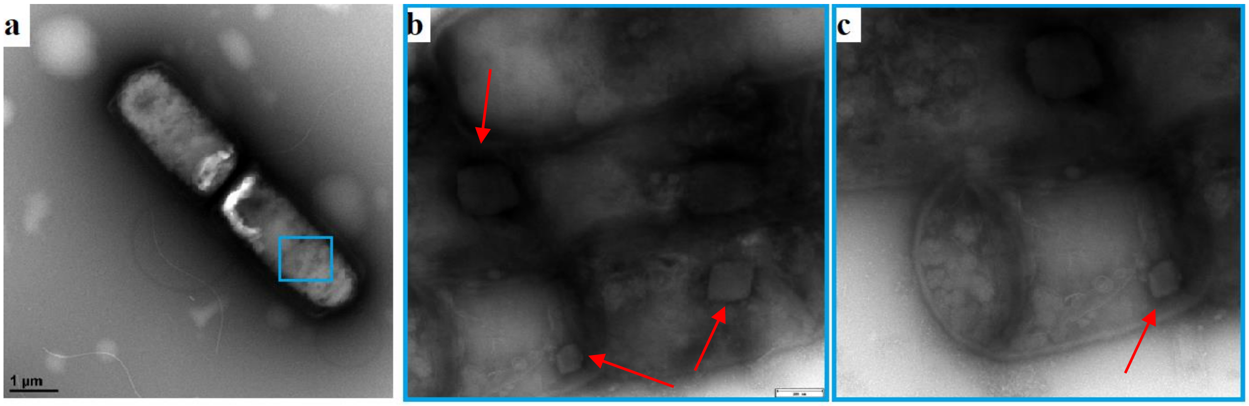


Figure C.1 Early sporulation stages of *Bacillus thuringiensis* subsp. *israelensis*. Negative stain electron micrographs of Bti cells early during sporulation when the nucleoid has condensed at the beginning of sporulation. This network is dense, with no detailed cellular structures observed (**a**). Upon imaging them with higher energy source that are able to penetrate the density of the material (blue box), the (**b-c**) presence of microcrystals (red arrows) are observed and the beginnings of the asymmetric division.

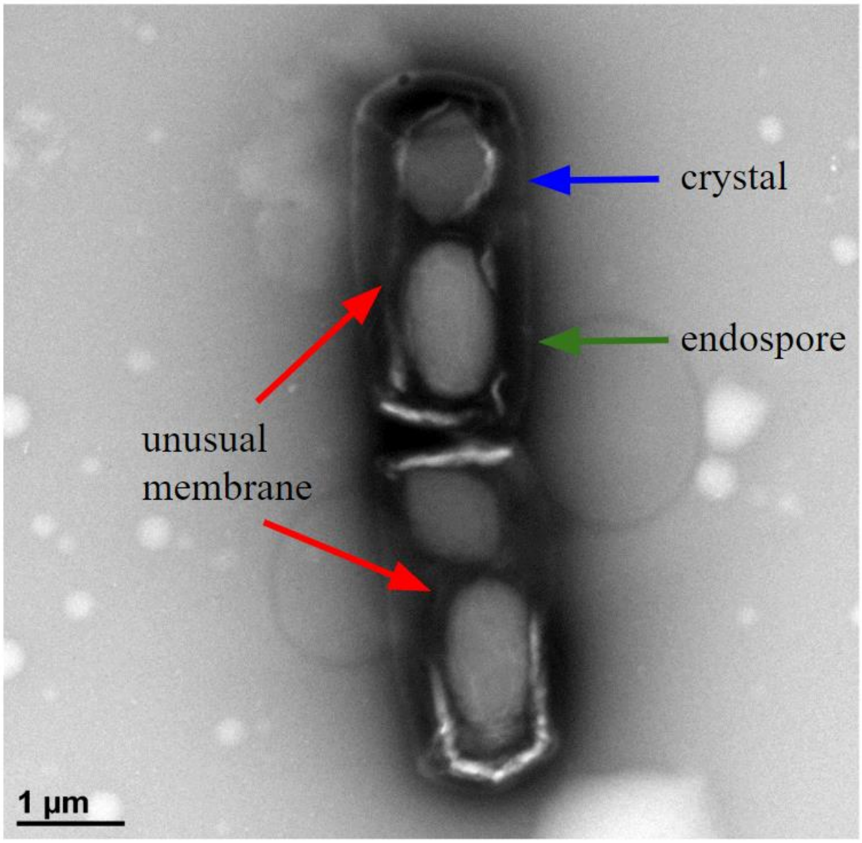


Figure C.2 ***Bacillus thuringiensis*** **subsp.** ***israelensis*** **intact cells with Cry11Ba crystal and endospore via negative stain electron microscopy.** *Bacillus thuringiensis* subsp. *israelensis* (*Bti*) having undergone the sporulation process and reached maturation before natural cell lysis. This is identified by the formation of complete crystalline inclusion (blue arrow) and endospore (green arrow), which occurs late in sporulation. The crystalline inclusion and spore size are usually similar in size, differ in shape, and size are dependent upon the *Bacillus* species as a vessel. With the negative staining EM, a thin membrane (red arrows) was observed forming during late sporulation after crystal and endospore maturation, which is assumed to be encapsulating the entire parasporal body throughout sporulation.

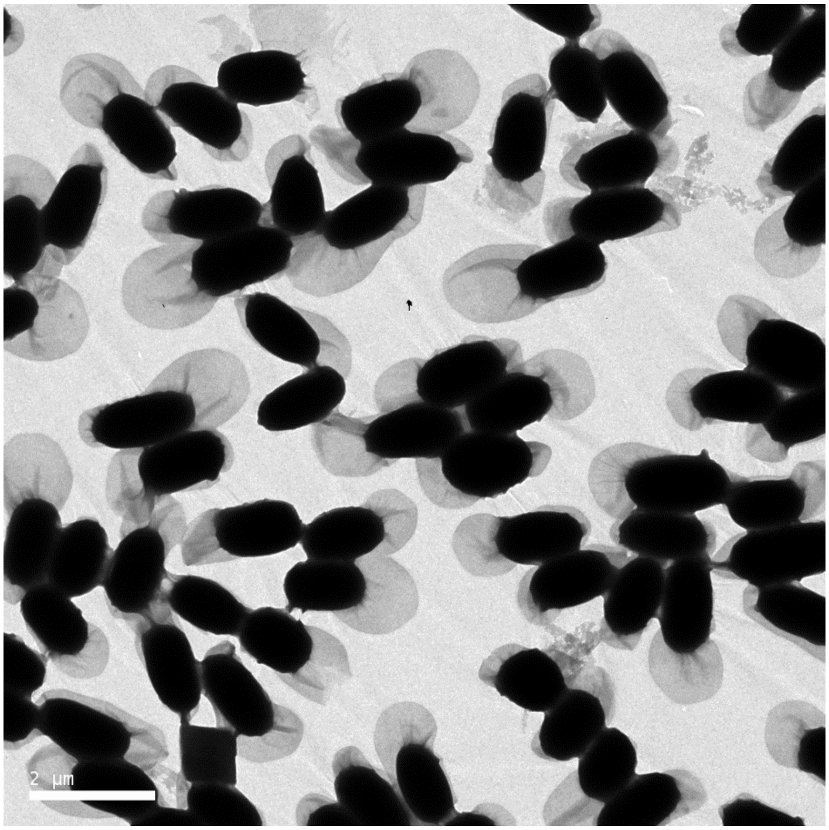


Figure C.3 Late sporulation stages of *Bacillus thuringiensis* subsp. *israelensis*. After purification of the Cry11Ba crystals, the lowest fraction collected typically contains spores and unlysed cells. The endospores are often time still attached to the cell membranes. This is hypothesized to be due to the mechanical natural lysis process causing the cell membranes to entangle around the endospores or they have membrane proteins that keep them coordinate and allow the crystals to release from the cell and the endospores remain protected by another layer.

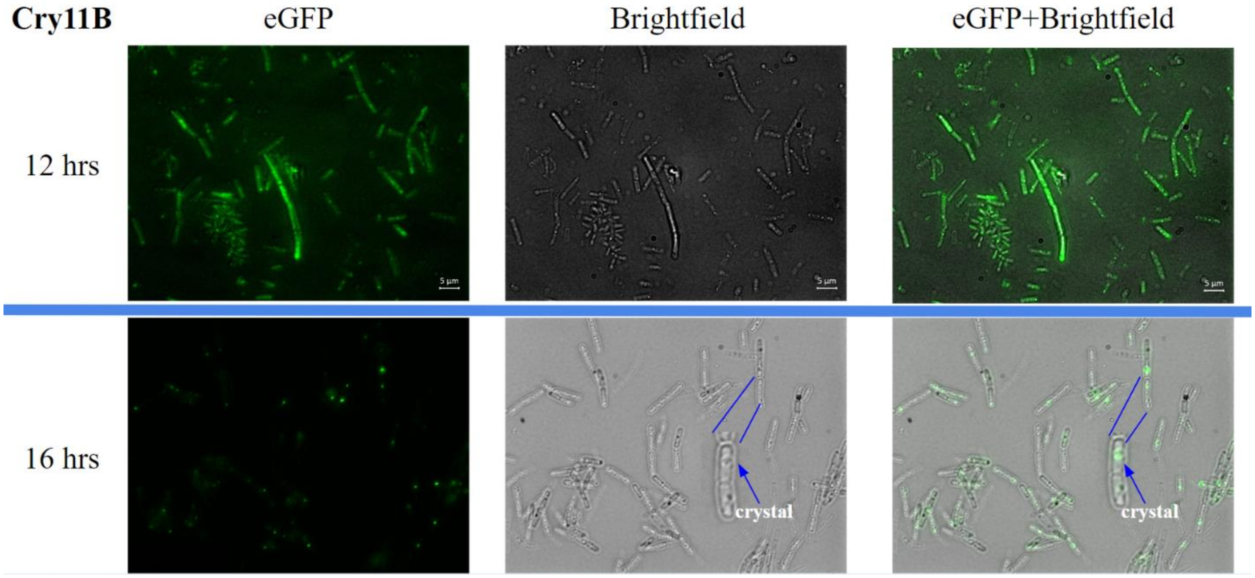


Figure C.4 Bt152-GFP tracking during sporulation stages of *Bacillus thuringiensis* subsp. *israelensis*. As the formation of the crystal and endospore progress during sporulation, there have been other proteins hypothesized to aid in positioning the two components during asymmetric division. Bt152 was found on the same pBtoxis[9] plasmid, which is the toxin-coding plasmid within Bti, thus it could be involved in the sporulation process. Bt152 was tagged with GFP and found to have (**a**) diffused signal early in sporulation indicating expression and found to (**b**) localize to the crystal (blue arrow) *in vivo* during sporulation. The previously observed membrane could be Bt152 or parasporal body protein to aid in this interaction and formation.

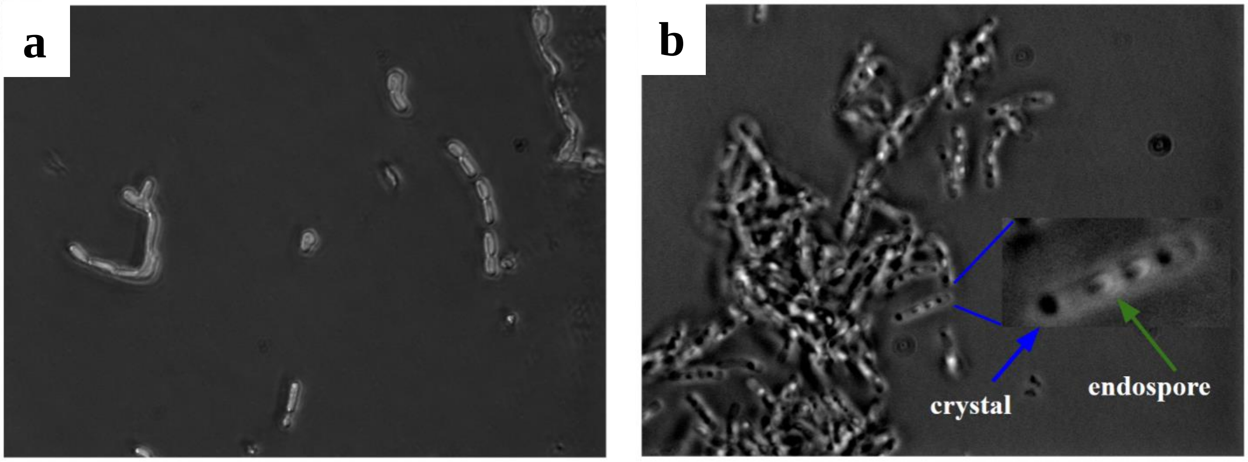


Figure C.5 Crystal and endospore development within *Bacillus thuringiensis* subsp. *israelensis*. The cells are able to be imaged via confocal microscopy to observe the internalized morphology change, where the cells are empty rod-shaped vessels at (**a**) 12 hrs and contain new species at (**b**) 16 hrs. This is further enhanced by phase contrast that shows the distinction of crystal (blue arrow, black square) and endospore (green arrow, gray oval). Without the phase contrast, these species would not be as distinct, but due to the limitations of confocal light microscopy detailed cellular features in sporulation are not observable.

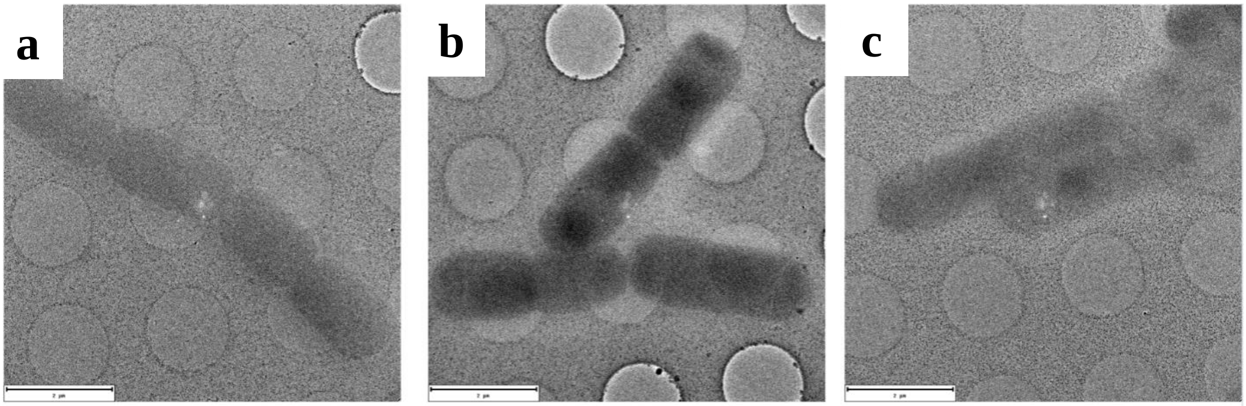


Figure C.6 Vitrification of *Bacillus thuringiensis* subsp. *israelensis* during late sporulation. The *Bti* cells are sensitive to their environment, which is the sensor that triggers the cells to engage in sporulation. In order to preserve the cells for cryo-ET, they are vitrified in ice, which preserves the cellular structures without affecting them as they freeze rapidly and prevent the formation of crystalline ice. The amorphous ice thickness (**a,b**) is the limiting factor for sample preparation and difficult as the cells are also 1 µm thick and needs to be (**c**) just thick enough to cover the cells, while also be thin enough to continue to observe the cells' features. This can be combatted by FIB milling, but the thickness must still be thin enough to prevent over-milling the lamellae.

C.4 References

(1) Rupp, B. *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*; Garland Science, 2009.

(2) Federici, B. A.; Park, H.-W.; Sakano, Y. Insecticidal Protein Crystals of Bacillus Thuringiensis. *Inclusions in prokaryotes* **2006**, 195–236.

(3) Sansinenea, E. Discovery and Description of Bacillus Thuringiensis. *Bacillus thuringiensis biotechnology* **2012**, 3–18.

(4) Khanna, K.; Lopez-Garrido, J.; Zhao, Z.; Watanabe, R.; Yuan, Y.; Sugie, J.; Pogliano, K.; Villa, E. The Molecular Architecture of Engulfment during Bacillus Subtilis Sporulation. *Elife* **2019**, *8*, e45257.

(5) Quentin, D.; Raunser, S. Electron Cryomicroscopy as a Powerful Tool in Biomedical Research. *J Mol Med* **2018**, *96* (6), 483–493. https://doi.org/10.1007/s00109-018-1640-y.

(6) Kochovski, Z.; Chen, G.; Yuan, J.; Lu, Y. Cryo-Electron Microscopy for the Study of Self-Assembled Poly(Ionic Liquid) Nanoparticles and Protein Supramolecular Structures. *Colloid Polym Sci* **2020**, *298* (7), 707–717. https://doi.org/10.1007/s00396-020-04657-w.

(7) Diaz-Mendoza, M.; Bideshi, D. K.; Federici, B. A. A 54-Kilodalton Protein Encoded by PBtoxis Is Required for Parasporal Body Structural Integrity in Bacillus Thuringiensis Subsp. Israelensis. *J Bacteriol* **2012**, *194* (6), 1562–1571. https://doi.org/10.1128/JB.06095-11.

(8) Sayadmanesh, A.; Ebrahimi, F.; Hajizade, A.; Rostamian, M.; Keshavarz, H. Expression and Purification of Neurotoxin-Associated Protein HA-33/A from Clostridium Botulinum and Evaluation of Its Antigenicity. *Iran Biomed J* **2013**, *17* (4), 165–170. https://doi.org/10.6091/ibj.1216.2013.

(9) Berry, C.; O’Neil, S.; Ben-Dov, E.; Jones, A. F.; Murphy, L.; Quail, M. A.; Holden, M. T. G.; Harris, D.; Zaritsky, A.; Parkhill, J. Complete Sequence and Organization of PBtoxis, the Toxin-Coding Plasmid of Bacillus Thuringiensis Subsp. Israelensis. *Appl Environ Microbiol* **2002**, *68* (10), 5082–5095. https://doi.org/10.1128/AEM.68.10.5082-5095.2002.