

**“Biophysical and Structural
Characterization of Bacterial
Proteins”**

Thesis submitted for the degree of

Doctor of philosophy (Sc.)

In

Biophysics and Molecular Biology

By

Paramita Saha

Department of Biophysics, Molecular Biology and Bioinformatics

University of Calcutta

2016

To my family

Acknowledgements

I take this opportunity to extend my vote of thanks to all those who made this thesis possible. First and foremost, I would like to express my sincere gratitude to my supervisor Dr. Mahua Ghosh for her continuous support, guidance and encouragement through out the research work. I am extremely grateful to Professor Jaydeb Chakrabarti for his valuable suggestions, which helped a lot to understand the work and improve the manuscripts. My special words of thanks to our collaborator Dr. Santasabuj Das and his research group, for helping and providing laboratory facility wherever required. I would like to thank my thesis committee member Professor Gautam Basu, Bose Institute for his valuable suggestions. Thanks to all my Ph.D committee members Professor Abhay Sankar Chakraborti, Professor P. K. Sengupta, Dr. Subhasis Mukherjee and Dr. Sarmistha Raychaudhury of Rajabazar Science College and Professor Ratan Gachhui of Jadavpur University, Kolkata for helpful comments. I would like to extend my vote of thanks to Professor Chanchal Kumar DasGupta and Dr. Debasis Das, Rajabazar Science College for allowing me to carry out the project work during the initial days of my Ph.D tenure. My special regards to all my teachers at school, undergraduate and postgraduate level.

I am thankful to Mr. Surajit Mukherjee, SNBNCBS for his help and assistance during the experiments. I would like to acknowledge the high field NMR and MALDI-TOF facility at TIFR, Mumbai. I extend my vote of thanks to Professor K. V. R. Chary, Dr. Ankona Datta, Ms. Anindita Sarkar, Mrs. Mamata Joshi, Mr. Manoj Naik, Mr. Jayesh Malkan, Mr. Devidas Jadhav, Mr. Mohan Dhabolkar and Mrs. Geetanjali A. Dhotre for their kind support during my visit to TIFR, Mumbai. I am thankful to Deepshikha Verma and Janeka Gartia, TIFR, Mumbai, for helping me in NMR analysis. I thank DBT-CU-IPLS facility and Mr. Souvik Roy, Ballygunge Science College for the calorimetry measurements.

Great friends always make days fun and fruitful. I thank my fellow group members: Dr. Amit Das, Dr. Manas Mondal, Samapan Sikdar, Suman Dutta, Sutapa Dutta, Piya Patra and Camelia Manna. I extend my word of thanks to other friends who were part of my days at SNBNCBS. I would like to thank the canteen staff of SNBNCBS for all their help and service.

This journey would not have been possible without the support of my family. My deepest gratitude goes to my parents, Rekha Saha and Dipak Chandra Saha. I thank my younger brother Debojit Saha for his love and affection.

I gratefully acknowledge SNBNCBS for providing me the financial support, a wonderful research environment and facilities, which helped a lot in carrying out the research work.

Finally, I owe thanks to everybody who has been important, whom I might not have mentioned, thank you very much.

Paramita Saha

*S. N. Bose National Centre for Basic Sciences,
Block JD, Sector III, Salt Lake,
Kolkata-700106, West Bengal, India.*

Table of Contents

Chapter 1 Introduction	7
Chapter 2 Cloning and preliminary characterization of STY3178	13
2.1. Introduction.....	13
2.2. Materials and Methods.....	13
2.2.1. Cloning.....	13
2.2.2. Overexpression.....	14
2.2.3. Extraction and purification of the overexpressed protein	14
2.2.4. Mass determination	15
2.2.5. Circular dichroism.....	15
2.2.6. Steady state fluorescence	16
2.2.7. Dynamic light scattering	16
2.3. Results.....	16
2.3.1. Cloning, overexpression and purification	16
2.3.2. Mass analysis	17
2.3.3. Secondary structure.....	17
2.3.4. Tertiary structure.....	18
2.3.5. Protein hydrodynamic size.....	19
2.4. Discussion	20
2.5. Conclusion	20
Chapter 3 Determination of oligomeric state	21
3.1. Introduction.....	21
3.2. Materials and Methods.....	21
3.2.1. Size exclusion chromatography	21
3.2.2. Nuclear magnetic relaxation measurements.....	22
3.3. Results.....	22
3.3.1. Oligomeric State	22
3.3.2. Rotational correlation time.....	23
3.4. Discussion	24
3.5. Conclusion	25
Chapter 4 Thermal stability of STY3178	26
4.1. Introduction.....	26
4.2. Materials and Methods.....	26
4.2.1. Secondary structure at high temperatures	26
4.2.2. Dynamic Light Scattering	27
4.2.3. Kinetics of unfolding.....	27
4.2.4. Kinetics of refolding	28
4.2.5. Tertiary structure at elevated temperature.....	28
4.2.6. Nuclear Magnetic Resonance spectroscopy	29
4.3. Results.....	29
4.3.1. Unfolding with temperature	29
4.3.2. Reversibility in unfolding	31
4.3.3. Dependence of heating or cooling rates	34
4.3.4. Kinetics measurements.....	35
4.4. Conclusion	37

Chapter 5 Response of STY3178 in various chemical environments	38
5.1. Introduction	38
5.2. Methods.....	40
5.2.1. Sample preparation for experiments	40
5.2.2. Circular Dichroism measurements	41
5.2.3. Intrinsic fluorescence measurements	41
5.2.4. Nuclear Magnetic Resonance.....	42
5.2.5. Dynamic light scattering	42
5.2.6. Isothermal titration calorimetry.....	42
5.2.7. Size exclusion chromatography in SDS	43
5.3. Results	43
5.3.1. Effect of pH.....	43
5.3.2. Effect of urea.....	45
5.3.1. Effect of guanidium chloride	51
5.3.2. Effect of SDS	56
5.4. Discussion	60
5.5. Conclusion	61
Chapter 6 Function of STY3178	62
6.1. Introduction	62
6.2. Materials and Methods.....	63
6.2.1. Interaction with antibiotics.....	63
6.2.2. Chaperone activity assay	66
6.3. Results	67
6.3.1. Antibiotic binding to STY3178.....	67
6.3.2. Structural changes upon antibiotic binding.....	70
6.3.3. Oligomeric state of antibiotic bound protein.....	72
6.3.2. Chaperone activity assay of STY3178.....	74
6.4. Discussion	75
6.5. Conclusion	80
Appendices.....	80
Chapter 7 Homology modeling and molecular dynamics simulation of STY3178 ...	91
7.1. Introduction	91
7.2. Methodology	91
7.2.1. Molecular model of STY3178	91
7.2.2. Molecular dynamics simulation of the model	92
7.2.3. Molecular model of trimeric assembly	93
7.3. Results	94
7.3.1. Molecular modeling	94
7.3.2. Molecular dynamics simulation at elevated temperatures	95
7.3.3. Structure of trimer model	103
7.4. Conclusion	106
References	107
List of Publications	119

Chapter 1| Introduction

Proteins are molecular machines that participate in various cellular processes.^{1, 2} They perform variety of biochemical functions and hence significant for living cells. Bacterial proteins are of fundamental importance because they often interact with the host cells. They also help the bacteria to survive in harsh environments like unsuitable temperature, pH changes, presence of antibiotics and denaturing agents. Bacterial resistance towards antibiotics and their survival result in new strains which are threat to public health.³ Functional and structural characterization of bacterial proteins is thus, essential in order to impair them by understanding the mechanism of infection caused.

Proteins are synthesized in the ribosome as polypeptide chains.⁴ In order to be biologically active, the polypeptide chain is folded into a specific three dimensional structure.⁵ Native state of protein is reported to be only 5-10 kcal/mole more stable than the denatured one.⁶ Different types of interactions present in a protein structure are hydrogen bond, van der Waals attraction, electrostatic and hydrophobic interactions.⁶ Folding of a polypeptide chain into its compact tertiary structure is often initiated by the hydrophobic residues present in protein.^{1, 7} The hydrogen bonds between amide and carbonyl groups of peptide backbone are responsible for the stability of secondary structural elements.^{6, 8} Protein structure is often prone to perturbations like temperature shocks, pH changes, high pressure, denaturing agents such as chaotropes and acids resulting in unfolding.^{7, 9-14} The transition from native to unfolded state is generally associated with loss of tertiary as well as secondary structure. The folding-unfolding process of protein could either be reversible or irreversible.

Complete genome sequences of many organisms are available nowadays.^{15, 16} These have resulted in new protein sequences.^{15, 16} Structural genomics initiative has characterized the newly identified proteins. However, many of them are yet to be characterized either functionally or structurally. Bioinformatics studies have identified proteins present in bacteria without any functional annotation as domain of unknown function (DUF).¹⁷ These are reported in Pfam database with nearly 2700 DUFs for bacteria.¹⁸

yfdX¹⁹⁻²¹ protein is member of DUF family with Pfam database ID PF10938. They are known to occur in a number of virulent bacteria like *Escherichia coli*, *Salmonella* Typhi, *Salmonella* Typhimurium, *Salmonella* Paratyphi, *Klebsiella pneumoniae*, *Pantoea ananatis*, *Edwardsiella tarda*, *Hafnia alvei* and *Plesiomonas shigelloides*.²²⁻³¹ These proteins in different bacteria are nearly 190-220 amino acids long.³² It was first identified in *E. coli*.^{19, 20} The

expression of *yfdX* gene was reported to be up-regulated in *E. coli* by a multidrug response regulator EvgA. This EvgA protein also regulates the expression of many other genes like *yfdW*, *yfdU*, *yfdV*, *yfdE*, *yhiU*, *yhiV*, *osmC* and *ompA* encoding the respective proteins.^{19-21, 33} These genes are reported to be involved in high osmolarity²⁰, efflux transport mechanism³⁴ and acid tolerance responses (ATR)^{33, 35}. The expression of genes activated by EvgA were quantified by DNA microarray^{19, 21} analysis and qRT-PCR²⁰. Enhancement obtained using DNA microarray analysis in the expression of genes *yhiU*, *yhiV*, *yfdX*, *yfdW*, *yfdU*, *yfdV* and *yfdE* was 1.8-, 4.4-, 54.4-, 6.8-, 1.0-, 3.0- and 4.0- folds, respectively.^{19, 21} In qRT-PCR, enhancement in expression of genes *yfdX*, *yfdW*, *yfdU*, *yfdV* and *yfdE* was 1600, 1300, 890, 500 and 170, respectively.²⁰ These data show that the expression of *yfdX* is highest among all other genes activated by EvgA. *yfdX* is the only gene which gets co-expressed in *E. coli* cells when EvgA is expressed.

Structural and functional characterizations of *yfdX* proteins are inadequate till date. There is only one *yfdX* protein from *Klebsiella pneumoniae* for which crystal structure is deposited in PDB (3DZA). However, no functional annotation is reported even for this protein. Occurrence in virulent bacteria, enhanced expression and highest level of up-regulation with EvgA in *E. coli* indicates that *yfdX* proteins may have functional importance.

In this thesis, we have characterized a *yfdX* protein, STY3178 from the multidrug resistant (MDR) CT18^{23, 25} strain of *Salmonella* Typhi. We first clone the gene *sty3178* in pET28a expression vector and overexpress it in *E. coli* by isopropyl- β -D-thiogalactoside (IPTG) induction. The affinity tagged overexpressed protein is purified using column chromatography. We determine the purity of STY3178 using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE).³² We perform mass spectrometry (MALDI-TOF) and confirm the molecular mass of the overexpressed protein to be 23.1 KDa, which is similar to the calculated mass of construct.³²

Biophysical characterization of STY3178 protein is performed using various techniques such as circular dichroism (CD), steady state fluorescence, dynamic light scattering (DLS), size exclusion chromatography (SEC), nuclear magnetic resonance (NMR) and isothermal titration calorimetry (ITC). We have determined the secondary structure and oligomeric state of protein in solution. We monitor the thermal and chemical stability of STY3178. The functional importance of STY3178 is determined by studying its interaction with antibiotics such as ciprofloxacin (Cpx), rifampin (Rfp) and ampicillin (Amp), to which CT18 strain of *Salmonella* Typhi is resistant.³⁶⁻⁴⁰ Further, based on output results of bioinformatics tool⁴¹, we perform chaperone activity assay for STY3178 and investigate this capability *in vitro*. Finally, a homology model is

generated using modeling servers. We perform molecular dynamics (MD) simulation of the monomer and oligomer to check the stability.

CD spectrum of STY3178 confirms a folded protein with characteristic signature of α -helix having two minima around 209 nm and 222 nm.³² The emission peak position of protein using steady state fluorescence is found around 342 nm for excitation wavelengths 257, 275, 280 and 295 nm. This peak position indicates emission from exposed tryptophan residues and occurrence of FRET in protein.³²

The mean hydrodynamic diameter (D_H) of STY3178 is around 6.5 nm as measured using DLS. This hydrodynamic size is higher for a 23 KDa protein compared to the reported^{24, 25} size of globular proteins with similar molecular weights. To address this size anomaly, we perform size exclusion chromatography (SEC). STY3178 elutes at a volume similar to a 66 KDa protein in SEC.³² This indicates STY3178 to be a trimer (~69 KDa) in solution. We further perform one dimensional ^1H - ^{15}N heteronuclear relaxation measurements to determine the longitudinal (T_1) and transverse (T_2) relaxation time for the protein. T_1 (~1.96 s) and T_2 (~0.033 s) values are obtained from the slope of intensity decay plots for various time delays. The rotation correlation time (τ_c) estimated using T_1 and T_2 is about 24.7 ns following the expression reported in literature^{42, 43}. We also calculate the τ_c using the Stokes-Einstein-Debye (SED) equation for the measured hydrodynamic radius (~3.25 nm) of protein and obtain a value around 30.5 ns, which is in good agreement with the experimental τ_c . Therefore, NMR relaxation measurements, DLS and SEC data confirm a trimeric assembly for STY3178 in solution.

The thermal unfolding studies of STY3178 show its stability up to 50°C. It has the capability to undergo reversible thermal unfolding as well.⁴⁴ We perform ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) experiments at five different temperatures from 25°C to 45°C. The spectrum remains almost unchanged even at 45°C compared to HSQC at 25°C. This is in agreement with our CD and fluorescence results, supporting protein stability in this temperature range. Further, we check the reversibility in unfolding of STY3178 by heating it at different temperatures for 30 minutes and then cooling it back to room temperature. Protein is able to refold back completely from 80°C as monitored by CD. We also monitor the effect of heating and cooling rates on unfolding and refolding of protein by measuring θ_{222} for six different rates 0.5°C/min, 1°C/min, 2°C/min, 3°C/min, 4°C/min and 5°C/min in the temperature range 20°C to 70°C. The kinetics of unfolding and refolding of protein follow Arrhenius behaviour with activation energies of about 246.9 and -58.7 kJ/mol, respectively.⁴⁴

We monitor stability of protein at various pH. The α -helical secondary structure of STY3178 is found to be stable in the pH range 2.5-10. However, the tertiary structure is stable only in 4.5-10 pH and affected in strong acidic pH of 2.5 and 3.⁴⁵ Protein becomes more helical up to 0.6 M urea. It unfolds above 2 M urea as observed in CD, fluorescence and NMR. In presence of GndCl, unfolding of protein occurs above 0.6 M concentration. We also measure the HSQC spectra of singly labeled STY3178 and find the structure to unfold at 3 M urea. Native protein-like dispersion of HSQC is observed upon reducing 3 M urea concentration by buffer exchange. Overall we find that the chaotrope induced unfolding of STY3178 is reversible in nature.⁴⁵

In presence of SDS, the secondary structure of protein changes slightly with increase in ellipticity at 209 nm and decrease at 222 nm.⁴⁶ However, the secondary structure does not show signature of an unfold protein as observed in presence of urea and GndHCl. The near UV-CD spectra, however, changes abruptly even at lowest concentration of SDS. This indicates that SDS affects protein tertiary or quaternary structure; although the α -helical nature is retained.⁴⁶ In DLS, we observe that the hydrodynamic size of protein decreases with increase in SDS concentration and D_H of about 4.2 nm is obtained for 69.4 mM (2 %) SDS. This lowering of protein size probably indicates dissociation of trimer into monomeric subunits. We confirm the dissociation of trimeric STY3178 into its monomer by performing SEC in presence of SDS. The native-like hydrodynamic size is regained upon decreasing SDS concentration in solution, indicating formation of trimer from the monomers. This dissociation of trimeric assembly in presence of SDS is also a reversible phenomenon.

We show the interaction of STY3178 with three antibiotics ciprofloxacin (CpX), rifampin (Rfp) and ampicillin (Amp) using steady state fluorescence and ITC.³² The fluorescence emission is substantially quenched in presence of all the three antibiotics at 280 nm excitation. This interaction is quantified by calculating the binding parameters using Stern- Volmer equation⁴⁷⁻⁴⁹. Dissociation constants (K_d) estimated for Cpx, Rfp and Amp binding are 25 μ M, 100 μ M and 2.5 mM, respectively. The number of binding sites is 0.99, 1.5 and 1.03 for Cpx, Rfp and Amp, respectively. We further verify these antibiotic interactions using ITC. The isotherms for Cpx and Amp are fitted to sequential binding model and Rfp binding data is fitted to a single site model in order to estimate the binding parameters. The dissociation constants (K_d) obtained by fitting the isotherm for Cpx is 1.72 μ M and 2.36 μ M. K_d for Rfp and Amp binding are 66.23 μ M and 134.4 μ M, respectively. Both fluorescence and ITC measurements confirm that Cpx binding with STY3178 is stronger than Rfp followed by Amp. The α -helical secondary

structure of protein remains intact upon antibiotic binding, measured using far UV-CD. Native protein has a broad shoulder spanning over 250-280 nm in near UV-CD spectra. This decomposes in presence of Cpx and Rfp, into two new peaks around 260 nm and 280 nm, which commensurate to phenylalanine and tyrosine fine structures^{50, 51}, respectively. DLS and NMR relaxation measurements indicate that STY3178 remains as trimer during interaction with the antibiotics.

We employ Bioinformatics tool MODexplorer⁴¹ for further functional characterization of STY3178. The results from this server include about 50 similar proteins, having >50 % structural similarity with STY3178. Most of these proteins contain helices similar to STY3178. Majority of these proteins have reported functions as chaperones or assist chaperone activity.⁴⁴ In order to verify this, we perform chaperone activity assay by monitoring thermal aggregation of two substrate proteins, alcohol dehydrogenase and insulin. The assay is carried out at 42°C by measuring the light scattering (A_{360}) at 360 nm in absence and presence of STY3178. We observe lowering of A_{360} for alcohol dehydrogenase and DTT-induced aggregation of insulin with increasing concentration of STY3178. This demonstrates that STY3178 is capable of preventing aggregation of both the substrate proteins and hence, acting as a chaperone.

Finally, we propose a model for STY3178 by homology modeling^{52, 53} using a template yfdX protein (PDB ID 3DZA) from *Klebsiella pneumoniae*.⁴⁴ Servers Phyre2⁵⁴ and RaptorX⁵⁵⁻⁵⁸ generate models containing residues A¹¹-Q¹⁸³ only. The rest of the C-terminal residues from S¹⁸⁴-H¹⁹⁹, are added de novo using Swiss-PdbViewer⁵⁹. The final model having residues A¹¹-H¹⁹⁹, is solvated using explicit solvent model and minimized by NAMD⁶⁰ with CHARMM27⁶¹ force field and TIP3P⁶² model of water after maintaining electro neutrality. The monomer model is simulated for 300 ns. It consists of ten α -helices and a β -sheet comprising of two strands. The simulated 310 K structure is used to perform MD simulation at higher temperatures 350 K and 400 K for determining its thermal stability.⁴⁴ The root mean square deviation (RMSD) at all temperatures is calculated with respect to C $_{\alpha}$ -atoms over the trajectories up to 300 ns of MD simulation. We estimate the root means square fluctuation (RMSF) of C $_{\alpha}$ -atoms of each residue at temperatures 310 K, 350 K and 400 K over the equilibrated trajectories 150-300 ns. Finally, the relative change in backbone dihedral angle fluctuation for each residue at 350 K ($r_{350} = \sigma_{350} / \sigma_{310}$) and 400 K ($r_{400} = \sigma_{400} / \sigma_{310}$) is estimated with respect to that at 310 K. The model structure at 350 K and 400 K do not show substantial unfolding though there are regions with increased fluctuations in backbone dihedral angles. This indicates that the model structure of STY3178 is thermally stable.

We experimentally find that STY3178 exists as a trimer in solution.³² Thus, we generate possible models for the trimeric structure of protein using 300 ns MD simulated monomer structure in ZDOCK⁶³ docking server. Energetically most favourable structure of oligomeric assembly is determined by minimizing the possible trimeric assemblies using NAMD⁶⁰ with CHARMM27⁶¹ force field and TIP3P⁶² water model. The most stable trimer is subjected to MD simulation. The radius of gyration⁶⁴ of simulated trimer model is calculated to be about 2.69 nm. We then estimate the hydrodynamic diameter of around 6.98 nm for this calculated radius of gyration, using the relation reported in literature⁶⁵. This is in good agreement with hydrodynamic diameter (~6.5 nm) determined using DLS. The proposed model of trimer has tryptophan and tyrosine pairs in near vicinity, which can explain the phenomenon of FRET observed in steady state fluorescence. The trimer interface consists of hydrophobic and electrostatic interactions.

Our study on protein STY3178 has revealed wide variety of interesting properties along with its sturdiness and important binding capabilities. Hence, our work opens up a new area of research on yfdX proteins, which is yet to be deciphered and might be relevant to understand the virulent bacteria. This thesis is organized into the following chapters: Chapter 2 covers the cloning of gene *sty3178*, its overexpression, protein purification and preliminary biophysical characterization. Chapter 3 addresses the anomaly in hydrodynamic size of STY3178. We determine the state of protein oligomeric assembly. In Chapter 4, we show thermal stability and reversible unfolding of protein. Chapter 5 shows the behaviour of STY3178 in various chemical environments. We shed light on the functional relevance of STY3178 using biophysical methods in Chapter 6. Finally in Chapter 7, we propose a model of STY3178 monomer based on homology modeling and a possible structure for the trimeric assembly.

Chapter 2| Cloning and preliminary characterization of STY3178

2.1. Introduction

Protein STY3178, a member of yfdX¹⁹⁻²¹ family is present in CT18^{23, 25} strain of *Salmonella* Typhi (*S. Typhi*)³². This protein is encoded by *sty3178* gene in *S. Typhi*^{23, 25}. Gene bank ID of *sty3178* gene including the signal peptide region is gi|16758993:c3049965-3049366. Its ID in the Uniprot⁶⁶ database is Q8Z3Y9. This gene is 600 base pairs long. We clone this desired gene without the signal peptide region having 573 base pairs in *E. coli* in order to overexpress and characterize the protein STY3178. We employ different techniques⁶⁷ of molecular biology for protein expression and purification using affinity chromatography. Preliminary biophysical characterization of the protein is done after extraction and purification. At first we run SDS-PAGE to determine the purity of STY3178. Our next measurement includes mass analysis of the overexpressed protein using mass spectrometry (MALDI-TOF). We then perform circular dichroism (CD) measurement in far UV (195-250 nm) region to determine the secondary structural elements present in STY3178. CD measurement in the near UV (250-330 nm) region is also performed to get information about the fine structure of aromatic residues present in the protein. We determine the nature of the fluorophores using steady state fluorescence measurements. The hydrodynamic size of the protein is determined using dynamic light scattering (DLS).

This chapter is organized as follows: Section 2.2 covers the experimental methods related to cloning, overexpression, extraction, purification and biophysical characterization of STY3178. In section 2.3, results of cloning, overexpression, purification and preliminary characterization of the protein are stated. Sections 2.4 and 2.5 would cover the discussion and conclusion, respectively, from the experimental findings.

2.2. Materials and methods

2.2.1. Cloning

Gene *sty3178* with 573 base pairs (without the signal peptide region) encoding our protein of interest is amplified by polymerase chain reaction (PCR) using genomic DNA of *Salmonella* Typhi (CT18 strain) as a template. The forward and reverse primers used during the PCR are: 5'-CATATGGCCGCAACAAACATGACTG-3'; and 5'-CTCGAGGATATTAATGCGCGGCGTCGTG-3' (Integrated DNA Technologies), respectively.

Both the primers contain restriction sites for enzymes NdeI (CATATG) and XhoI (CTCGAG). The amplified PCR product is used to run agarose gel (1 %) and the desired product is extracted by gel purification using Qiagen kit. The purified product is then inserted into pTZ57R/T (TA) vector (Fermentas) using T4 DNA ligase enzyme. The insert is transformed into Top 10 *E. coli* cells (Novagen) and incubated overnight at 37°C in Luria-agar plate containing ampicillin (0.286 mM). The *E. coli* cells with the plasmid containing the desired gene are selected by performing blues/ white screening using 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (Xgal) as a marker. Plasmid is then isolated from the selected white colonies using plasmid isolation protocol (Qiagen). The isolated plasmid is digested with NdeI and XhoI. The digested plasmid is run on agarose gel (1 %) and purified by gel purification method (Qiagen). The expression vector pET28a is also digested with the same restriction enzymes NdeI and XhoI. The purified digested product is sub-cloned into the pET28a expression vector using T4 DNA ligase. Cloning of the desired gene is confirmed by sequencing using T7 primers specific to pET28a system (Applied Biosystems). The plasmid carrying the desired gene is then transformed into *E. coli* BL21 (DE3) cells (Novagen) and incubated in Luria agar plate containing 0.124 mM kanamycin.

2.2.2. Overexpression

E. coli BL21 (DE3) cells with the plasmid pET28a carrying the gene *sty3178* are grown overnight in 10 ml Luria-Bertani (LB) media containing kanamycin (0.124 mM) at 37°C in a shaker incubator (Innova 42 New Brunswick Scientific) at 250 rpm. This overnight pre-culture is used to inoculate 1 liter of fresh LB media and the cells are grown till 0.9 optical density (OD₆₀₀). Cells are then induced with 0.2 mM isopropyl- β -D-thiogalactoside (IPTG) for 4 hrs keeping other conditions same. After 4 hrs of induction, cells are harvested by spinning down in a centrifuge (Eppendorf) at 5000 g for 10 minutes at 4°C. Overexpression of protein STY3178 is confirmed by running 12 % sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE).

2.2.3. Extraction and purification of the overexpressed protein

Harvested cells are lysed by sonication (Sartorius LABSONIC) at 30 % amplitude and 0.7 cycle in ice-bath after resuspension in a lysis buffer containing 50 mM potassium phosphate (pH 7), 250 mM sodium chloride (NaCl) and 1 mM phenylmethanesulfonyl fluoride (PMSF). The crude cell lysate is then subjected to centrifugation at 14000 g for 10 minutes at 4°C. The

desired protein is obtained in the supernatant after centrifugation as observed in SDS-PAGE. Purification of the protein is performed by affinity chromatography using nickel-nitrilotriacetic acid (Ni-NTA) beads (Qiagen). The supernatant is allowed to bind the Ni-NTA beads. Non-specific proteins bound to the Ni-NTA beads are removed by washing with a buffer containing 50 mM potassium phosphate (pH 7), 250 mM NaCl, 1 mM PMSF and 30 mM imidazole. Finally, the desired protein is eluted using an elution buffer containing 50 mM potassium phosphate (pH 7), 250 mM NaCl, 1 mM PMSF and 250 mM imidazole. Purity of the recombinant protein is determined by running SDS-PAGE. Amicon spin concentrator (10 KDa cut off) is used to remove imidazole from the purified protein by exchanging the elution buffer with that containing 50 mM potassium phosphate (pH 7), 250 mM NaCl and 1 mM PMSF. Concentration of the protein is determined using the absorbance value at 280 nm (BioSpectrometer, Eppendorf) and applying Beer-Lambert law. Extinction coefficient value (ϵ_{280}) required for concentration determination is estimated from ProtParam⁶⁸ tool of expasy server using the construct sequence as an input. The obtained value of ϵ_{280} is $18450 \text{ M}^{-1}\text{cm}^{-1}$.

2.2.4. Mass determination

Mass analysis of STY3178 is performed in MALDI-TOF Bruker Ultraflex extreme spectrometer. The matrix used during the measurement is sinapinic acid, dissolved in a mixture of acetonitrile and trifluoroacetic acid (1:1). STY3178 is mixed in 1:1 ratio with the dissolved matrix sinapinic acid prior to measurement.

2.2.5. Circular Dichroism (CD)

The CD measurement is carried out in Jasco J-815 CD spectrometer at 20°C in the far UV (195-250 nm) and near UV (250-330 nm) regions. The path lengths of the cells used in far UV and near UV measurements are 1 mm and 10 mm, respectively. The respective protein concentrations used in far UV- and near UV-CD are 10 μM and 30 μM . The reported data are an average of three scans and background corrected by subtracting the corresponding spectrum of the buffer containing 50 mM potassium phosphate (pH 7), 250 mM NaCl and 1 mM PMSF.

The ellipticity value at 222 nm of the CD data (far UV-region) is used to determine the helical content of the protein following the relation^{51, 69}, $f_H = \frac{([\theta]_{222} - 3,000)}{(-36,000 - 3,000)}$, (1)

where f_H and $[\theta]_{222}$ are fractional helicity and mean residue ellipticity at 222 nm wavelength.

2.2.6. Steady state fluorescence measurement

The fluorescence measurements are performed using Jobin Yvon Horiba Fluorolog with entrance and exit slit widths of 2 nm. Protein concentration of 10 μM is used. The sample is excited at 257, 275, 280 and 295 nm wavelengths. The final spectra for the protein at each excitation wavelength are obtained by subtracting the corresponding buffer containing 50 mM potassium phosphate (pH 7), 250 mM NaCl and 1 mM PMSF. Each spectrum is averaged over two sets of measurements.

2.2.7. Dynamic light scattering (DLS)

DLS measurement is performed in Nano-S Malvern instrument at 20°C. 10 μM protein is subjected to laser scattering with wavelength 632.8 nm and measuring angle 173°. Measurement is obtained as a mean of five successive counts. The sample is filtered through 0.22 μm syringe filter (Millipore) prior to measurement.

2.3. Results

2.3.1. Cloning, overexpression and purification

We confirm cloning of the gene *sty3178* without the N-terminal signal peptide in pET28a expression vector by sequencing (Applied Biosystems). The clone contains six histidine (His) tags in the N-terminal region for the purpose of purification by affinity chromatography. Overexpression of the recombinant protein STY3178 in *E. coli* BL21 (DE3) cells is induced by 0.2 mM IPTG and confirmed by running 12 % SDS-PAGE. In Figure 2.1, lane 2 shows the un-induced cells before adding IPTG. The overexpressed protein band is observed around 25 KDa in SDS-PAGE (Figure 2.1, lane 3). However, the estimated molecular weight of the construct protein sequence is about 23.11 KDa. Cells with the overexpressed protein are resuspended in a buffer with composition 50 mM potassium phosphate (pH 7), 250 mM NaCl and 1 mM PMSF. The resuspended cells are then lysed by sonication as mentioned in method section. Protein STY3178 is obtained in the supernatant after extraction (Figure 2.1, lane 4). The extracted protein is purified by affinity chromatography using Ni-NTA beads. The beads are washed with a buffer having 50 mM potassium phosphate (pH 7), 250 mM NaCl, 1 mM PMSF and 30 mM imidazole. Finally, STY3178 is eluted with a buffer with composition 50 mM potassium phosphate (pH 7), 250 mM NaCl, 1 mM PMSF and 250 mM imidazole. Purity of STY3178 is confirmed by SDS-PAGE (Figure 2.1, lane 8).

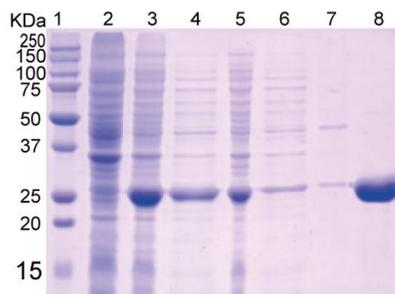


Figure 2.1. Overexpression, extraction and purification of STY3178. SDS-PAGE representing molecular weight marker in lane 1; un-induced cells, lane 2; induced cells in lane 3; crude extract of soluble proteins, lane 4; insoluble cell debris, lane 5; flow through from Ni-NTA affinity column, lane 6; wash from Ni-NTA affinity column, lane 7; elute from affinity column, lane 8.

2.3.2. Mass analysis

We perform MALDI-TOF to determine the exact molecular mass of the purified protein. The mass spectrum contains two peaks. The first peak has m/z ratio of 11.5 KDa for the doubly charged species and the second peak at 23.1 KDa for the singly charged one (Figure 2.2). This confirms the molecular weight of STY3178 to be 23.1 KDa, although protein migrates around 25 KDa in SDS-PAGE (Figure 2.1).

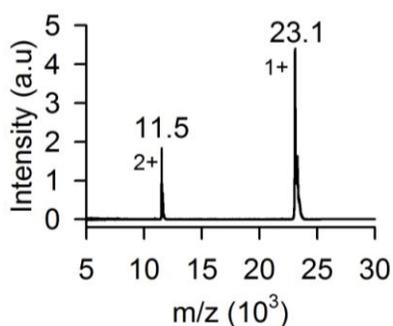


Figure 2.2. Mass spectrum of STY3178. MALDI-TOF spectrum of pure protein with peaks at 23.1 KDa for the singly charged species and 11.5 KDa for the doubly charged one.

2.3.3. Secondary structure

The overexpressed and purified STY3178 is found to be well folded protein as observed in far UV (195-250 nm) CD (Figure 2.3). The spectrum contains two minima around 209 nm and 222 nm, which is a characteristic of α -helical proteins (Figure 2.3). The percentage of helix is estimated using the ellipticity value at 222 nm (θ_{222}) following the equation^{26, 27},

$f_H = \frac{([\theta]_{222} - 3,000)}{(-36,000 - 3,000)}$. The helical content obtained for the protein is ~43.9 %. We use the

CD data (200-240 nm) of the protein in Dichroweb server⁷⁰⁻⁷² and apply K2D⁷³ method to estimate the percentage of secondary structural elements. The estimated α -helix and β -sheet content obtained from Dichroweb server analysis are around 50 % and 20 %, respectively. The helix content estimated using both the analysis methods is quite comparable.

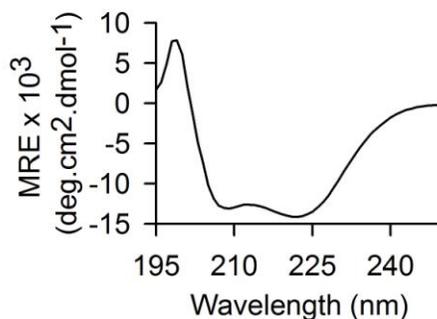


Figure 2.3. Secondary structure of STY3178. Far UV (195-250 nm) CD spectrum of protein showing α -helical nature with two characteristic minima around 209 nm and 222 nm.

2.3.4. Tertiary structure

We use steady state fluorescence measurement to study the nature of the fluorophore (phenylalanine, tyrosine and tryptophan) environment in STY3178. STY3178 is excited at four different wavelengths 257, 275, 280 and 295 nm at room temperature. The excitation wavelengths used will solely excite the aromatic residues phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp). The emission peak position for all these four excitations are found around 342 nm (Figure 2.4a). This corresponds to emission from exposed tryptophan (Trp) residue (s). The primary sequence of protein contains two Phe, five Tyr and two Trp residues. However, the fluorescence emission for excitation at 257 nm and 275 nm are similar to that of Trp. This suggests a possibility of Förster resonance energy transfer (FRET) between Phe-Trp and Tyr-Trp pairs present in the protein. The difference spectra for (275-295) nm and (257-295) nm excitation wavelengths show signature of FRET between Tyr-Trp (Figure 2.4a, blue) and Phe-Trp (Figure 2.4a, grey) pairs, respectively. The intensity of difference spectrum for (257-295) nm is less than that of (275-295) nm. This might be due to low quantum yield of Phe compared to Tyr or Trp. All the three aromatic amino acid residues could get excited at 257 nm and the emission wavelength (280 nm) of Phe overlaps with excitation wavelengths of Trp and Tyr, which also causes lowering of emission for (257-295) nm difference spectrum.

We further monitor the protein tertiary structure in near UV (250-330 nm) region using CD measurements. The near UV-CD spectrum of the native protein contains a broad shoulder spanning from 250 nm to 280 nm (Figure 2.4b). This broad shoulder is a signature of aromatic residues Phe, Tyr and Trp present in the protein.

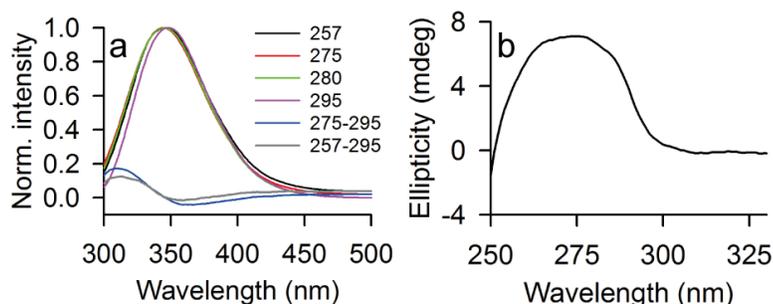


Figure 2.4. Tertiary structure of STY3178. (a) Steady state fluorescence emission spectra for excitation wavelengths 257 (black), 275 nm (red), 280 (green) and 295 nm (magenta) are shown. The difference fluorescence spectra for (275-295) nm and (257-295) nm are represented in blue and grey, respectively. (b) Shows the near UV (250-330 nm) CD signature of protein with broad shoulder in the region 250-280 nm.

2.3.5. Protein hydrodynamic size

The hydrodynamic size of STY3178 is measured using dynamic light scattering. The value of hydrodynamic diameter obtained for the protein in solution is about 6.5 nm (Figure 2.5).

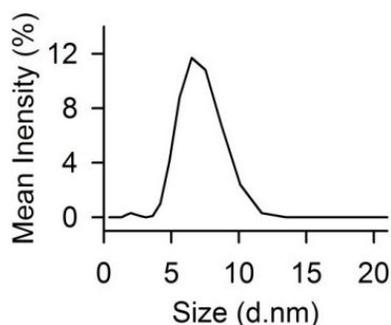


Figure 2.5. DLS spectrum of STY3178 showing hydrodynamic diameter around 6.5 nm.

2.4. Discussion

STY3178 is predicted to be a hypothetical protein by Kegg⁷⁴ and STRING⁷⁵ databases. On the other hand, it is predicted as a putative membrane protein by Topsisan⁷⁶ database whereas Uniprot⁶⁶ and NCB mention it a periplasmic protein. We find STY3178 in the supernatant after extraction which confirms it to be a soluble protein. Our observation is also in agreement with previous study on *E. coli* yfdX¹⁹ protein where it was found in the soluble cytoplasmic fraction.

The hydrodynamic size of this protein measured using DLS is found to be higher compared to proteins of similar molecular weights.³² It shows an anomaly considering the relation between hydrodynamic size and molecular weight (MW) of different standard globular proteins.⁷⁷ For example, the reported hydrodynamic radii of soyabean trypsin inhibitor (20 KDa) and carbonic anhydrase (29 KDa) are 2.4 nm and that of ovalbumin (45 KDa) is 2.8 nm.^{78, 79} This suggests that a hydrodynamic radius of 3.25 nm for STY3178 would correspond to molecular weight higher than 50 KDa.

2.5. Conclusion

yfdX protein, STY3178 from multidrug resistant CT18 strain of *S. Typhi* is characterized using different biophysical techniques. This protein is majorly α -helical in nature along with some β -sheet content. The fluorescence measurements indicate emission from exposed tryptophan residue (s). Protein mass and hydrodynamic size is determined using MALDI-TOF and DLS, respectively. The hydrodynamic diameter of the protein shows a size anomaly, which is addressed in the next chapter.

Chapter 3| Determination of oligomeric state

3.1. Introduction

In the previous chapter, we report preliminary characterization of protein STY3178. Dynamic light scattering (DLS) indicates an anomaly in hydrodynamic size of STY3178 compared to its molecular mass. To address this, size exclusion chromatography (SEC) is performed for STY3178. SEC can provide the information regarding the monomeric or oligomeric state of the protein. We apply Stokes-Einstein-Debye (SED) equation, $\tau_c = 4\pi\eta R^3 / 3KT$ (where η and R are the viscosity and hydrodynamic radius, respectively) to estimate the rotation correlation time (τ_c) of STY3178 for the measured hydrodynamic radius (R). The SED equation correlates the hydrodynamic radius (R) with the rotation correlation time (τ_c) of the molecule. We further carry out longitudinal and transverse relaxation measurements using nuclear magnetic resonance (NMR) to determine the rotation correlation time (τ_c) of STY3178.

The organization of this chapter is as follows: Section 3.2 includes the experimental details employed for determination of protein oligomeric state. In section 3.3, results of the experiments are stated. Section 3.4 includes the discussion regarding the possible oligomeric assemblies of the protein. Finally, we conclude in section 3.5 stating the overall finding.

3.2. Materials and methods

3.2.1. Size exclusion chromatography (SEC)

Size exclusion chromatography is performed using a glass column (2.5 × 45 cm, Sigma) packed with superdex75 media (GE Healthcare). The packed column is equilibrated with a buffer containing 50 mM potassium phosphate (pH 7), 250 mM NaCl, 1 mM PMSF prior to loading protein. 500 μM protein is subjected to the pre-equilibrated column. Fractions are collected for every 3 ml volume using a peristaltic pump (GE Healthcare) with a flow rate of 1 ml/min. Absorbance at 280 nm (A_{280}) are measured for all the fractions in order to identify the desired protein. The column is calibrated using standard proteins namely, lysozyme (14.4 KDa), carbonic anhydrase (29 KDa), ovalbumin (43 KDa), BSA (66 KDa) and conalbumin (75 KDa). SEC measurements are performed three times for all the proteins and mean of these data are reported.

3.2.2. Nuclear magnetic resonance (NMR): Relaxation measurements

Uniformly ^{15}N -labeled STY3178 is prepared by growing the *E. coli* cells carrying the plasmid pET28a with the gene *sty3178* in M9 minimal media supplemented with $^{15}\text{NH}_4\text{Cl}$ as a source of nitrogen. The other conditions of the cell culture are kept same as reported in chapter 2. The extraction and purification protocol followed is same as mentioned in purification section of chapter 2.

The one dimensional (1D) ^1H - ^{15}N heteronuclear longitudinal (T_1) and transverse (T_2) relaxation experiments are carried out using 600 MHz Varian spectrometer equipped with triple resonance probe at 30°C . The final concentration of ^{15}N -labeled protein used during the measurement is $355\ \mu\text{M}$ in a buffer having 30 mM phosphate (pH 7), 150 mM NaCl and 10% D_2O . The recycle delays (d_1) of 8, 9 and 12 seconds are used for T_1 measurements. The free induction decay (FID) is collected for the delay points 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 1.0, 1.2, 1.5 and 1.8 seconds in T_1 measurement. Similarly, d_1 used for T_2 measurement is 4 seconds. The delay points used in FID collection for T_2 are 0.01, 0.03, 0.05, 0.07, 0.09, 0.11, 0.13, 0.15 and 0.17 seconds. All the FIDs are acquired for 256 scans. The intensity in the range 8.5-10.5 ppm is obtained by processing the data using VnmrJ for each set of delay. The intensities for all the d_1 time delays for both T_1 and T_2 measurements are first normalized for each individual data set. The normalized intensities are then averaged over all the data sets acquired at each delay point. The plot of intensity versus time is fitted to a single exponential equation, $I = a * \exp(-Rt)$ using SigmaPlot, where I denote the normalized intensity. R and t are rate of relaxation and delay time, respectively. T_1 and T_2 values are calculated as a reciprocal of respective R values. The total rotational correlation time (τ_c) is calculated using the relation^{42,43} given below-

$$\tau_c = \frac{1}{4\pi V_N} \sqrt{6 \frac{T_1}{T_2} - 7}, \quad (1)$$

where V_N is nitrogen (^{15}N) resonance frequency in Hertz.

3.3. Results

3.3.1. Oligomeric State

We perform size exclusion chromatography (SEC) to determine the oligomeric state of STY3178. It is carried out using Superdex75 media (GE healthcare) packed in a glass column (Sigma). The chromatogram for STY3178 elution from the column is shown in Figure 3.1a.

The elution volume (~120 ml) is compared with many standard proteins in the molecular weight range of 14.4 KDa to 75 KDa (Figure 3.1b). These standard proteins elute from the column at similar volumes for repeated experiments as shown by small error bars (Figure 3.1b). STY3178 elution volume matches in the molecular weight regime of 66 KDa protein BSA. This suggests that STY3178 exists as a trimer in solution with a molecular mass ~69 KDa.

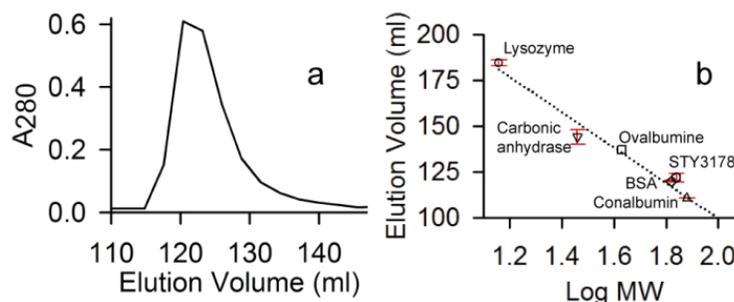


Figure 3.1. Oligomeric state from SEC. (a) The size exclusion chromatogram of STY3178 showing elution around 120 ml in superdex75 column. (b) The elution volume versus logarithm of molecular weight of proteins lysozyme (14.4 KDa, circle); carbonic anhydrase (29 KDa, inverted triangle); ovalbumin (43 KDa, square); BSA (66 KDa, diamond); STY3178 (hexagon) and Conalbumin (75 KDa, triangle). The error bars (red) are the standard errors estimated using the average elution volume for the repeated experiments.

3.3.2. Rotation correlation time (τ_c) from NMR relaxation measurements

We experimentally estimate τ_c of STY3178 using NMR spectroscopy by acquiring data for one dimensional (1D) ^1H - ^{15}N heteronuclear longitudinal (T_1) and transverse (T_2) relaxation measurements. We perform T_1 measurement for three different d_1 delays since T_1 relaxation times are typically longer than T_2 relaxation times. The normalized peak intensities are averaged over the repeated sets of measurements. The average peak intensities in 8.5-10.5 ppm region of the spectra plotted as a function of corresponding delay times for both T_1 and T_2 experiments (Figures 3.2a,b). The T_1 and T_2 values are obtained from the integrated peak intensities by fitting it to exponential decay curves. T_1 and T_2 values obtained by fitting the intensity decay curves are about 1.95 and 0.033 seconds, respectively. The rotational correlation time τ_c calculated using these T_1 and T_2 values in equation^{54, 55},

$$\tau_c = \frac{1}{4\pi V_N} \sqrt{6 \frac{T_1}{T_2} - 7} \text{ is about } 24.7 \text{ ns.}$$

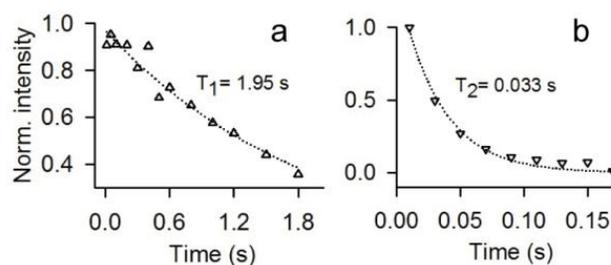


Figure 3.2. NMR relaxation measurement for ¹⁵N labelled STY3178. (a) and (b) Show the plot of normalized intensities with delay times for T₁ and T₂ relaxation measurements. The delay times used for T₁ measurements are 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 1.0, 1.2, 1.5, 1.8 s and for T₂ measurements are 0.01, 0.03, 0.05, 0.07, 0.09, 0.11, 0.13, 0.15, 0.17 s. The corresponding T₁ and T₂ values are extracted by fitting the curved with single exponential decay equation.

3.4. Discussion

We estimate the rotational correlation time (τ_c) of STY3178 using Stokes-Einstein-Debye (SED) equation, $\tau_c = 4\pi\eta R^3 / 3KT$ (where η and R are the respective viscosity and hydrodynamic radius) for the experimentally determined hydrodynamic radius (~3.25 nm). The τ_c of about 30.5 ns is obtained using this relation. This value is much higher compared to the protein of similar molecular weight such as α -chymotrypsin (25 KDa), where the reported⁸⁰ τ_c is 13.4 ns. The experimentally determined τ_c from NMR relaxation measurements is in good agreement with the calculated τ_c for the measured hydrodynamic radius using the SED equation. Thus, all these experimental observations for the protein STY3178 confirm its trimeric state of oligomerization in solution.

As reported in chapter 2 that STY3178 is majorly α -helical protein with some β -sheet as secondary structural elements. Similar observation is found for secondary structural elements of an orthologues yfdX protein from *Klebsiella pneumoniae* (PDB 3DZA). The experiments also suggest that STY3178 exist as an oligomer in solution similar to 3DZA. However, STY3178 is a trimer, whereas 3DZA forms a homotetramer. Any possibility of stable tetrameric aggregation state for STY3178 is ruled out based on our experimental observations. The molecular weight (MW) of STY3178 tetramer would be 92 KDa, had it been a stable tetramer. The hydrodynamic radius (R_H) of proteins with MW higher than 90 KDa is compared in Table

3.1 and the reported values are found to be larger than 3.25 nm. In SEC calibration curve (Figure 3.1b), the elution volume (~120 ml) for STY3178 would be a mismatch point from the standard protein plot if it is a tetramer. These observations suggest that yfdX proteins seem to have a propensity for the formation of oligomeric states but the degree of oligomerization may vary within the family. Similar feature is observed for other oligomeric proteins such as family of small heat shock proteins⁸¹.

Table 3.1. Hydrodynamic size^a (R_H) of different proteins in various molecular weights regimes

Protein	Molecular weight (KDa)	R_H (nm)
Myoglobin	19.5	2.12
Chymotrypsinogen	25.0	2.4
Trypsin inhibitor	28.0	2.47
Carbonic anhydrase	29.0	2.6
Haemoglobin	65.0	3.30
Bovine serum albumin (BSA)	67.0	3.55
Alcohol dehydrogenase	80.0	3.7
Transferrin	87.1	4.0
Amyloglucosidase	99.0	3.9
Hexokinase	102.0	4.3

^a reported in www.malvern.co.uk

3.5. Conclusion

In this chapter, we resolve the anomaly in hydrodynamic size and molecular weight of STY3178. We reveal its trimeric assembly formation in solution state using size exclusion chromatography. The SEC results are further supported by the rotational correlation time determined using NMR relaxation methods. Our present study indicates that yfdX proteins have a propensity to form oligomeric assemblies although their state of oligomers may vary.

Chapter 4| Thermal stability of STY3178

4.1. Introduction

The thermal response for different proteins is diverse. Thermal stability of a protein could shed light towards understanding its function. There are proteins which can tolerate high temperatures and undergo reversible unfolding whereas others form irreversible aggregates upon heating. For example, polymerases, like Taq⁸² or pfu⁸³ from thermophilic bacteria show very high temperature stability. Proteins showing reversible unfolding includes lysozyme^{84, 85}, chymotrypsinogen⁸⁶, soybean kunitz trypsin inhibitor⁸⁷, bovine pancreatic trypsin inhibitor^{88, 89}, porcine S100A12⁹⁰ (calcium binding protein), HSA^{91, 92} and triosephosphate isomerase⁹³.

In this chapter, we study the thermal stability of a yfdX protein, STY3178 from *Salmonella* Typhi. The primary objective is to shed light on its function by characterizing thermal unfolding profile. We show the thermal stability and reversibility in unfolding of STY3178 using circular dichroism (CD), steady state fluorescence, dynamic light scattering (DLS) and nuclear magnetic resonance (NMR) experiments. We also monitor the effect of various heating and cooling rates on protein folding. The activation energies are estimated by measuring the kinetics of unfolding and refolding using CD.

The organization of this chapter is as follows: The experimental methods for studying the thermal unfolding are discussed in section 4.2. The results related to the experiments are given in section 4.3 and conclusion is covered in section 4.4.

4.2. Materials and methods

4.2.1. Secondary structure at different temperatures

All the circular dichroism (CD) measurements are performed in Jasco J-815 CD spectrometer attached with a peltier unit (Jasco) for controlling the temperature. Quartz cell of 3 mm path length is used during the measurement. The data are recorded in the far UV-region (200-250 nm) using protein concentration of 10 μ M. Protein unfolding is monitored by increasing the temperature with increments of 10°C in the range of 20°C to 70°C. Sample is equilibrated for 10 minutes at each temperature (30°C, 40°C, 50°C, 60°C and 70°C) prior to measurement. Similarly, protein refolding is achieved by decreasing the temperature from 70°C to 20°C with decrements of 10°C (60°C, 50°C, 40°C, 30°C, and 20°C). The experiments are performed in a buffer with composition 50 mM potassium phosphate (pH 7), 250 mM NaCl

and 1 mM PMSF. All the spectra are background corrected by subtracting the data for the buffer at each temperature.

We monitor the transition from folded to unfolded state by measuring the ellipticity value at 222 nm (θ_{222}) upon gradually heating the protein from 20°C to 70°C. Different heating rates used for probing the transition are 0.5°C/min, 1°C/min, 2°C/min, 3°C/min, 4°C/min and 5°C/min. Similarly, six different cooling rates (0.5°C/min, 1°C/min, 2°C/min, 3°C/min, 4°C/min and 5°C/min) are used for refolding the protein. The spectra of each sample are background corrected by subtracting the data of the buffer collected for each rate of heating or cooling. The fraction of the folded protein (f_N) at each temperature is calculated using the relation⁹⁴, $f_N = (\theta_D - \theta_T)/(\theta_D - \theta_N)$, (1)

where θ_N and θ_D are the respective measured ellipticity values of folded and unfolded protein. θ_T is ellipticity at each temperature.

In another set of experiment, we heat STY3178 (10 μ M) at various temperature 50°C, 60°C, 70°C, 80°C, 90°C and 100°C for 30 minutes in a water bath (Labnet). The unfolded protein samples are allowed to cool down to room temperature. Far UV-CD is recorded with the cooled samples in order to monitor the refolding.

4.2.2. Dynamic Light Scattering (DLS)

DLS measurements are performed at temperatures 20°C, 30°C, 40°C, 50°C and 60°C using Nano-S Malvern instrument attached with a thermostat cell holder. All the samples are equilibrated at each temperature for 10 minutes prior to measurement. Sample preparation, laser wavelength and measuring angle are same as reported in chapter 2.

4.2.3. Kinetics of unfolding

Kinetics of protein unfolding is monitored by the change in ellipticity value (θ_{222}) at 222 nm wavelength. Protein is heated at the following temperatures 53°C, 55°C, 57°C, 59°C, 60°C, 63°C and 65°C for 4 hours and θ_{222} is recorded. Buffer (50 mM potassium phosphate, pH 7, 250 mM NaCl, 1 mM PMSF) is equilibrated at each temperature prior to addition of STY3178. The dead time of mixing is 15 seconds. All the data are background corrected.

Fraction of folded protein (f_N) is estimated for each temperature using θ_{222} value and applying equation (1). f_N values are plotted as a function of time (t) for each temperature (T)

and fitted with single exponential decay equation, $f_N = y_f + ae^{-k_u t}$, where k_u is rate of unfolding. We assume first order kinetics and apply Arrhenius equation, $k_u \sim e^{-E_a^u/RT}$ to estimate activation energy (E_a^u) of unfolding from the slope of $\ln k_u$ versus $1/T$ plot. T and R denote the absolute temperature and universal gas constant, respectively.

4.2.4. Kinetics of refolding

We follow the refolding kinetics by measuring the ellipticity change at 222 nm (θ_{222}) of the unfolded protein. We first heat the protein (10 μ M) at 70°C for 10 minutes in order to unfold it. The unfolded protein at 70°C is subjected to sudden cooling by lowering the temperature of the cell holder. The temperatures at which protein is subsequently refolded by cooling are 30°C, 32°C, 34°C, 36°C, 38°C and 40°C. There is a time lag between the beginning of the measurement and the increase in f_N value as a result of refolding. The lag time is controlled by the instrument itself and is different for different temperatures (99.5 sec-30°C, 99.5 sec-32°C, 99.5 sec-34°C, 101 sec-36°C, 99.5 sec-38°C and 101 sec-40°C). We finally average our data over three sets of measurements. The change in folded protein fraction, Δf_N is defined as $\Delta f_N = f_N - f_R$, (f_R is the folded fraction at which f_N increases) in order to account for the difference between the lag time and initiation of refolding. The f_R values are non-zero, however, small for various temperatures. Similarly, the time at which f_N starts increasing is denoted by Δt and given by subtracting the lag time from each set of measurement. The plot of Δf_N versus Δt is fitted with single exponential equation, $f_N = y_0 + a(1 - e^{-k_f t})$ which provided the rate of refolding (k_f). The plot of natural logarithm of k_f ($\ln k_f$) with $1/T$ is fitted linearly. This Arrhenius plot is utilized to estimate the activation energy of refolding (E_a^f).

4.2.5. Tertiary structure from steady state fluorescence

Fluorescence measurements are carried out in Jobin Yvon Horiba Fluorolog equipped with a peltier unit (Wavelength Electronics). The excitation and emission slits used during measurement is 2 nm. 10 μ M protein is heated gradually at various temperatures ranging from 20°C to 70°C with increments of 5°C followed by equilibration of 10 minutes at each temperature (25°C, 30°C, 35°C, 40°C, 45°C, 50°C, 55°C, 60°C, 65°C and 70°C). Sample is excited at 275, 280 and 295 nm after equilibration. Fluorescence emission is collected in the range 300 nm to 500 nm for all the excitation wavelengths. Similarly, the same sample is

cooled down by gradually lowering the temperature from 70°C to 20°C in order to monitor its refolding. The final spectra at various temperatures are background corrected by subtracting the spectrum of the buffer containing 50 mM potassium phosphate (pH 7) 250 mM NaCl, 1 mM PMSF.

4.2.6. Nuclear Magnetic Resonance spectroscopy

E. coli cells are grown in M9 minimal media containing $^{15}\text{NH}_4\text{Cl}$ in order to uniformly label the protein with ^{15}N . Protein extraction and purification methods are same as described earlier in chapter 2. ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) experiments are performed in a 600 MHz Varian spectrometer equipped with a room temperature probe at different temperatures 25°C, 30°C, 35°C, 40°C and 45°C. 400 μM protein is used for the measurements in a buffer containing 30 mM phosphate buffer (pH 7), 150 mM NaCl and 10 % D_2O . All the HSQC measurements are acquired as 64 scans and 128 number of points. NMRPIPE⁹⁵ is used for data processing. Referencing of the HSQC spectra at 25°C, 30°C, 40°C and 45°C are performed with respect to that at 35°C using NMRVIEW⁹⁶.

4.3. Results

In chapter 2, we have shown that STY3178 is an α -helical protein with two characteristic minima around 209 and 222 nm. Its emission maxima is around 342 nm measured from steady state fluorescence. We monitor the effect of heat on secondary and tertiary structure of STY3178.

4.3.1. Unfolding with temperature

The α -helical secondary structure remains stable and almost intact up to 50°C upon heating gradually with increment of 10°C in the temperature range 20°C to 70°C (Figure 4.1a). Substantial changes are observed in secondary structure of the protein above 50°C. The two characteristic minima around 209 and 222 nm diminish completely at 60°C (Figure 4.1a). This shows melting of the protein secondary structure. CD spectrum does not show any further change at 70°C compared to that at 60°C (Figure 4.1a). This indicates complete unfolding of the protein occurs at 60°C.

Similarly, we monitor the effect of temperature on the environment of the fluorophores using steady state fluorescence. The emission spectra for 280 nm excitation shows gradual decrease in intensity as temperature increases from 20°C to 70°C (Figure 4.1b). Similar

observation is found for 275 and 295 nm excitations. The emission peak position for 280 nm excitation is plotted as a function of temperature and observed to remain unchanged up to 45°C (Figure 4.1c). It starts shifting towards higher wavelength when heated above 45°C. The emission peak position is red shifted by 10 nm between 45°C to 55°C. This indicates protein unfolding starts beyond 45°C. The peak position does not change further above 55°C. However, the emission intensity continues to decrease till 70°C.

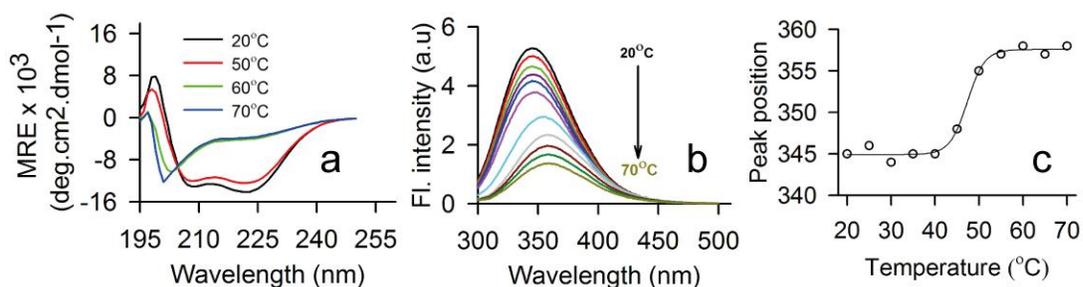


Figure 4.1. Secondary and tertiary structure of STY3178 at elevated temperatures. (a) Shows the far UV-CD spectra of the protein at various temperatures 20°C (black), 50°C (red), 60°C (green) and 70°C (blue). (b) Shows the fluorescence emission spectra of protein at temperatures 20°C (black), 25°C (red), 30°C (green), 35°C (purple), 40°C (blue), 45°C (magenta), 50°C (cyan), 55°C (grey), 60°C (brown), 65°C (dark green) and 70°C (dark yellow). (c) Shows the plot of peak position as a function of temperature for excitation at 280 nm during unfolding of protein.

We perform NMR measurements to observe the tertiary structural changes in protein with increasing temperature. At 25°C, HSQC spectrum of ¹⁵N-labelled protein contains well dispersed peaks in the region 6-10.5 ppm (Figure 4.2a). Large number of peaks is found to cluster around 7.5-8.5 ppm region in the spectrum (Figure 4.2a). This is indicative of α -helical protein. It also contains well dispersed peaks other than the clustered region which suggests presence of some β -sheet elements in the structure. The HSQC spectrum is also monitored at four other temperatures 30°C (Figure 4.2b), 35°C (Figure 4.2c), 40°C (Figure 4.2d) and 45°C (Figure 4.2e). We find more number of peaks in the HSQC spectra with increase in temperature. This is due to increase in molecular tumbling with temperature. The overall spectra are very similar in the temperature range 25°C to 45°C (Figure 4.2f). This indicates that the tertiary structure is quite stable at 45°C supporting CD and fluorescence results. Only few peaks show sensitivity with increasing temperature in the HSQC spectra (Figure 4.2f). Those peaks showing sensitivity are majorly from the well dispersed regions.

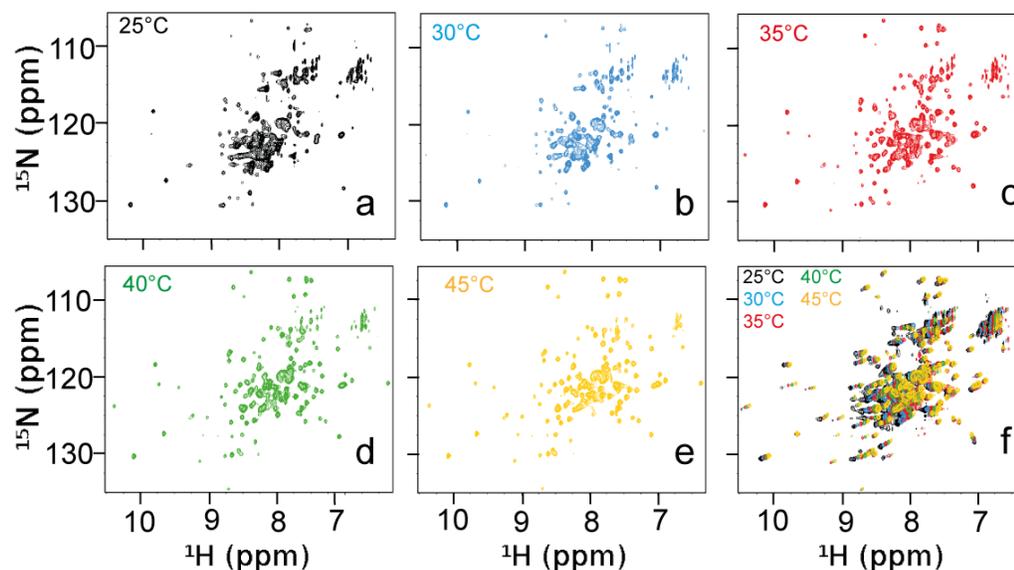


Figure 4.2. ^{15}N - ^1H HSQC spectra (6-10.5 ppm) of STY3178 at various temperatures (a) 25°C (black), (b) 30°C (blue), (c) 35°C (red), (d) 40°C (green) and (e) 45°C (yellow). (f) Represents the superimposed spectra of protein at all the temperatures mentioned in (a)-(e).

4.3.2. Reversibility in unfolding

We next determine the possibility of STY3178 refolding into its native structure from an unfolded state. The thermally unfolded protein is gradually cooled from 70°C to 20°C by decreasing 10°C at a time and CD spectra is recorded. The protein remains in unfolded state up to 50°C (Figure 4.3a). Refolding of the protein is observed upon cooling down further to 40°C as evident from the α -helical secondary structure (Figure 4.3a). The CD spectrum of the refolded protein (Figure 4.3a) is similar to that of the native protein at 20°C (Figure 4.3a) prior to unfolding. This observation establishes reversibility in thermal unfolding of STY3178.

We further monitor the capability of protein to refold back from thermally unfolded state using steady state fluorescence measurement. Figure 4.3b shows the emission spectra of the protein after cooling from 70°C to lower temperatures for excitation at 280 nm. The emission intensity increases as temperature decreases to 20°C compared to the unfolded state at 70°C. We also observe a blue shift of about 8 nm in emission maxima along with increase in intensity upon cooling (Figure 4.3c). Similarly results are observed for 275 and 295 nm excitations. This further suggests refolding of the protein which results in a blue shift. However, the final emission maxima is red shifted by 2-3 nm (Figure 4.3c) and its intensity is less in the refolded protein (Figures 4.3b) compared to that of native state (Figure 4.1b,c). This

suggests slight rearrangement in the conformation of amino acids side chains. Therefore, both steady state fluorescence and CD results are in line indicating reversibility in thermal unfolding of the protein.

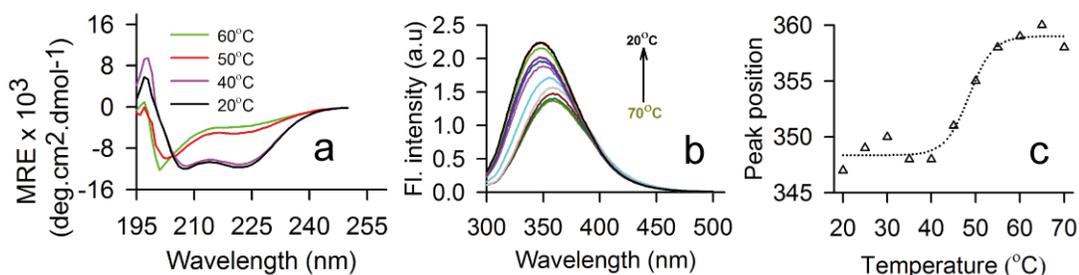


Figure 4.3. Structure of STY3178 upon cooling. (a) Shows the far UV-CD spectra after cooling from 70°C to lower temperatures 60°C (green), 50°C (red), 40°C (magenta) and 20°C (black). (b) Shows the steady state fluorescence spectra of protein at 280 nm excitation upon cooling to temperatures 20°C (black), 25°C (red), 30°C (green), 35°C (purple), 40°C (blue), 45°C (magenta), 50°C (cyan), 55°C (grey), 60°C (brown), 65°C (dark green). The unfolded protein spectrum at 70°C is shown in dark yellow. (c) Shows the plot of peak position versus temperature for excitation at 280 nm during protein refolding.

We further determine the conformational stability of STY3178 by heating it for a long period of time. It is heated at various high temperatures (50°C, 60°C, 70°C, 80°C, 90°C and 100°C) for 30 minutes using a water bath and allowed to cool down to room temperature after that. The far UV-CD of the protein samples after cooling are recorded and compared with the native spectrum. Protein heated in the temperature range 50°C to 80°C, refolds completely into a native-like structure (Figure 4.4). There is a slight decrease in ellipticity of the refolded protein upon heating at 90°C indicating partial loss of secondary structure (Figure 4.4). Refolding could not be achieved for the sample heated at 100°C (Figure 4.4). Thus, the reversibility in protein unfolding is maintained completely up to 80°C and partially for 90°C.

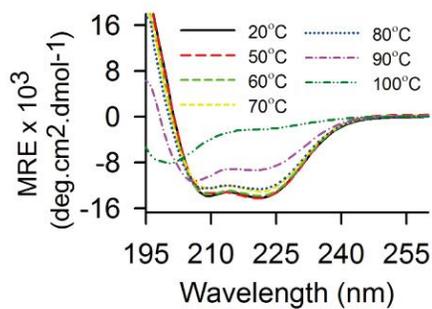


Figure 4.4. Secondary structure of the protein upon cooling from elevated temperatures. Far UV-CD spectra of refolded protein recorded at room temperature after heating at 50°C (red), 60°C (green), 70°C (yellow), 80°C (blue), 90°C (magenta) and 100°C (dark green) for 30 minutes. The native protein spectrum at 20°C is shown in black for comparison.

To determine the stability of the protein trimeric assembly, the hydrodynamic size of STY3178 is monitored in the temperature range 20°C to 60°C (Figure 4.5, black) using DLS. The hydrodynamic diameter of the native protein at 20°C is around 6.5 nm as reported in chapter 2. The size varies within 1 nm with increase in temperature which is not a substantial change to consider. This suggests that the oligomeric state of the protein is stable and its size remains almost unchanged upon heating. Similar native-like hydrodynamic size (~6.5 nm) is observed upon cooling the unfolded protein from 60°C to 20°C (Figure 4.5, red).

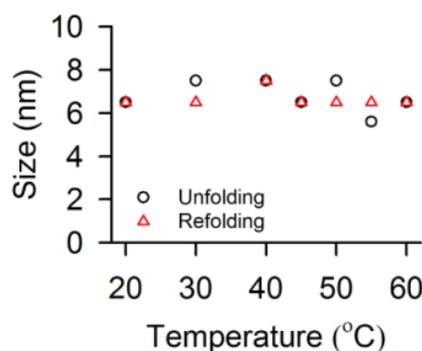


Figure 4.5. Effect of temperature on the hydrodynamic size of STY3178. The hydrodynamic diameter is plotted as function of temperatures 20°C, 30°C, 40°C, 50°C and 60°C during unfolding (black) and refolding (red).

4.3.3. Dependence on heating and cooling rates

The thermal unfolding of a protein could be either kinetically or thermodynamically controlled. We heat the protein at different heating rates and record the ellipticity at 222 nm (θ_{222}). f_N is calculated from the θ_{222} value as described in methods. It does not unfold below 50°C for all the heating rates. f_N remains unchanged up to 50°C and decreases on further increasing temperature (Figure 4.6a). A difference in half denaturation temperature (T_m) at $f_N=0.5$, is observed for different heating rates. T_m is found to be low for slower heating rates and increases for higher heating rates as shown in Table 4.1. In other words, rapid unfolding is achieved upon exposing the protein for long time at elevated temperature. This is a characteristic of kinetically controlled process^{94, 97, 98}. Unfolded protein at 70°C is cooled at different cooling rates and θ_{222} is measured similarly. f_N is plotted as a function of temperature for various cooling rates (Figure 4.6b). It indicates refolding of the protein achieved below 50°C. There is a hysteresis in unfolding and refolding of the protein. Hysteresis is caused due to the insufficient time for the protein to reach equilibrium during faster scan rate. Half renaturation temperature (T_m') is found to be lower than half denaturation temperature (T_m) as shown in Table 4.1. Hysteresis is observed to decrease with decreasing rates of heating and cooling.

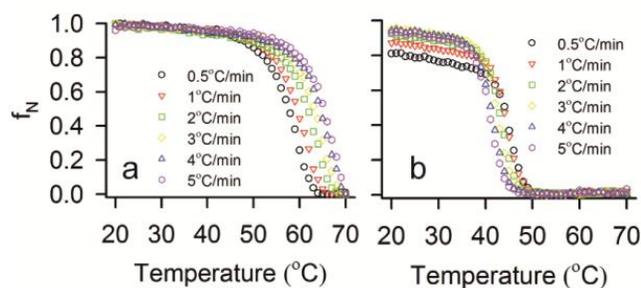


Figure 4.6. Unfolding and refolding of STY3178. (a) The variation of f_N with temperature for different heating rates 0.5°C/min (black), 1°C/min (red), 2°C/min (green), 3°C/min (yellow), 4°C/min (blue) and 5°C/min (magenta) is shown. (b) Change in f_N upon cooling the protein from 70°C at different cooling rates 0.5°C/min (black), 1°C/min (red), 2°C/min (green), 3°C/min (yellow), 4°C/min (blue) and 5°C/min (magenta) is represented.

Table 1. Half denaturation (T_m) and renaturation temperature (T_m') of protein STY3178 for different heating and cooling rates.

Rate (°C/min)	T_m (°C)	T_m' (°C)
0.5	56.57	46.57
1	58.65	45.73
2	61.15	44.59
3	62.40	43.96
4	63.65	43.24
5	64.59	42.40

4.3.4. Kinetics measurement

The kinetics of unfolding of STY3178 is measured using CD. The ellipticity changes at 222 nm (θ_{222}) is recorded for 4 hours at temperatures 53°C, 55°C, 57°C, 59°C, 60°C, 63°C and 65°C. f_N values approaching to zero after long time is indicative of complete unfolding (Figure 4.7a). We find f_N decreases below 50 % within 50 minutes of heating the protein at the above mentioned temperatures. The plot of f_N with time follows single exponential decay. The rate of unfolding (k_u) is estimated from the exponential fit of f_N plotted as a function of time and shown for few temperatures in Figure 4.7b. The plot of $\ln k_u$ versus $1/T$ follows the Arrhenius behavior and the activation energy of unfolding (E_a^u) is calculated from the slope of this plot (Figure 4.7c). E_a^u of about 246.9 kJ/mol is obtained for STY3178 and is comparable to that of other proteins reported in literature⁹⁹⁻¹⁰¹. We do not perform kinetics measurement below 53°C temperature because complete unfolding of protein could not be achieved despite long measurement time.

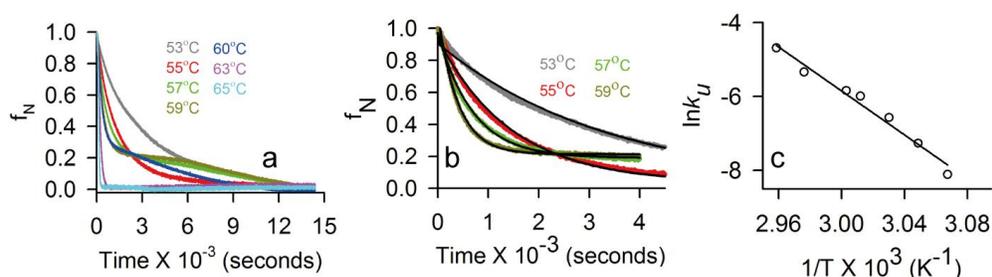


Figure 4.7. Unfolding kinetics of STY3178. (a) Fraction of folded protein (f_N) is plotted as a function of time upon heating at temperatures 53°C (grey), 55°C (red), 57°C (green), 59°C (dark yellow), 60°C (blue), 63°C (magenta) and 65°C (cyan). (b) Represents the exponential fitting of f_N versus time for temperatures 53°C (grey), 55°C (red), 57°C (green) and 59°C (dark yellow). (c) Shows the linear dependence of $\ln k_u$ over $1/T$ (Kelvin⁻¹) following Arrhenius behaviour.

The refolding kinetics of STY3178 is measured similarly using the change in θ_{222} . The thermally unfolded protein at 70°C is allowed to cool down to lower temperatures (30°C, 32°C, 34°C, 36°C, 38°C and 40°C) in order to follow the refolding kinetics. We observe a time lag (t_L) of about 100 seconds after which f_N increases rapidly ($f_N \approx 0.95$) within 300 seconds as a result of refolding (Figure 4.8a). The reported spectra over the entire time of measurement (0 to 600 seconds) are an average of three sets of data acquired at different temperatures. The change in Δf_N is defined as $\Delta f_N = f_N - f_R$, where f_R represents the folded fraction at time $t = t_L$. The f_R has small and non-zero values for various temperatures. Similarly, the increase in time is defined as $\Delta t = t - t_L$. The plot of Δf_N as a function of Δt show single exponential rise as shown for few temperatures in Figure 4.8b. The rate of refolding (k_f) is estimated from exponentially fitted curves. The plot of $\ln k_f$ versus $1/T$ shows Arrhenius behavior upon refolding with activation energy (E_a^f) of about -58.66 kJ/mol (Figure 4.8c). The quality of linear fit in refolding kinetics (Figure 4.8c) is poorer compared to that in unfolding (Figure 4.7c) as reflected in the R^2 values of the fits (0.97 for unfolding and 0.85 for refolding). This leads us to determine the dependences of $\ln(k_u/T)$ and $\ln(k_f/T)$ on $1/T$.⁹³ The plots confirm the linear dependences for both processes (Figure 4.8d). This indicates that the activation heat capacities⁹³ are negligible for both unfolding and refolding.

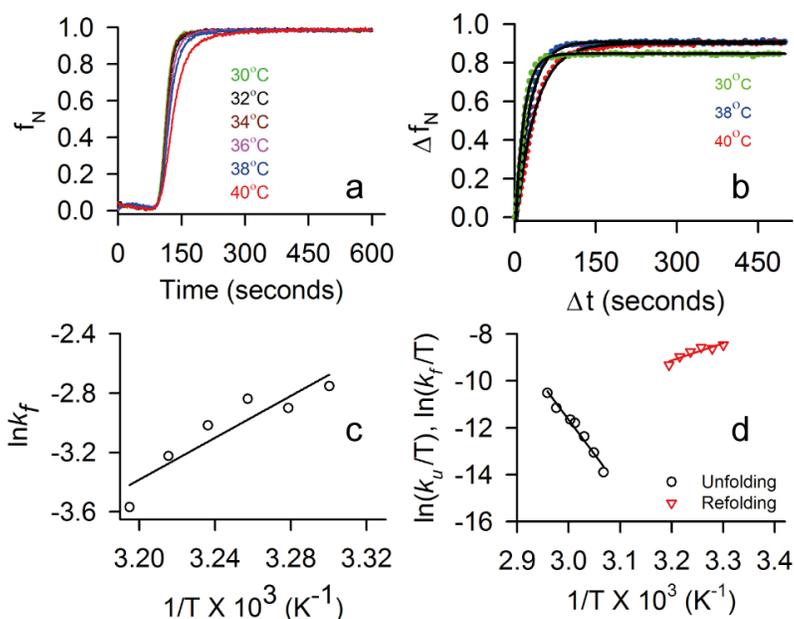


Figure 4.8. Refolding kinetics of STY3178. (a) The change in f_N with time is shown during refolding for temperatures 30°C (green), 32°C (black), 34°C (brown), 36°C (magenta), 38°C (blue) and 40°C (red).

(b) Shows the variation in Δf_N ($\Delta f_N = f_N - f_R$, where f_R is the folded fraction from which the rise in f_N starts) with time (Δt) at which f_N starts increasing. It is fitted exponentially and shown for temperatures 30°C (green), 38°C (blue) and 40°C (red). (c) Arrhenius plot of $\ln k_f$ versus $1/T$ (Kelvin⁻¹) showing linear dependence. (d) Shows the plots of $\ln(k_u/T)$ and $\ln(k_f/T)$ as functions of $1/T$ for unfolding (black) and refolding (red) kinetics of STY3178.

4.4. Conclusion

The studies related to thermal response of STY3178 using CD, steady state fluorescence, DLS and NMR show reversible unfolding and structural stability at elevated temperatures. STY3178 can refold back upon cooling from 80°C. There is hysteresis in unfolding and refolding of the protein as observed from the time dependent CD measurements in the temperature range 20°C to 70°C. The unfolding and refolding kinetics shows the Arrhenius behavior with the activation energy barriers 246.9 kJ/mol and -58.66 kJ/mol, respectively.

Chapter 5| Response of STY3178 in various chemical environments

5.1. Introduction

Proteins unfold due to change in solution pH or in presence of chemical agents such as urea, guanidium chloride (GndHCl), sodium dodecyl sulphate (SDS). Among these, pH of the medium has a significant role in biological processes.¹⁰²⁻¹⁰⁴ The protonation state of amino acids is governed by the pH of solution.^{102, 105} As a result, protein structure, stability and function are also influenced by pH changes.¹⁰⁵⁻¹¹¹ Other chemical agents disrupt the structure of protein by different mechanisms. For example, urea perturbs the protein structure either by altering solvent environment around it or by directly interacting with the polypeptide backbone.^{112, 113} GndHCl also unfolds proteins in a similar manner¹¹⁴, although, it is more potent in unfolding the protein compared to urea.¹¹⁵ The anionic surfactant SDS is known for its denaturing property as well as for stabilizing the α -helices of protein.¹¹⁶⁻¹¹⁹

Small proteins often show reversibility in chemical unfolding.¹⁰ For example, bovine pancreatic trypsin inhibitor (BPTI) is known to have resistance towards chemical denaturation.¹²⁰ Cytochrome c from yeast¹²¹ and calcium binding protein, Porcine S100A12⁹⁰ are capable of showing reversible chemical unfolding. Proteins well studied for reversibility in chemical unfolding include serum albumin from bovine and human.^{91, 92, 122, 123} However, in literature, only a few multi-domain proteins are reported for refolding into a native-like structure from an unfolded state. These include skeleton proteins namely erythroid, non-erythroid spectrins¹²⁴ and oligomeric heat shock protein, chaperonin 10 (cpn10) from human¹²⁵.

There are helical membrane proteins, known for undergoing reversible chemical denaturation. For example, bacteriorhodopsin^{126, 127} is a membrane protein to be refolded *in vitro*. The integral membrane protein diacylglycerol kinase¹²⁸ from *E. coli* also shows reversible chemical denaturation. Another membrane protein, KcsA¹²⁹ from potassium channel superfamily and disulphide bond reducing protein DsbB¹³⁰ from *E. coli*, are reported for reversible refolding capability. An α -helical membrane protein, CopA from *Archaeoglobus fulgidus* is also reported to regain its native-like structure and function up on denaturation by GndHCl.¹³¹ Apart from these aforementioned proteins, there are few transporter proteins like EmrE and GaIP from small multidrug resistant transporter (SMR) and major facilitator

superfamily (MFS), respectively, which show similar reversible unfolding in presence of chaotropes.^{132, 133}

Surfactants, on the other hand, are more potent in inducing protein unfolding compared to urea and GndHCl.¹³⁴ They are required only in millimolar concentration whereas urea and GndHCl are required in molar concentration range for protein unfolding due to weak interaction with peptide backbone.¹³⁴ However, surfactants can suppress protein aggregation as well.^{135, 136} For example, surfactant Hecameg is reported to dissociate cytochrome *b₆f* complex into its monomer.¹³⁷ The most commonly used surfactant SDS also prevents aggregation of proteins such as EmrE¹³⁸ belonging to SMR protein family, BRICHOS¹³⁹ domain trimer, multi-oligomeric haemoglobin¹⁴⁰ and insulin¹⁴¹. The other reported properties of SDS are stabilization of α -helical protein above its critical micellar concentration (CMC).¹⁴²⁻¹⁴⁴ On the other hand, β -sheeted proteins get stabilized below the CMC of SDS.¹⁴⁴⁻¹⁴⁶ Apart from these, the tertiary structure of protein can get altered in presence of SDS while secondary structure remains unchanged as reported for cytochrome *c*.^{144, 147-153}

We report in previous chapters that multidrug and acid response regulator, EvgA activates genes which are involved in either acid tolerance response (ATR), high osmolarity or efflux transport mechanism in *E. coli*.²⁰ Genes *osmC* and *ompA* are related to high osmolarity responses^{20, 154-156} whereas *yhiU*²⁰ and *yhiV*²⁰ are efflux transporters. Those related to acid tolerance are *yfdXUWEV*.^{33, 35} The growth of bacterial cells in acidic medium is reported to be affected by deletion of these genes.³⁵ Functional characterization of *yfdW*³⁵, *yfdU*³⁵, *yfdE*³⁴ and *yfdV*³⁴ are performed. *yfdW* is formyl-coenzyme A (CoA) transeferase whereas *yfdU* is oxalyl-CoA decarboxylase.³⁵ *yfdE* and *yfdV* are suggested to be involved in oxalate-induced ATR mechanism in *E.coli*.³⁴ However, response of *yfdX* protein in different chemical environments is not reported.

In this chapter, we show the effect of various chemical environments on structural stability of *yfdX* protein, STY3178. We have chosen different pH and chemical agents such as urea, guanidium chloride (GndHCl) and sodium dodecyl sulphate (SDS), for perturbing the native structure of STY3178. We employ techniques like circular dichroism (CD), steady state fluorescence, dynamic light scattering (DLS), nuclear magnetic resonance (NMR) and isothermal titration calorimetry (ITC) to monitor the changes induced by different perturbing agents.

This chapter is organized as follows: Section 5.2 includes the methods related to sample preparation followed by experimental techniques used for monitoring the stability of STY3178

in presence of chemicals. Results covering the effect of pH, urea, GndHCl and SDS are included in section 5.3. Discussion and conclusion are given in sections 5.4 and 5.5, respectively.

5.2. Methods

5.2.1. Sample preparation for experiments

pH dependent measurements: The effect of acidic and alkaline media on STY3178 is monitored for the pH values 2.5, 3, 4, 4.5, 5.2, 6, 6.5, 7, 7.5, 8, 8.5, 9 and 10. Buffers used for achieving the desired pH are as follows: glycine-HCl (pH 2.5 and 3), sodium acetate-acetic acid (pH 4, 4.5 and 5.2), potassium phosphate (pH 6, 6.5 and 7), tris-HCl (pH 7.5, 8 and 8.5) and glycine-NaOH (pH 9 and 10). STY3178 (10 μ M) is incubated overnight in 50 mM buffers having the aforesaid pH values.

Urea and guanidine hydrochloride (GndHCl): 5 μ M protein is used for all urea and GndHCl unfolding experiments. Buffer contains 50 mM potassium phosphate (pH 7), 250 mM sodium chloride (NaCl) and 1mM phenylmethanesulfonyl fluoride (PMSF) having urea or GndHCl. The concentrations of urea or GndHCl used for unfolding experiments are 0.2 M, 0.4 M, 0.6 M, 0.8 M, 1 M, 1.2 M, 1.4 M, 1.6 M, 1.8 M, 2 M, 2.5 M, 3 M, 4 M and 6 M. STY3178 in buffers having various concentrations of urea or GndHCl is equilibrated over night at room temperature.

For carrying out refolding measurements, 200 μ M protein is first unfolded in 8 M of either urea or GndHCl by incubating overnight at room temperature. This unfolded protein in 8 M chaotrope is diluted by various folds to achieve lower concentrations of urea or GndHCl. Protein concentration is kept fixed at 5 μ M for all the refolding experiments. In a different experimental set up, unfolded protein (200 μ M STY3178 in 8 M urea or GndHCl solution) is diluted instantaneously in buffer with no chaotrope to reduce urea or GndHCl concentration by 40, 30, 20 and 10 folds. Measurements are performed immediately after dilution without any equilibration.

Sodium dodecyl sulphate (SDS): 5 μ M protein is incubated overnight in buffers (50 mM potassium phosphate, pH 7, 250 mM NaCl and 1mM PMSF) having SDS. Various SDS concentration used during the experiment are 0.69 mM (0.02 %), 1.39 mM (0.04 %), 2.1 mM (0.06 %), 2.78 mM (0.08 %), 3.47 mM (0.1 %), 8.7 mM (0.25 %), 17.35 mM (0.5 %), 34.7 mM (1 %) and 69.4 mM (2 %). In another set, 200 μ M protein is first equilibrated overnight in

presence of 173.5 mM (5 %) SDS. The concentration of SDS is then reduced by different dilutions keeping protein concentration fixed at 10 μ M.

Samples for background correction are prepared similarly without adding protein for each set of measurement in presence of respective chemical agents.

5.2.2. Circular Dichroism measurements

CD measurements for protein samples (mentioned above) in various pH buffers, urea, GndHCl and SDS are carried out in Jasco J-815 CD spectrometer. All measurements in far-UV-region (200-250 nm) and near UV-CD region (250-330 nm) are performed using 1 mm and 10 mm path length cells, respectively. Scan speed used during the measurements is 100 nm min^{-1} . All spectra are averaged over minimum two sets of measurements. Background correction is done by subtracting the data for corresponding buffer solutions with respective pH, urea, GndHCl and SDS from that of the sample.

The fraction of folded protein (α_i) is estimated using the ellipticity values (θ_{222}) at 222 nm for various concentrations of chemical agents using the relation, $\alpha_i = (\theta_i - \theta_U) / (\theta_N - \theta_U)$, where θ_i denote the ellipticity of protein at different concentrations of chemical agents. θ_N and θ_U are ellipticity of native and unfolded protein, respectively. Folding constant (K) is then calculated from α_i value at a certain concentration of chemical agents using the relation¹⁵⁷, $K = P_i \alpha_i / (n P_i (1 - \alpha_i))^n$, described by Greenfield. P_i defines monomer protein concentration expressed in moles per liter and n is the number of subunits. The free energy associated with protein folding is estimated using K value in the expression, $\Delta G = -RT \ln K$, where R and T are the universal gas constant (1.98 kcal/mol) and absolute temperature (Kelvin), respectively. The free energy (ΔG_0) in absence of chemical agents is estimated from the plot of free energy of folding (ΔG_i) as a function of chaotrope concentration (D_i), assuming the process to be two-state and a linear dependence ΔG_i over D_i according to the relation, $\Delta G_i = \Delta G_0 + m[D_i]$. ΔG_i defines the free energy of protein at a given concentration of chemical agent and m is the slope.¹⁵⁷

5.2.3. Intrinsic fluorescence measurements

The measurements of steady state fluorescence are recorded in Jobin Yvon Horiba Fluorolog using a quartz cell of 10 mm path length. The excitation wavelengths used during

measurement are 275, 280 and 295 nm. Concentration of protein is kept fixed at 10 μM throughout the measurements. pH dependent measurements are performed using 2 nm slit width and that in presence of urea, GndHCl and SDS are carried out using slit width of 3 nm. All spectra are background corrected by subtracting the data of respective buffers.

5.2.4. Nuclear magnetic resonance measurement

The protocol used for preparation of ^{15}N -labelled protein and ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) experiments are same as described in chapter 4. Data are collected at 35°C in 30 mM phosphate buffer (pH 7), 150 mM NaCl and 10 % D_2O with 64 scans. HSQC spectra of 400 μM protein are recorded in presence of urea (unlabeled) with concentrations 0, 0.5, 1, 2 and 3 M. Protein in 3 M urea is buffer exchanged several times with that of 0 M urea buffer using a spin concentrator (Amicon, 10 KDa cut off). This resulted in a lower urea concentration of about 0.013 M. HSQC of protein after buffer exchange is then recorded similar to the native one. All HSQC spectra are processed in NMRPIPE⁹⁵ and analyzed in NMRVIEW⁹⁶.

5.2.5. Dynamic Light Scattering

Hydrodynamic diameter of protein is monitored in presence of different pH, urea, GndHCl and SDS using Nano-S Malvern instrument at 20°C. The wavelength of laser and angle of scattering used for all measurements is 632.8 nm and 173°, respectively. 10 μM protein is equilibrated overnight in buffers with various pH, urea, GndHCl and SDS concentration as described in sample preparation section. Measurements for monitoring refolding are carried out similarly in presence of lower concentration of chemical agents. Each measurement is an average of five successive scans.

5.2.6. Isothermal titration calorimetry (ITC)

ITC measurements are performed for STY3178 (497 μM) by titrating it with 500 mM SDS, loaded in syringe at 298 K in a MicroCal iTC200 calorimeter (GE healthcare). The measurement is carried out in a buffer containing 30 mM phosphate (pH7), 150 mM NaCl and 1mM PMSF. Reference power used during ITC measurement is 10 $\mu\text{cal/s}$. First injection of 0.4 μl and rest 25 injections of 1.5 μl are injected in the cell with constant stirring of 50 rpm. Each injection is done for 3 seconds with 200 seconds spacing and an initial delay of 60 seconds

prior to first injection. The heat of demicellization upon dilution of SDS is subtracted from that of SDS titration into buffer. We plot the integrated data using MicroCal origin.

5.2.7. Size exclusion chromatography in presence of SDS

We perform SEC using superdex 75 column as per the protocol described in chapter 3. The column is pre-equilibrated with buffer containing 50 mM phosphate (pH 7), 250 mM NaCl, 1 mM PMSF and 3.47 mM (0.1 %) or 69.4 mM (2 %) SDS. STY3178 (400 μ M) is first incubated in the above buffer containing 3.47 mM (0.1 %) or 69.4 mM (2 %) SDS. Protein is eluted at a flow rate of 1 ml/minute using a peristaltic pump (GE Healthcare). Various standard proteins (lysozyme, 14.4 KDa; carbonic anhydrase, 29 KDa; ovalbumine, 45 KDa; BSA, 66 KDa and conalbumin, 75 KDa) are used for column calibration as reported in chapter 3.

5.3. Results

5.3.1. Effect of pH

Secondary structure of STY3178 is monitored in the far UV-region (200-250 nm) at different pH using CD. The spectra of protein in different buffers in the pH range 2.5-10 show characteristic minima at 222 nm (Figure 5.1a). We plot the ellipticity at 222 nm (θ_{222}) as a function of pH (Figure 5.1b). This indicates that STY3178 is quite stable and α -helical secondary structure is retained in acidic (pH 2.5) as well as in alkaline (pH 10) media. However, we find an increase in ellipticity for the protein in buffers with pH 7.5-10 compared to the native protein at neutral pH 7. We could not measure the ellipticity at 209 nm for the samples due to limitation of the instrument.

Near UV (250-350 nm) CD spectra of STY3178 at various pH between 2.5 to 10 are monitored to determine the stability of tertiary structure. Native protein at neutral pH 7, has a broad shoulder spanning over 250- 280 nm in near UV-CD spectrum (Figure 5.1c). This broad shoulder disappears in strong acidic pH 2.5 and 3. However, it is retained at pH 4.5 with lower ellipticity compared to the native protein at pH 7 (Figure 5.1c). For the pH range 6 to 10, broad shoulder has ellipticity almost same as that of the native protein (Figure 5.1c). This establishes the stability of protein in the pH range 4.5-10.

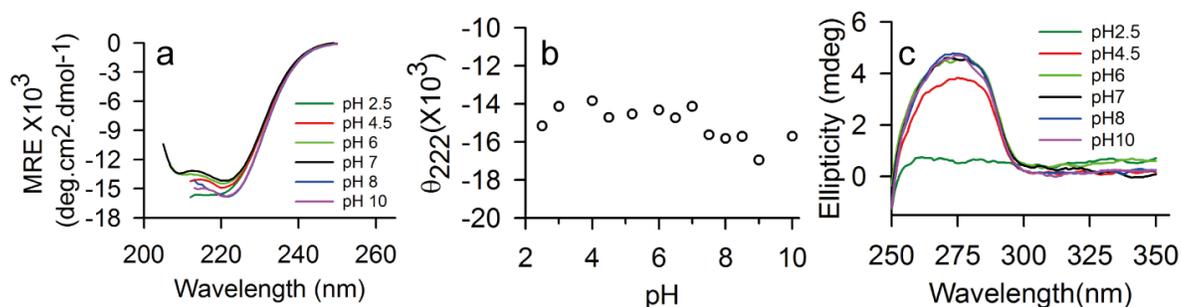


Figure 5.1. Effect of pH on STY3178. (a) Shows the far UV-CD spectra of protein in buffers with pH 2.5 (dark green), 4.5 (red), 6 (green), 7 (black), 8 (blue) and 10 (magenta). (b) Represents the plot of ellipticity at 222 nm (θ_{222}) versus pH in the range 2.5-10. (c) Near UV-CD spectra of protein in buffers having pH values 2.5 (dark green), 4.5 (red), 6 (green), 7 (black), 8 (blue) and 10 (magenta).

Effect of pH on the nature of fluorophores is monitored using steady state fluorescence. Protein samples in buffers with various pH (2.5 to 10) are excited at 275, 280 and 295 nm. The emission spectra upon 280 nm excitation are shown for buffers with pH 2.5, 4.5, 6, 7, 8 and 10 (Figure 5.2a). The emission intensity is found to be low for samples in buffers with pH range 2.5 to 6.5. We observe an increase in emission intensity compared to the native protein for buffers having pH 7.5, 8, 8.5, 9 and 10. The emission peak position is blue shifted to 339 nm in acetate buffers (pH 4, 4.5 and 5.2) compared to that of native protein at pH 7. It remains same as that of the native protein for rest of the pH values. Similar observation is made for 275 and 295 nm excitations. Plots of peak position as a function of pH for excitation wavelengths 275, 280 and 295 nm support these results (Figure 5.2b). These observations are comparable to the far UV-CD results.

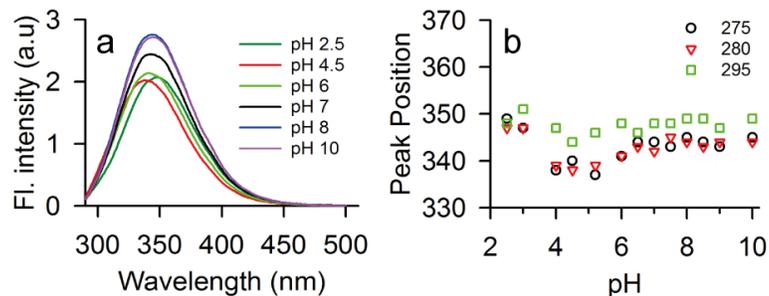


Figure 5.2. Fluorescence emission of STY3178 in different pH. (a) Shows the emission spectra of protein in buffers with pH 2.5 (dark green), 4.5 (red), 6 (green), 7 (black), 8 (blue) and 10 (magenta) at 280 nm excitation. (b) Peak positions are plotted as a function of pH for excitation at 275 (black), 280 (red) and 295 (green) nm.

The hydrodynamic size of STY3178 is measured to determine the effect of pH on the oligomeric state of protein. Native STY3178 at pH 7 has a hydrodynamic diameter of about 6.5 nm, as reported in chapter 2. This corresponds to a trimeric assembly for STY3178 (shown in chapter 3). Figure 5.3a shows the DLS spectra of protein for few representative pH values 2.5, 4.5, 6, 7, 8 and 10. The hydrodynamic diameter remains around 6.5 nm at all these pH values. We plot the hydrodynamic diameter against pH values in the range 2.5-10 (Figure 5.3b). It shows that protein size does not change in acidic as well as alkaline pH. However, a decrease in intensity of the peak around 6.5 nm is observed for pH 2.5, 3, 4, 4.5 and 5.2 compared to pH 7 (Figure 5.3a).

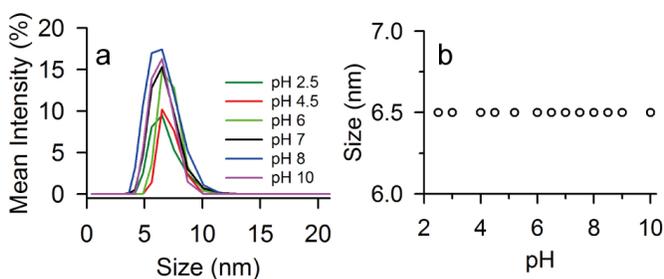


Figure 5.3. Hydrodynamic size of STY3178 in various pH media.

(a) DLS spectra of protein for pH 2.5 (dark green), 4.5 (red), 6 (green), 7 (black), 8 (blue) and 10 (magenta). **(b)** Shows the hydrodynamic diameter plotted as a function of pH in the range 2.5 to 10.

5.3.2. Effect of urea

We monitor the secondary structure of STY3178 in presence of different concentrations of urea using CD. The ellipticity values at 209 and 222 nm decrease up to 0.6 M urea (Figure 5.4a). It starts increasing above 0.6 M urea (Figure 5.4a). The ellipticity change at 222 nm (θ_{222}) is plotted as a function of urea concentration (Figure 5.4b). Decrease in θ_{222} up to 0.6 M urea indicates increase in helix content of protein (Figure 5.4b). θ_{222} increases above 0.6 M urea and reaches a value similar to that of native state at 1.6 M urea (Figure 5.4b). The ellipticity increases further beyond 1.6 M urea concentration (Figure 5.4b). This suggests that unfolding is initiated above 1.6 M urea and major secondary structural changes occur in presence of 2.5 M urea. θ_{222} does not change much in the urea concentration range 2.5 M to 6 M. This indicates unfolding of STY3178 above 2.5 M urea concentration.

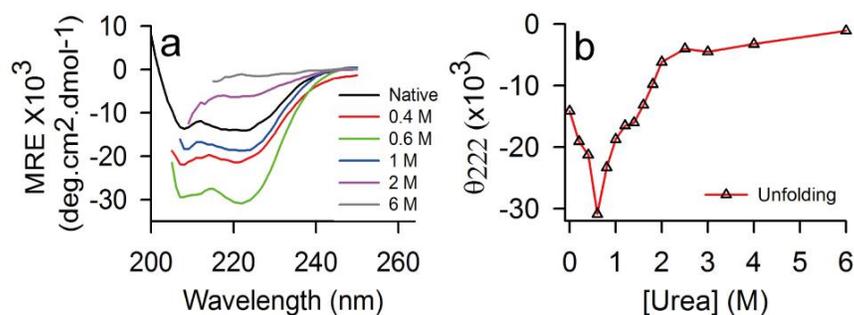


Figure 5.4. Secondary structure of STY3178 in presence of urea.

(a) Represents the CD spectra of protein in native state (black) and in presence of different urea concentrations 0.4 M (red), 0.6 M (green), 1M (blue), 2 M (magenta) and 6 M (grey). (b) Shows the plot of ellipticity (θ_{222}) as a function of different concentration of urea during unfolding.

The changes in tertiary structure of protein in presence of urea are monitored using near UV-CD. There is a broad shoulder spanning from 250 to 280 nm in native protein spectrum due to presence of aromatic residues namely tryptophan, tyrosine and phenylalanine (Figure 5.5a). The ellipticity of this broad shoulder decreases at 0.4 M and 0.6 M urea (Figure 5.5a). It reduces further at 1 M urea indicating tertiary structural changes in STY3178 (Figure 5.5a). This broad shoulder diminishes further at urea concentration 2 M and above, suggesting loss of tertiary structure of protein (Figure 5.5a).

The effect of urea on tertiary structure of protein is further monitored using steady state fluorescence measurements. Emission maximum of native protein is around 342 nm for excitation wavelengths 275, 280 or 295 nm as reported in chapter 2.³² We show the emission spectrum for excitations at 275 nm and 295 nm in Figures 5.5b and c, respectively. The spectra remain unchanged at 1 M urea compared to the native protein for both the excitation wavelengths (Figure 5.5b,c). Peak position shows red shift at 2 M urea and does not change further up to 6 M urea (Figure 5.5b,c). However, there is a small shoulder around 300 nm in 2 M and 6 M urea for 275 nm excitation, which corresponds to tyrosine emission. The emission spectrum at 295 nm excitation in presence of 6 M urea also contains a shoulder around 305 nm (Figure 5.5c). This shoulder is not present in 1 M and 2 M urea.

We calculate the difference spectra (275-295 nm) of protein in presence of urea and compare it with that of the native spectrum. The nature of difference spectrum in 1 M urea remains same as the native protein with a peak around 303 nm (Figure 5.5b). This indicates that

the phenomenon of FRET still occurs in 1 M urea as reported in chapter 2. We find the intensity of difference spectrum is reduced at 2 M urea. It almost diminishes in presence of 6 M urea. This suggests that the FRET between tyrosine and tryptophan is less efficient in 2 M and negligible in 6 M urea due to structural alterations. The observations of fluorescence emission at 280 nm excitation are same as that of 275 nm. We plot the change in emission peak position for excitation wavelengths 275, 280 and 295 nm as a function of different urea concentrations (Figure 5.5d). The emission maxima remain same up to 1.2 M urea and are red shifted above this urea concentration for all the three excitation wavelengths (Figure 5.5d). At 2 M urea, huge red shift of about 11 nm is observed, indicating protein unfolding (Figure 5.5d).

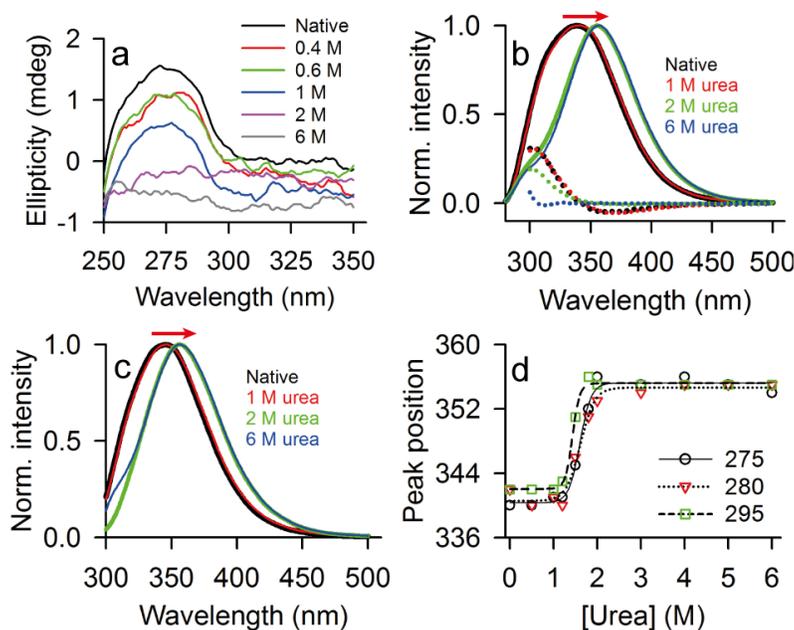


Figure 5.5. Tertiary structure of STY3178 in presence of urea. (a) Shows the near UV (250-350 nm) CD spectra for different urea concentrations 0.4 M (red), 0.6 M (green), 1M (blue), 2 M (magenta) and 6 M (grey). (b) and (c) Show the fluorescence emission spectra of protein in presence of urea for excitation wavelengths 275 and 295 nm, respectively. The difference spectra of 275-295 nm are shown using dotted lines in (b). Concentrations of urea are 1 M (red), 2 M (green) and 6 M (blue) for both (b) and (c). Spectrum of native protein is shown in black for (a), (b) and (c). (d) Shows the variation in emission peak position plotted as a function of different urea concentrations during unfolding for excitation wavelengths 275 nm (black circle), 280 nm (red triangle) and 295 nm (green square).

The hydrodynamic size of protein is monitored in presence of increasing urea concentration. Native protein has a hydrodynamic diameter of about 6.5 nm, as reported in chapter 2. The plot of hydrodynamic diameter versus urea concentration shows that the size remains same as that of native protein up to 2 M urea (Figure 5.6). It increases by 1 nm between 2 M to 4 M urea (Figure 5.6). Hydrodynamic size increases around 12 nm further for protein above 4 M urea (Figure 5.6). This indicates formation of higher order aggregates as a result of unfolding.

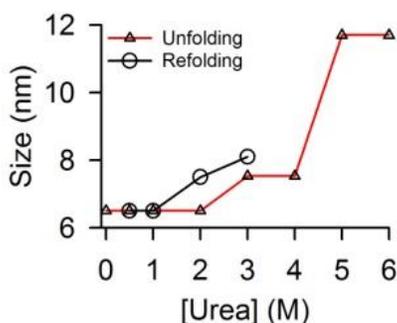


Figure 5.6. Protein hydrodynamic diameter plotted for different concentration of urea during unfolding (triangle) and refolding (circle).

We further monitor the structural changes in protein by performing HSQC measurements in presence of urea. Native protein ^1H - ^{15}N HSQC consists of well dispersed peaks along with clusters in 7.5-8.5 ppm (Figure 5.7a). Presence of clustered peaks shows α -helical nature of the protein. The dispersion of spectrum indicates presence of some β -sheet also (Figure 5.7a). HSQC in presence of 0.5 M urea resembles with native protein spectrum (Figure 5.7b). There are few peaks which show changes in chemical shift in presence of 1 M urea as marked by arrows in Figure 5.7c. Substantial changes are observed in chemical shift position of NH-peaks at 2 M urea (Figure 5.7d). More peaks start shifting towards the central region of HSQC, indicating initiation of protein unfolding. HSQC spectrum is collapsed with cross peaks centered around 7.5-8.5 ppm in presence of 3 M urea indicating complete unfolding of protein (Figure 5.7e). We then reduce 3 M urea concentration of unfolded protein by buffer exchange and record its HSQC. ^1H - ^{15}N HSQC spectrum of protein after reducing urea concentration to about 0.013 M, is shown in Figure 5.7f. It consists of well dispersed NH-cross peaks similar to that of native state (Figure 5.7a). This indicates that STY3178 can refold back to native-like structure.

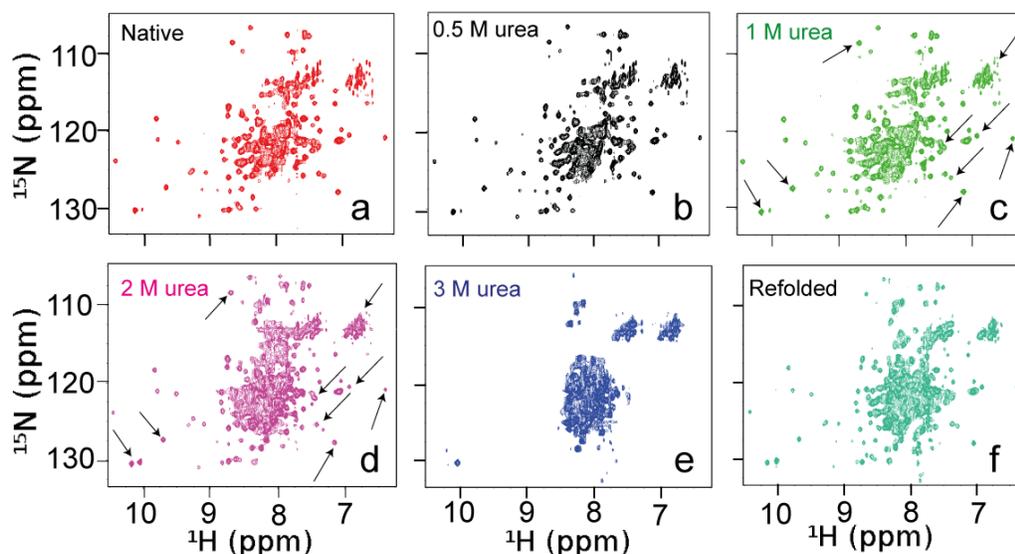


Figure 5.7. ^1H - ^{15}N HSQC spectra of singly labeled STY3178 in presence of urea. HSQC of protein are shown in (a) native state (red) and in presence of (b) 0.5 M (black), (c) 1 M (green), (d) 2 M (magenta) and (e) 3 M (blue) urea. (f) Represents the spectrum (cyan) of refolded protein upon reducing urea to 0.013 M. All these spectra are recorded at 35°C with 64 scans.

In far UV-CD experiments, we equilibrate the unfolded protein (in 8 M urea) in buffers containing lower urea concentrations prior to measurements, as described in method section. CD spectrum shows native-like signature upon reducing urea concentration to 0.4 M, 0.6 M and 1 M (Figure 5.8a). However, no refolding is observed for 2 M to 6 M urea. θ_{222} values of refolded protein are plotted against different urea concentrations (Figure 5.8b). The decrease in θ_{222} below 2 M urea suggests regaining of native-like protein structure (Figure 5.8b). We compare θ_{222} versus urea concentration plot for unfolding and refolding (inset in Figure 5.8b) and observe hysteresis during the process. In another set of experiment, we record the far UV-CD signature after immediate dilution of unfolded protein in 8 M urea by 40, 30, 20 and 10 folds without any equilibration (Figure 5.8c). Upon 40 folds instantaneous dilution of urea (resulting in concentration of 0.2 M), refolding of about 79.3 % is achieved compared to the native protein ellipticity. We observe refolding of 72.6 % for 30 (0.27 M) folds and 65% for both 20 (0.4 M) and 10 (0.8 M) folds of urea dilution. Protein with all these folds of urea dilution has refolded immediately. However, its ellipticity is slightly less compared to native protein spectrum (Figure 5.8a, black).

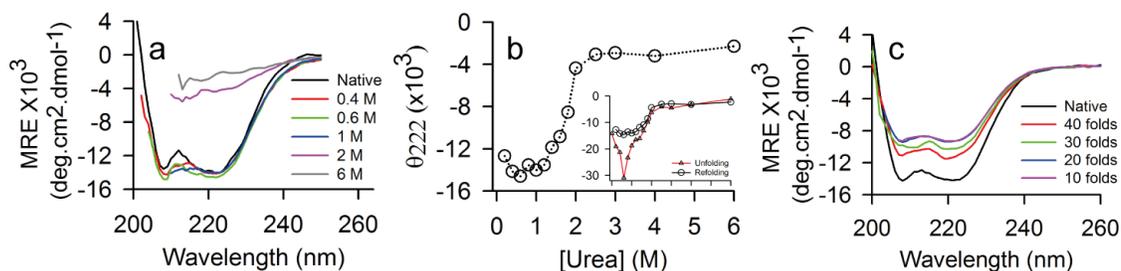


Figure 5.8. Protein secondary upon refolding from urea unfolded state. (a) CD spectra of refolded protein upon reducing urea concentration to 0.4 M (red), 0.6 M (green), 1 M (blue), 2 M (magenta) and 6 M (grey). (b) Plot of ellipticity at 222 nm (θ_{222}) versus different concentration of urea during refolding of protein. The inset shows a comparative plot of θ_{222} versus urea concentration during unfolding (red) and refolding (black). (c) Shows the CD spectra upon urea dilution by 40 (red), 30 (green), 20 (blue) and 10 (magenta) folds. Native protein spectrum is shown in black for both (a) and (c).

The changes in tertiary structure upon protein refolding are further monitored using near UV-CD and fluorescence emission. Near UV-CD spectra show native-like conformation upon reducing urea concentration below 1 M (Figure 5.9a). The broad shoulder is recovered completely for 0.4 M and 0.6 M urea (Figure 5.9a). Protein remains in unfolded state in presence of 2 M and 6 M urea. Fluorescence emission spectra for 275 and 295 nm excitations show blue shift upon reducing urea concentration to 0.4 M and 1 M (Figure 5.9b,c). The emission spectra at 6 M remain red shifted as observed during unfolding (Figure 5.9b,c). The difference spectrum (275- 295 nm) shows emission around 303 nm in presence of 0.4 M and 1 M urea, similar to that of the native protein (Figure 5.9b). This suggests that FRET occurs in the refolded protein as well. Excitation at 280 nm shows observations similar to 275 nm excitation. Plots of peak position versus urea concentration clearly indicate blue shift due to protein refolding below 3 M urea for all the three excitation wavelengths 275, 280 and 295 nm (Figure 5.9d).

We also monitor the hydrodynamic diameter of STY3178 upon refolding and find size of very high order (~ 400 nm) upon reducing urea concentration from 8 M to 4, 5 or 6 M (Figure 5.6). Further dilution of urea concentration to 3 M and 2 M results in lower size of about 8 nm for STY3178 (Figure 5.6). At 1 M and below urea concentration, a hydrodynamic size (~ 6.5 nm) similar to that of native protein is observed (Figure 5.6). This demonstrates that urea unfolding of STY3178 is completely reversible.

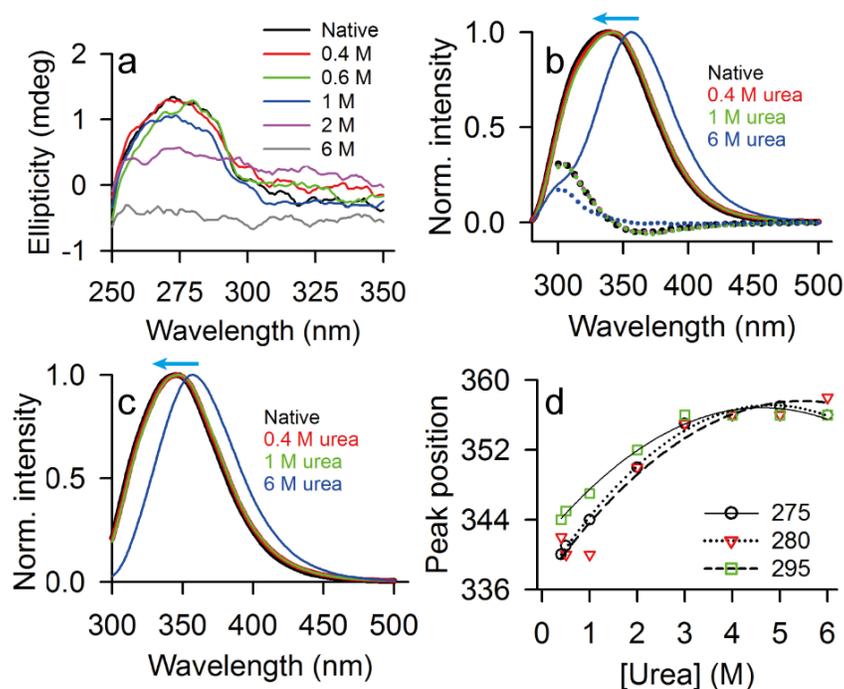


Figure 5.9. Tertiary structure of the refolded protein. (a) Shows the near UV-CD spectra in the range 250-355 nm after reducing urea concentration to 0.4 M (red), 0.6 M (green), 1M (blue), 2 M (magenta) and 6 M (grey). Fluorescence emission spectra of native (black) and refolded protein in 0.4 M (red), 1 M (green) and 6 M (blue) urea are shown upon excitation at (b) 275 nm and (c) 295 nm. The difference spectra for 275-295 nm are shown using dotted lines for urea concentrations 0 M (black), 0.4 M (red), 1 M (green) and 6 M (blue) in (b). (d) The change in emission peak position upon protein refolding for different concentration of urea at excitation wavelengths 275 nm (black circle), 280 nm (red triangle) and 295 nm (green square).

5.3.3. Effect of guanidium chloride

The changes in protein secondary structure are monitored similarly, in presence of GndHCl using CD. Native protein structure remains stable up to 0.4 M GndHCl with two characteristic minima at 209 and 222 nm (Figure 5.10a). The ellipticity starts increasing at 0.6 M GndHCl (Figure 5.10a). The two minima diminish completely in GndHCl concentration range 1 M to 6 M (Figure 5.10a). The plot of θ_{222} versus GndHCl concentration shows unfolding of protein at 0.8 M and above GndHCl (Figure 5.10b). θ_{222} remains almost unchanged in 1 M to 6 M GndHCl concentration range.

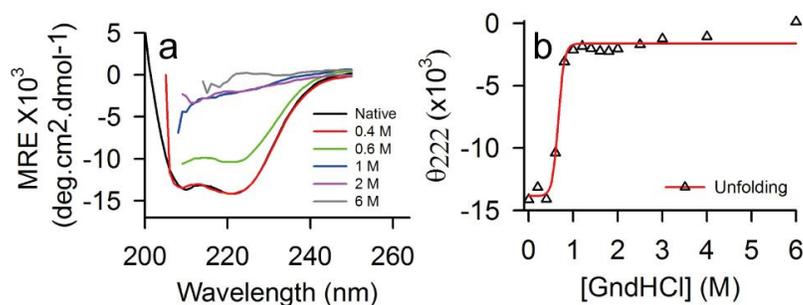


Figure 5.10. Secondary of STY3178 in presence GndHCl. (a) Shows the CD spectra of protein in native state (black) and in presence of different GndHCl concentrations 0.4 M (red), 0.6 M (green), 1M (blue), 2 M (magenta) and 6 M (grey). (b) Shows the change in ellipticity (θ_{222}) as a function of different concentrations of GndHCl.

Near UV-CD spectra show that protein retains its structure up to 0.4 M GndHCl with a broad shoulder in the range 250-280 nm (Figure 5.11a). The ellipticity of broad shoulder decreases at 0.6 M GndHCl and disappears above that. Fluorescence emission peak shows that the tertiary structure is stable up to 1 M GndHCl concentration for both 275 and 295 nm excitations (Figure 5.11b,c). It shows red shift of about 12 nm above 1 M GndHCl (Figure 5.11b,c). We observe a shoulder around 303 nm along with the red shifted peak in presence of 2 M and 6 M GndHCl for 275 nm excitation. This could be due to emission from tyrosine residue (s). There is no such shoulder observed at 295 nm excitation (Figure 5.11c). The difference spectrum (275-295 nm) in 1 M GndHCl is same as that of the native protein, indicating occurrence of FRET (Figure 5.11b). The intensity of difference spectra decreases in 2 M and 6 M GndHCl (Figure 5.11b). Similar observation is obtained for 280 nm excitation. We plot the changes in peak position as a function of GndHCl concentration for excitation wavelengths 275, 280 and 295 nm (Figure 5.11d). Peak positions are observed to be red shifted by 15-18 nm above 1 M GndHCl for all the three excitation wavelengths indicating unfolding of protein.

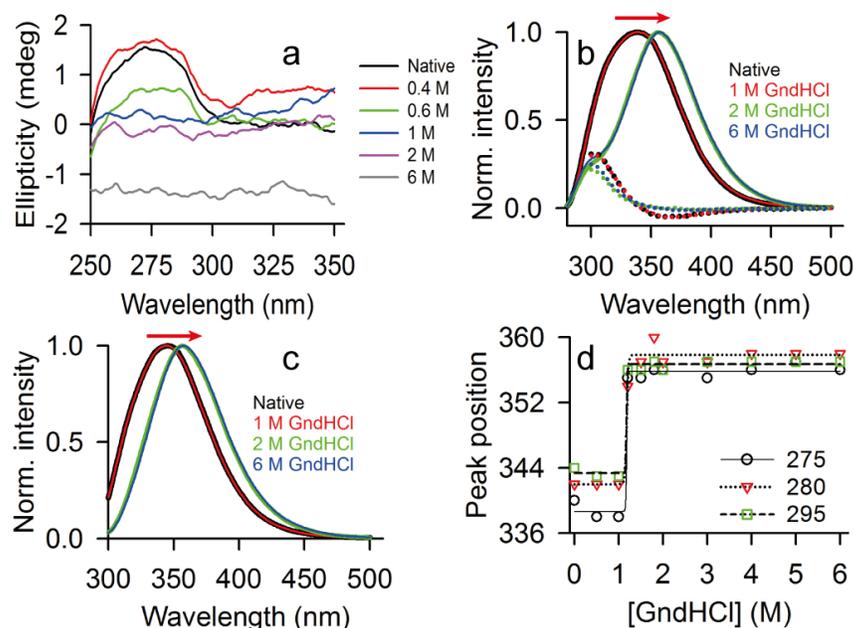


Figure 5.11. Tertiary structure of STY3178 in presence of GndHCl.

(a) Near UV-CD (250-350 nm) spectra of protein in presence of 0.4 M (red), 0.6 M (green), 1 M (blue), 2 M (magenta) and 6 M (grey) concentration of GndHCl are shown. (b) and (c) Show the fluorescence emission spectra of protein for excitations at 275 and 295 nm, respectively. The difference spectra for 275-295 nm are shown using dotted lines in (b). GndHCl concentrations are 1 M (red), 2 M (green) and 6 M (blue) for both (b) and (c). Native protein spectrum is shown in black for (a) to (c). (d) Shows the change in emission peak position in presence of different concentration of GndHCl during unfolding at 275 nm (black circle), 280 nm (red triangle) and 295 nm (green square) excitations.

The hydrodynamic size of STY3178 is measured similarly with increasing GndHCl concentration (Figure 5.12, triangle). Native hydrodynamic size (~6.5 nm) is retained up to 1 M GndHCl. There is an increase in size of about 1 nm in GndHCl concentration range 1 M to 3 M. The aggregation size increases further around 12 nm in presence of 5 M and 6 M GndHCl. This increase in hydrodynamic size indicates formation of higher aggregates due to unfolding similar to that of urea.

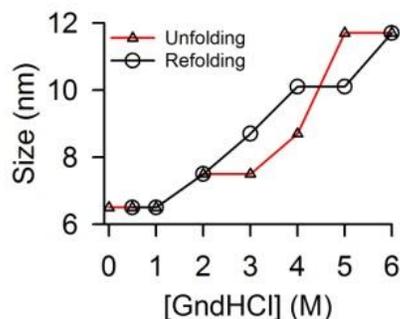


Figure 5.12. Hydrodynamic diameter of STY3178 in presence of GndHCl. The change in hydrodynamic size as a function of GndHCl concentration is shown for unfolding (triangle) and refolding (circle).

We monitor the refolding of STY3178 from unfolded state induced by 8 M GndHCl similarly. CD spectrum shows native-like protein signature for 40, 30 and 20 folds of dilution. Refolding of about 75% is achieved for 40 to 20 folds of dilution compared to that of native protein (Figure 5.13a). However, protein does not refold back upon 10 folds dilution which corresponds to 0.8 M GndHCl (Figure 5.13a). This observation is in agreement with the unfolding experiments where protein loses its structure at 0.8 M GndHCl. In equilibrium refolding measurements, we incubate the unfolded protein in lower GndHCl concentration for longer time. We find that the protein regains native-like CD signature upon lowering GndHCl concentration to 0.4 M (Figure 5.13b). It is partially refolded at 0.6 M and remains unfolded for 1 M, 2 M and 6 M GndHCl (Figure 5.13b). The plot of θ_{222} against GndHCl concentration shows decrease in ellipticity below 0.8 M GndHCl, indicating refolding of protein (Figure 5.13c). A comparison of θ_{222} for unfolding and refolding is shown in Figure 5.13c (inset). It indicates no hysteresis during the process in presence of GndHCl.

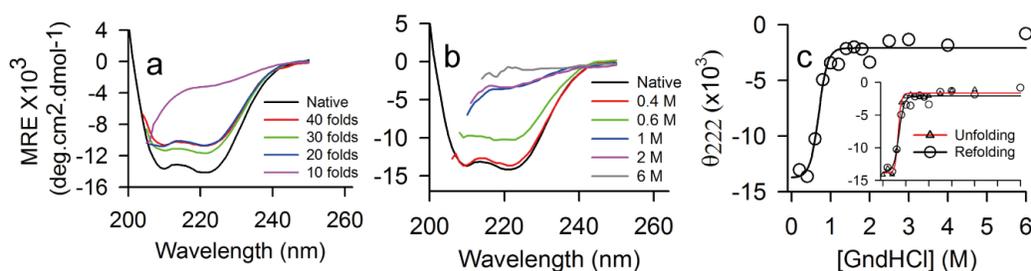


Figure 5.13. Protein secondary structure upon refolding from GndHCl unfolded state. (a) Shows the CD signature upon dilution of GndHCl by 40 (red), 30 (green), 20 (blue) and 10 (magenta) folds. (b) CD spectra of refolded protein for different GndHCl concentrations 0.4 M (red), 0.6 M (green), 1 M (blue), 2 M (magenta) and 6 M (grey). Native protein spectrum is shown in black for both (a) and (b). (c) Represents the plot of θ_{222} against GndHCl concentration. The inset shows the plot of θ_{222} for unfolding and refolding, indicating no hysteresis during the process in presence of GndHCl.

of ellipticity (θ_{222}) as a function of different concentration of GndHCl during refolding of protein. The inset shows a comparative plot of θ_{222} versus GndHCl concentration for unfolding (red) and refolding (black).

Near UV-CD spectra of protein upon reducing GndHCl to 0.4 M and 0.6 M indicate a native-like structure as a result of refolding (Figure 5.14a). It remains unfolded in presence of 1 M, 2 M and 6 M GndHCl (Figure 5.14a). We find a blue shift in fluorescence spectra of protein upon lowering GndHCl concentration to 0.4 M for 275 and 295 nm excitations (Figure 5.14b,c). Spectra remain red shifted in presence of both 1 M and 6 M GndHCl (Figure 5.14b,c). The difference spectrum (275-295 nm) upon lowering GndHCl concentration to 0.4 M, has a shoulder around 303 nm, similar to the native protein (Figure 5.14b). This shows that FRET occurs in the refolded protein similar to native state. In presence of 1 M GndHCl, similar shoulder is observed around 303 nm. However, its intensity is less compared to that of protein in 0 M and 0.4 M GndHCl (Figure 5.14b). The fluorescence emission results for 280 nm excitation are similar to that of 275 nm excitation. The plot of peak position as a function of GndHCl concentration clearly shows that refolding starts below 1 M GndHCl for all three excitation wavelengths 275, 280 and 295 nm (Figure 5.14d). We also monitor the hydrodynamic diameter of STY3178 upon reducing GndHCl concentration (Figure 5.12). The size reduces from 12 nm in presence of 6 M GndHCl to around 10 nm in 5 M and 4 M GndHCl (Figure 5.12). This decreases further in 4 M GndHCl. Native-like hydrodynamic size is finally observed in 1 M and below GndHCl (Figure 5.12). Thus, unfolding of STY3178 in presence of GndHCl is also completely reversible.

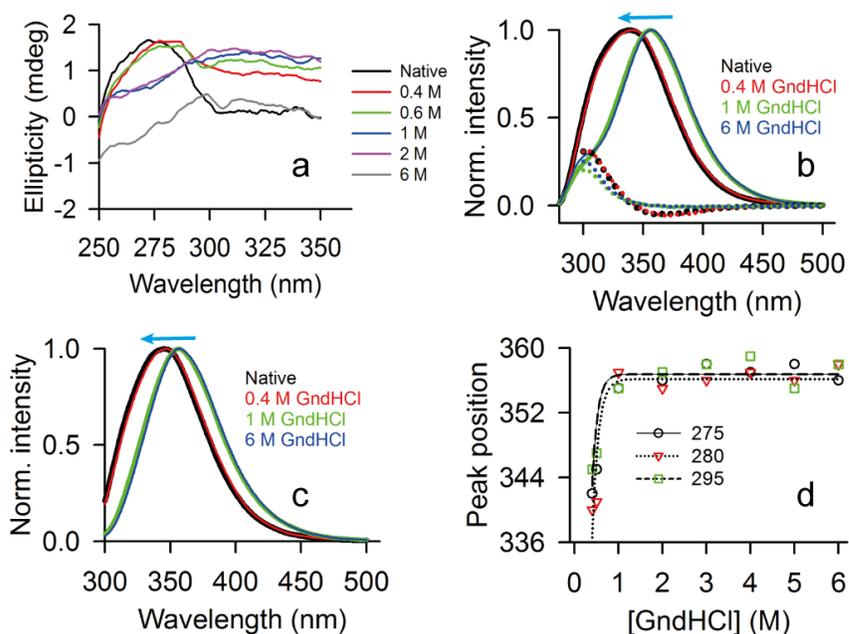


Figure 5.14. Tertiary structure upon protein refolding from GndHCl unfolded state. (a) Shows the near UV-CD signature of STY3178 upon decreasing GndHCl concentration to 0.4 M (red), 0.6 M (green), 1 M (blue), 2 M (magenta) and 6 M (grey) from 8 M. (b) Fluorescence emission spectra upon refolding for GndHCl concentrations 1 M (red), 2 M (green) and 6 M (blue). Native protein spectrum is shown in black for both (a) and (b). (c) Shows the plot of emission peak position upon excitation at 275 nm (black circle), 280 nm (red triangle) and 295 nm (green square) during refolding.

5.3.4. Effect of SDS

The secondary structure of STY3178 in far UV (200-250 nm) region is monitored in presence of SDS using CD. The spectrum of protein in SDS is compared with that of native state (Figure 5.15a). There are changes in the ellipticity values of two minima at 209 and 222 nm in presence of 69.4 mM (2 %) SDS (Figure 5.15a). The spectrum is not similar to that of an unfolded protein. However, helix content decreases to 34.7 % in presence of SDS compared to that (44 %) of native protein.

The plots of ellipticity values at 209 and 222 nm as a function of increasing SDS concentration are shown in Figure 5.15b. Ellipticity at 222 nm decreases slightly and that at 209 nm increases in presence of 3.5 mM (0.1 %) SDS. These ellipticity values remain almost unchanged up to 173 mM SDS concentration.

Near UV-CD spectra of protein in presence of SDS are found to be different from that of native protein (Figure 5.15c). STY3178 has a broad shoulder in near UV-CD as reported in chapter 2. Ellipticity of this broad shoulder decreases even at lowest concentration of 0.69 mM (0.02 %) SDS (Figure 5.15c). This remains unchanged for other higher concentrations of 3.5 mM (0.1%) and 17.4 mM (0.5%) SDS. This lowering of signal could be due to interaction of SDS with the aromatic residues of protein, which contribute to the near-UV CD spectral signature.

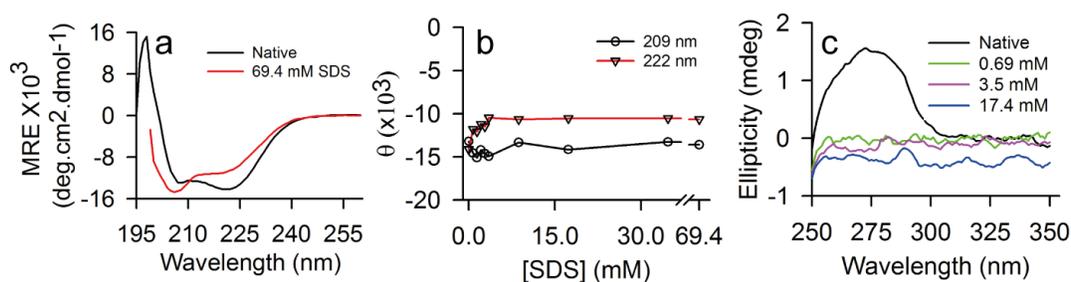


Figure 5.15. Effect of SDS on secondary and tertiary structure of STY3178. (a) CD spectra of protein are shown in absence (black) and presence (red) of 69.4 mM SDS. (b) The plots of ellipticity at 209 nm (circle) and 222 nm (triangle) versus concentration of SDS are shown. (c) Shows the near UV (250-355 nm) CD signature of STY3178 in presence of SDS with concentrations 0.69 mM (green), 3.5 mM (magenta) and 17.4 mM (blue). The native protein spectrum is shown in black.

Steady state fluorescence emission in presence of SDS indicates no substantial change in peak position for excitations at 275, 280 and 295 nm. Figure 5.16a shows the emission spectra of protein in SDS for excitation wavelength 280 nm. The emission intensity decreases with increase in SDS concentration for all the excitation wavelengths (Figure 5.16b). This suggests that the tertiary structure of the protein may be affected in presence of SDS, however, the α -helical nature is retained as observed in far UV-CD.

The hydrodynamic diameter of protein in presence of SDS remains same as that (~ 6.5 nm) of the native protein up to 8.7 mM (0.25 %) SDS (Figure 5.16c). It decreases to about 5.6 nm in 17.4 mM SDS. In presence of 69.4 mM (2%) SDS, it further decreases to about 4.2 nm (Figure 5.16c). We reduce SDS in protein and observe that the hydrodynamic size increases with decrease in SDS concentration (Figure 5.16c). The native-like size (~ 6.5 nm) is found upon reducing SDS to 8.7 mM (0.25 %). The reduced hydrodynamic diameter (~ 4.2 nm) at 69.4 mM SDS is similar to that of smaller proteins such as soyabean trypsin inhibitor and

carbonic anhydrase with molecular weights 20 KDa and 29 KDa, respectively.³² This suggests possible dissociation of STY3178 trimer into monomer above a certain SDS concentration.

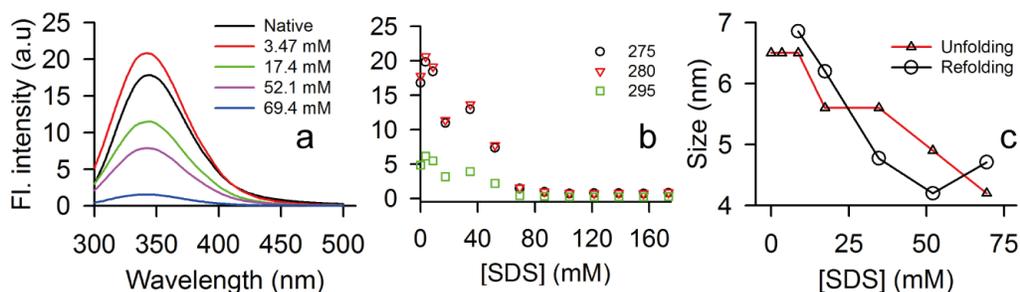


Figure 5.16. Fluorescence emission and hydrodynamic size in presence of SDS.

(a) Shows the fluorescence emission spectra of STY3178 in presence of SDS with concentrations 3.47 mM (red), 17.4 mM (green), 52.1 mM (magenta) and 69.4 mM (blue). The spectrum of native STY3178 is shown in black. (b) The changes in fluorescence emission intensity as a function of SDS concentration are shown for excitation wavelengths 275 (black), 280 (red) and 295 nm (green). (c) The hydrodynamics size of protein shown for increasing (triangle) and decreasing (circle) concentration of SDS.

We next perform size exclusion chromatography (SEC) to confirm the dissociation of STY3178 trimer in presence of SDS. DLS measurement shows that STY3178 remains as trimer in 3.47 mM SDS and dissociates into monomer in presence of 69.4 mM SDS. Thus, we perform SEC for STY3178 in presence of 3.47 mM (0.1 %) and 69.4 mM (2 %) SDS. Figure 5.17a shows the chromatogram of protein elution from the column in presence of 3.47 mM and 69.4 mM SDS. The elution volume of protein in 3.47 mM SDS is around 120 ml and matches with that of native STY3178, reported in chapter 3. This confirms that STY3178 remains in trimeric state in 3.47 mM SDS. However, protein in 69.4 mM SDS elutes at a volume of 160 ml (Figure 5.17a, solid). We compare this elution volume with that of several other standard proteins eluted from the same superdex 75 column and find elution at 160 ml corresponds to a protein with molecular weight ~23 KDa (Figure 5.17b). This confirms SDS-induced dissociation of trimeric STY3178 into its monomers.

ITC measurements are performed to determine the interaction of STY3178 with SDS. The thermogram of SDS titration into STY3178 solution at 298 K is shown in Figure 5.17c. The corresponding isotherm is represented in Figure 5.17d. ITC isotherm shows two regions during the titration (Figure 5.17d). In the first region, up to 100 molar ratio at 14th injection, the

final SDS concentration is 33.2 mM (0.95 %). In this SDS concentration range, the hydrodynamic diameter decreases from native size (~6.5 nm) to around 5.6 nm, as observed in DLS (Figure 5.16c). The second region in ITC isotherm corresponds to SDS concentration of about 63.2 mM (1.8 %) at the final injection (Figure 5.17d). DLS shows further reduction in hydrodynamic diameter from 5.6 nm to 4.2 nm for this second region of SDS concentration (Figure 5.16c). The change in enthalpy upon SDS binding is comparatively low which probably indicates predominance of hydrophobic interaction of SDS with STY3178.

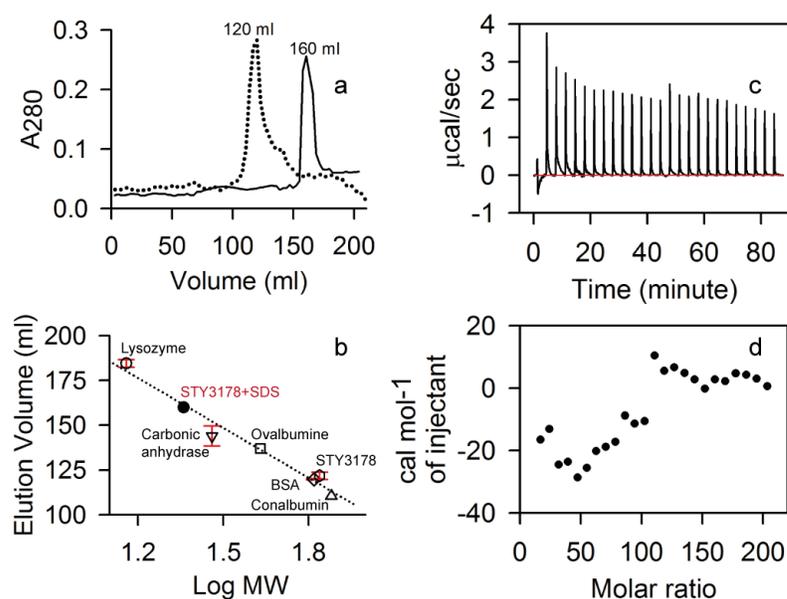


Figure 5.17. SEC and ITC in presence of SDS. (a) Shows the size exclusion chromatogram of STY3178 elution in presence of 3.47 mM (dotted) and 69.4 mM (solid) SDS. (b) The calibration curve of superdex75 column (also reported in chapter 3) showing the plot of elution volume versus logarithm of molecular weight of different standard proteins lysozyme (14.4 KDa, circle); carbonic anhydrase (29 KDa, inverted triangle); ovalbumin (43 KDa, square); BSA (66 KDa, diamond) and conalbumin (75 KDa, triangle). STY3178 in absence and presence of SDS are shown using hexagon and filled circle, respectively. The thermogram and corresponding isotherm of titration of 500 mM SDS into STY3178 solution are shown in (c) and (d), respectively.

5.4. Discussion

Our experimental results show that STY3178 can refold back to its native-like structure from an unfolded state induced by urea and GndHCl. This reversibility in unfolding can be utilized to calculate the free energy changes associated with protein folding.¹⁵⁸ We estimate the free energy change (ΔG_i) in presence of urea and GndHCl as described in method section. Protein unfolding in presence of urea suggests existence of some state at 0.6 M urea concentration, which is more α -helical, compared to the native state (Figure 5.4a,b). A two-state model (Native \rightleftharpoons Denatured) is assumed to calculate the free energy change. Folding constant (K) is estimated using the fraction of folded protein (α_i) at a given concentration of any chemical agent. ΔG_i is then calculated using K values in equation, $\Delta G = -RT \ln K$. A linear dependence of ΔG_i over urea concentration is observed (Figure 5.18a). The change in free energy (ΔG_0) at zero concentration of urea is estimated about -24.3 kcal/mol. Similarly, ΔG_i as a function of different concentrations of GndHCl is estimated and value of ΔG_0 obtained upon extrapolation of the plot to zero molar concentration of GndHCl is about -21 kcal/mol (Figure 5.18a). We observe a different phenomenon in presence of SDS. It does not unfold STY3178 unlike urea and GndHCl. The two characteristic minima are observed even in presence of 69.4 mM (2 %) SDS (Figure 5.15a). We find only ~9 % decrease in helix content of protein in SDS. However, a reversible dissociation of trimer into monomer occurs in presence of 69.4 mM (2 %) SDS (Figure 5.16c). This transition of trimer into monomer could be define by $N_3 \rightleftharpoons 3N$, where N is the folded protein monomer. We calculate the corresponding ΔG_0 associated with this transition using the ellipticity value at 222 nm as described in methods and obtained a value of about -13.35 kcal/mol in presence of SDS (Figure 5.18b).

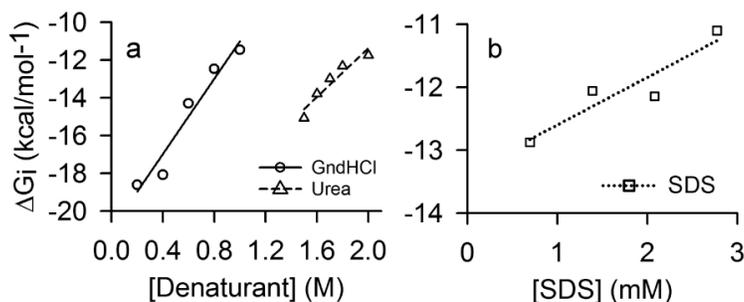


Figure 5.18. The change in free energy (ΔG_i) of STY3178 as a function of concentration of chemical agents is shown for (a) urea (triangle) and GndHCl (circle); (b) SDS (square).

5.5. Conclusion

The prime finding of this study is the reversible chemical unfolding capability of STY3178 in presence of urea and GndHCl. In presence of SDS, the trimeric protein dissociates into monomer. This dissociation is reversibly such that upon reducing SDS concentration, trimer is reformed. We report the changes in free energy associated with chemical unfolding in presence of urea and GndHCl as well as trimer dissociation into monomer in SDS.

Chapter 6 | Function of STY3178

6.1. Introduction

We have studied the protein, STY3178 from CT18^{3, 5} strain of *Salmonella* Typhi (*S.*Typhi). This strain is reported to be multidrug resistant.^{3, 5} STY3178 is a member of yfdX protein family. yfdX protein occurs in a number of multidrug resistant bacteria. This protein is encoded by yfdX gene. Expression of this gene is reported to be up-regulated by a multidrug response regulator EvgA in *E.coli*.^{19, 20, 33} This EvgA regulator also modulates the expression of other genes involved in acid tolerance response³⁵ (ATR), high osmolarity¹³⁷⁻¹³⁹ and efflux transport mechanism²⁰ along with yfdX. It is the only gene which also gets co-expressed when EvgA is over produced in *E.coli*.¹⁹ This enhanced expression and ubiquitous presence of yfdX in drug resistant bacteria suggest that they might not be completely non-functional. It raises a question whether yfdX protein has got any relation with drug resistance of bacteria.

We have shown in chapter 4 that STY3178 undergoes reversible thermal unfolding. It has the ability to completely refold back to native-like state even after heating at 80°C keeping the trimeric assembly unchanged. Proteins with high thermal stability and reversibility in unfolding can act as chaperones. For example, heat shock proteins (Hsps) such as IbpB¹⁵⁹ and GroES¹⁶⁰ show reversible thermal unfolding. Other types of proteins capable of undergoing reversible unfolding are periplasmic Skp¹⁶¹ and cytoplasmic SlyD¹⁶² and SecB¹⁶³ from *E. coli*. All these proteins act as chaperones in the cell. This leads us to investigate if STY3178 possess chaperone-like property since it shows reversibility in thermal unfolding along with stability.

In this chapter, we determine the interaction of STY3178 with antibiotics. We have chosen three antibiotics to which CT18^{3, 5} strain of *S.* Typhi is resistant.^{30, 32-34} These are ciprofloxacin, rifampin and ampicillin. We monitor the interaction of STY3178 with these three antibiotics using biophysical techniques such as steady state fluorescence and isothermal titration calorimetry (ITC). The effect of antibiotics on protein structure is determined using circular dichroism (CD), steady state fluorescence, dynamic light scattering (DLS) and nuclear magnetic resonance (NMR) experiments. We further predict the function of STY3178 using bioinformatics tool and perform chaperone activity assay using two different substrate proteins insulin and alcohol dehydrogenase. Our finding indicates that STY3178 has relevance in bacterial survival.

The organization of the chapter is as follows: section 6.2 covers materials and methods employed for functional characterization of STY3178. This section is further classified into two sub-sections namely interaction with antibiotics and chaperone activity assay. In section 6.3, results related to interaction of STY3178 with antibiotics and chaperone activity assay are reported. Section 6.4 covers the discussion on antibiotic binding residues and possible cellular localization of the protein. In section 6.5, conclusion of the overall finding is given.

6.2. Materials and methods

6.2.1. Interaction with antibiotics

6.2.1.1. Circular dichroism measurement in presence of antibiotics

CD measurements are carried out in Jasco J-815 CD spectrometer at 20°C. Far UV (200-250 nm) and near UV (250-330 nm) CD spectra of protein are collected using 1 mm and 10 mm path-length quartz cells, respectively. Protein concentration used for far UV and near UV-CD is 10 μM and 30 μM respectively. Data are collected in absence and presence of antibiotics. Concentration of antibiotics used for these measurements are: i) 5, 10, 20, 30 and 50 μM for ciprofloxacin (Cpx); ii) 10, 20, 50, 100 and 200 μM for rifampin (Rfp); iii) 60, 200 and 300 μM for ampicillin (Amp). CD spectra of isolated amino acids are recorded in the near UV region (250-330 nm) for 400 μM tyrosine (Tyr), 1000 μM phenylalanine (Phe) and 400 μM tryptophan (Trp) in presence of antibiotics. The concentration of antibiotics used with isolated amino acids is 50 μM Cpx, 100 μM Rfp and 200 μM Amp. All reported spectra are an average of three scans and background corrected by subtracting the data of buffer. We calculate the percentage of helix for protein in presence of these three antibiotics similarly (described in

chapter 2) using the relation^{51, 69}, $f_H = \frac{([\theta]_{222} - 3,000)}{(-36,000 - 3,000)}$.

In another set of CD measurement, we determine the thermal stability of the protein bound to antibiotics. STY3178 (5 μM) is heated at a rate of 4°C/min in presence of 25 μM Cpx, 100 μM Rfp and 100 μM Amp. The ellipticity values at 222 nm (θ_{222}) is measured in the temperature range 20°C to 70°C using a 3 mm path length cell. We calculate the fraction of the folded protein (f_N) at each temperature using the θ_{222} value in the equation⁹⁴, $f_N = (\theta_D - \theta_T)/(\theta_D - \theta_N)$, as described in chapters 4. θ_N and θ_D denote the respective ellipticity values of folded and unfolded protein. θ_T is the ellipticity at each temperature. We

determine the enthalpy of unfolding by calculating the folding constant (K) using the folded protein as stated in chapter 5.¹⁶⁴ We consider the linear portion of the data with maximum slope and fit it with equation of straight line, $y = mx + c$, where m denote the slope.¹⁶⁴ The enthalpy is calculated by multiplying the slope of linear fit with universal gas constant ($R=1.98 \text{ cal mol}^{-1}$).

6.2.1.2. Dynamic Light Scattering (DLS)

DLS measurement of STY3178 is performed in presence of antibiotics using Nano-S Malvern instrument at 20°C. Laser of wavelength 632.8 nm with measuring angle 173° is used during the measurement (as reported in chapter 2). Hydrodynamic diameter of protein-antibiotic complex is measured in presence of 50 μM Cpx, 200 μM Rfp and 200 μM Amp. Each reported measurement is a mean of five successive counts. All samples are filtered through 0.22 μm syringe filter (Millipore) prior to each measurement.

6.2.1.3. Steady state fluorescence measurement

Steady state fluorescence experiments are carried out in Jobin Yvon Horiba Fluorolog with 2 mm slit width. 10 μM protein is excited at wavelengths 257, 275, 280 and 295 nm with increasing concentrations of all three antibiotics Cpx, Rfp and Amp. The concentration of antibiotics used are as follows: (i) 2, 5, 10, 20, 30, 40 and 50 μM for Cpx, (ii) 10, 20, 50, 100 and 150 μM for Rfp and (iii) 60, 100, 150, 200, 250, 300, 400, 500 and 1000 μM for Amp. Each data is averaged over two sets of measurements.

We analyze the data of protein fluorescence quenching in presence of antibiotics using Stern-Volmer equation⁴⁷⁻⁴⁹, $(F_0 - F)/F = K_{SV}[Q]$, where F_0 and F are the respective fluorescence intensities in absence and presence of antibiotics at concentration Q . K_{SV} denotes the Stern-Volmer quenching constant. The modified form of Stern-Volmer equation⁴⁷⁻⁴⁹, $\log[(F_0 - F)/F] = \log K + n \log[Q]$, is used to estimate the binding constant (K) and number of binding sites (n). Dissociation constant (K_d) is calculated as a reciprocal of K . ΔG is obtained using the equation, $\Delta G = -2.303RT \ln K$.⁴⁹

6.2.1.4. Isothermal Titration Calorimetry (ITC)

We perform the ITC measurements using MicroCal iTC200 calorimeter (GE healthcare) for confirming the interaction of STY3178 with antibiotics Cpx, Rfp and Amp. Dialysis of

protein is performed against the buffer containing 30 mM phosphate (pH7), 150 mM NaCl and 1mM PMSF. Concentration of protein used in the cell during all titration measurements is 300 μ M. Concentrations of Cpx, Rfp and Amp loaded in the syringe are 20 mM, 10 mM and 100 mM, respectively. All titrations of protein-antibiotics and buffer-antibiotics are performed in buffer having composition 30 mM phosphate (pH7), 150 mM NaCl and 1mM PMSF at 298 K temperature. Reference power used in ITC is 10 μ cal/s. Each antibiotic titration is performed for total 30 injections with an initial delay of 60 seconds and constant stirring at 50 rpm. First injection of 0.4 μ l is carried out for a time period of 0.8 seconds. Rest of the 29 injections of 0.6 μ l each are added over a time period of 1.2 seconds. Spacing between each injection is 200 seconds. We subtract the heat of dilution of respective antibiotics from that of the protein-antibiotics and plot the integrated data after subtraction using MicroCal origin. Cpx and Amp isotherms are fitted with sequential binding models whereas Rfp data is fitted with single site binding model for estimation of binding constant (K), enthalpy (ΔH) and entropy (ΔS) changes. Dissociation constant (K_d) is calculated as a reciprocal of K . The change in free energy (ΔG) associated with each antibiotic binding is calculated using the experimental ΔH and ΔS values in Gibbs equation, $\Delta G = \Delta H - T\Delta S$.

6.2.1.5. NMR relaxation measurement of protein-antibiotic complex

The measurements of longitudinal (T_1) and transverse (T_2) relaxation time of STY3178 bound to antibiotics are performed similarly as that of native protein (in chapter 3) at 30°C in 600 MHz Varian spectrometer. 400 μ M STY3178 is used for each relaxation measurement. Concentration of Cpx and Rfp used are 1600 μ M and that of Amp is 3200 μ M. Recycle delay (d_1) of 8 seconds is used for T_1 measurement in presence of Rfp and Amp. Free induction decay (FID) is collected for Rfp and Amp bound protein with time delay points 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0, 1.2, 1.5 and 1.8 seconds during T_1 measurement. In presence of Cpx, T_1 measurement is performed using d_1 of 9 seconds and FID is collected for delay points 0.01, 0.05, 0.1, 0.2, 0.3, 0.6, 0.8, 1.0, 1.2, 1.5 and 1.8 seconds. In case of T_2 measurements, 4 seconds d_1 delay is used for Rfp and Amp bound protein and that of Cpx bound state is 6 seconds. The delay points during FID collection are 0.01, 0.03, 0.05, 0.07, 0.09, 0.11, 0.13, 0.15 and 0.17 seconds for T_2 measurements in presence of all three antibiotics. All data are collected as 256 scans. Data processing and analysis for estimation of T_1 and T_2 are performed similarly as

described in chapter 3. The total rotational correlation time (τ_C) is calculated using T_1 and T_2

values in the equation^{42, 43},
$$\tau_C = \frac{1}{4\pi V_N} \sqrt{6 \frac{T_1}{T_2} - 7}.$$

6.2.1.6. HSQC measurements in presence of antibiotics

¹⁵N-labelled STY3178 is prepared following the protocol given in chapter 4. We perform ¹H-¹⁵N heteronuclear single quantum coherence (HSQC) measurements in absence and presence of antibiotics using 600 MHz Varian spectrometer at 35°C. 400 μM singly labeled protein in a buffer (30 mM phosphate buffer, pH 7, 150 mM NaCl and 10 % D₂O), is titrated with Cpx, Rfp and Amp. Antibiotic concentrations used during titration are the following: (i) 10, 50, 100, 200, 400, 800 and 1600 μM for both Cpx and Rfp; (ii) 200, 400, 1100, 1600, 3200 and 6400 μM for Amp. All HSQC data are acquired as 64 scans, processed in NMRPIPE⁹⁵ and analyzed using NMRVIEW⁹⁶.

6.2.2. Chaperone activity assay

Chaperone activity assay is performed using two different substrate proteins namely human insulin and yeast alcohol dehydrogenase (ADH). Aggregation of 50 μM insulin in presence of 20 mM DTT is monitored by measuring the absorbance value at 360 nm (A_{360}) in absence and presence of STY3178.¹⁶⁵ A 10 mm path length cell is used for the measurement in BioSpectrometer (Eppendorf). STY3178 and insulin are equilibrated for 10 minutes at 42°C before adding DTT. Different molar ratios of STY3178 and insulin used during the assay are 0.1:1, 0.25:1 and 0.5:1. Similarly, we monitor the A_{360} value of STY3178 with time in presence of 20 mM DTT. The dead time of the reaction is ~12 seconds.

Aggregation of 1 μM ADH is monitored by measuring A_{360} value in absence and presence of STY3178 at 42°C in Shimadzu spectrophotometer. Both the proteins are equilibrated on ice for 5 minutes prior to measurement. The molar ratios of ADH to STY3178 used are 1:0.25, 1:0.5 and 1:1. Buffer used during the experiments contains 50 mM phosphate (pH 7), 250 mM NaCl and 1 mM PMSF.

6.3. Results

6.3.1. Antibiotic binding to STY3178

Interaction of STY3178 with three antibiotics, namely, ciprofloxacin (Cpx), rifampin (Rfp) and ampicillin (Amp) is probed using steady state fluorescence spectroscopy, isothermal titration calorimetry (ITC) and circular dichroism measurements. Fluorescence emission of protein is measured for 280 nm excitation with increasing concentration of the antibiotics (Q). The emission is quenched substantially in presence of all the antibiotics compared to that of the native protein (Figure 6.1). A slight shift in emission peak position is observed for Cpx and Rfp binding (Figures 6.1a,b) which is not considerable for Amp binding (Figure 6.1c). Similar results are obtained for 275 nm excitation. Fluorescence quenching constant K_{SV} is estimated using the Stern-Volmer equation⁴⁷⁻⁴⁹, $(F_0 - F)/F = K_{SV}[Q]$ as described in method section. The value of K_{SV} obtained from linear fitting of $(F_0 - F)/F$ versus $[Q]$ for Cpx, Rfp and Amp binding are $0.05 (\pm 2 \times 10^{-3})$, $0.26 (\pm 0.034)$ and $0.0003 (\pm 2 \times 10^{-5}) \text{ M}^{-1}$, respectively (Table 6.1). The binding constant (K) for each antibiotic is calculated using the modified form of Stern-Volmer equation⁴⁷⁻⁴⁹, $\log[(F_0 - F)/F] = \log K + n \log[Q]$ as reported in methods. The respective plots of $\log[(F_0 - F)/F]$ versus $\log[Q]$ for Cpx, Rfp and Amp are shown in Figures 6.1d-f. K values and number of binding sites (n) are obtained from the respective intercept and slopes of the curves fitted linearly (Table 6.1). K for Cpx, Rfp and Amp binding to protein are $0.04 (\pm 0.0023)$, $0.01 (\pm 0.004)$ and $0.0004 (\pm 4 \times 10^{-5}) \text{ M}^{-1}$, respectively (Table 6.1). Dissociation constants (K_d) are estimated from the reciprocal of K (Table 6.1). K_d values obtained for binding of Cpx, Rfp and Amp with STY3178 are $25.0 (\pm 1.2)$, $100.0 (\pm 20.5)$ and $2500.0 (\pm 370.5) \mu\text{M}$, respectively. We calculate the free energy change (ΔG) associated with each of these protein-antibiotic interactions using K values in the relation, $\Delta G = -2.303RT \ln K$. The respective ΔG values are -26.5 -22.8 $-14.6 \text{ kJ mol}^{-1}$ for Cpx, Rfp and Amp interactions (Table 6.1). Fluorescence data indicate Cpx and Amp bind to a single site with $n \approx 1$. However, Rfp has a possibility of binding in more than one site with $n \approx 1.5$. Overall, we find that Cpx interaction with the protein is strong followed by Rfp and Amp.

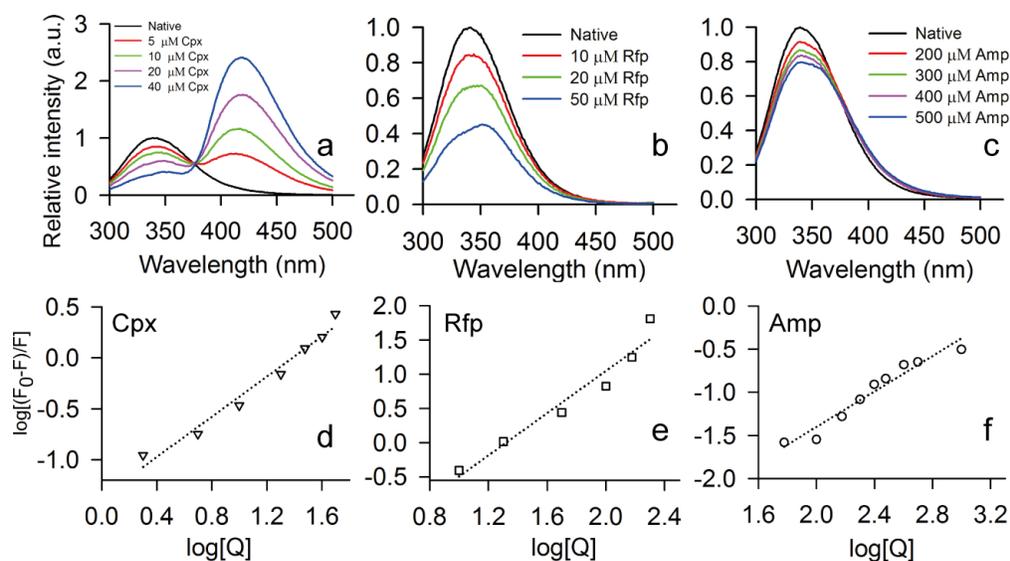


Figure 6.1. Interaction of STY3178 with antibiotics from fluorescence. (a), (b) and (c) Show the fluorescence emission spectra of STY3178 in presence of ciprofloxacin (Cpx), rifampin (Rfp) and ampicillin (Amp), respectively, for 280 nm excitation wavelength. The concentrations of antibiotics are as follows: 5 (red), 10 (green), 20 (magenta) and 40 (blue) μM for Cpx; 10 (red), 20 (green) and 50 (blue) μM for Rfp; 200 (red), 300 (green), 400 (magenta) and 500 (blue) μM for Amp. The spectrum of native protein is shown in black colour for (a)-(c). (e), (f) and (g) Show the plot of $\log[(F_0 - F)/F]$ against $\log[Q]$ for Cpx (inverted triangle), Rfp (square) and Amp (circle), respectively.

Table 6.1 STY3178-antibiotic interaction data from steady state fluorescence measurement

Antibiotic bound	K_{SV} (μM^{-1})	K (μM^{-1})	K_d (μM)	n	ΔG (kJ mol^{-1})
Cpx	0.05 (± 0.002)	0.04 (± 0.0023)	25.0 (± 1.2)	0.99	-26.5
Rfp	0.26 (± 0.034)	0.01 (± 0.004)	100.0 (± 20.5)	1.54	-22.8
Amp	0.0003 (± 0.00002)	0.0004 (± 0.00004)	2500.0 (± 370.5)	1.03	-14.6

We substantiate the interaction of STY3178 with antibiotics using ITC measurements. The thermograms for STY3178 interaction with Cpx, Rfp and Amp at 298 K are shown in Figures 6.2a-c. The corresponding binding isotherms are represented in Figures 6.2d-f. The isotherms are fitted with a sequential binding model for Cpx and Amp whereas Rfp data is fitted to a single site binding model. The number of binding sites chosen for fitting Cpx and Amp isotherms is six and two, respectively. The binding constant (K), enthalpy changes (ΔH) and entropy changes (ΔS) are obtained by fitting the data with above stated binding models. We

consider the first two binding constant values for Cpx to be relevant. Similarly, first binding constant of Amp is considered to be significant. K values for binding of Cpx are $0.58 (\pm 0.04)$ and $0.423 (\pm 0.0079) \mu\text{M}^{-1}$ with respective K_d values of 1.72 and $2.36 \mu\text{M}$ (Table 6.2). The values of K for Rfp and Amp binding are $0.0151 (\pm 0.00911)$ and $0.00744 (\pm 0.0037) \mu\text{M}^{-1}$, respectively (Table 6.2). The values of K_d calculated as a reciprocal of K for Rfp and Amp binding are 66.23 and $134.41 \mu\text{M}$, respectively (Table 6.2). ΔH and $T\Delta S$ values for binding of antibiotics are tabulated in Table 6.2. These values are used to calculate the free energy change (ΔG) following the relation, $\Delta G = \Delta H - T\Delta S$. The calculated ΔG values are shown in Table 6.2 along with other binding parameters. The binding parameters obtained from ITC confirm the interaction of STY3178 with Cpx, Rfp and Amp with K_d values lying in micro molar (μM) range for all antibiotics. Cpx binding is found to be stronger than Rfp and Amp. ΔG of Cpx bound protein is more favourable followed by Rfp and Amp. This trend of K_d and ΔG values from ITC are similar to those obtained from steady state fluorescence measurements, although ITC data indicates much stronger binding of antibiotics with STY3178.

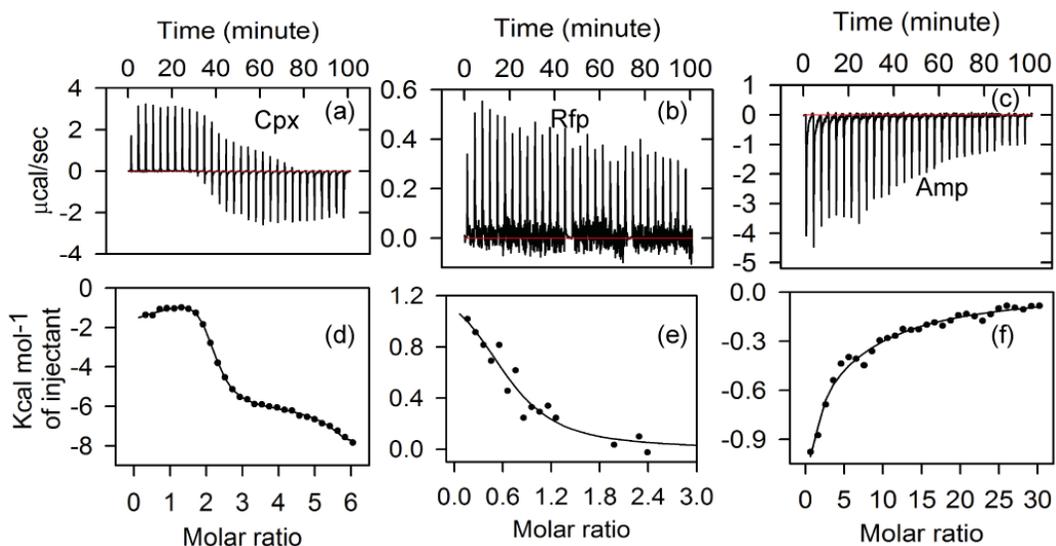


Figure 6.2. STY3178 interaction with antibiotics using ITC. The thermogram of STY3178 binding with (a) ciprofloxacin (Cpx), (b) rifampin (Rfp) and (c) Ampicillin (Amp). The respective fitted isotherms are shown in (d)-(f).

Table 6.2 STY3178-antibiotic binding parameters from ITC

Antibiotic bound	K (μM^{-1})	K_d (μM)	ΔH (kJ mol^{-1})	$T\Delta S$ (kJ mol^{-1})	ΔG (kJ mol^{-1})
Cpx	$K_1=0.58$ (± 0.04)	$K_{d1}=1.72$	$\Delta H_1=-6.49$ (± 0.55)	$T\Delta S_1=26.53$	$\Delta G_1=-33.02$
	$K_2=0.423$ (± 0.0079)	$K_{d2}=2.36$	$\Delta H_2=-4.03$ (± 1.13)	$T\Delta S_2=28.16$	$\Delta G_2=-32.19$
Rfp	$K_1=0.0151$ (± 0.00911)	$K_{d1}=66.23$	$\Delta H_1=6.00$ (± 1.62)	$T\Delta S_1=29.91$	$\Delta G_1=-23.91$
Amp	$K_1=0.00744$ (± 0.0037)	$K_{d1}=134.41$	$\Delta H_1=-6.54$ (± 1.07)	$T\Delta S_1=15.65$	$\Delta G_1=-22.19$

The binding parameters, K and K_d for interaction of several proteins¹⁶⁶⁻¹⁷² with different antibiotics and small molecules are tabulated in Appendix A. We observe K_d values to be in the range 1-100 μM for more than $\sim 80\%$ cases. Another $\sim 15\%$ belong to the range 100-200 μM and less than 5 % in the range 200-300 μM . The data from ITC measurements lie in the lowest K_d range which indicates stronger binding of STY3178 with antibiotics in the biologically relevant regime.

6.3.2. Structural changes upon antibiotic binding

Binding of antibiotics with STY3178 is confirmed from ITC and fluorescence. The structural changes in protein upon this binding is determined using CD, steady state fluorescence, DLS and NMR. The near UV-CD spectra of protein in antibiotic free and bound states are measured (Figures 6.3a-c). Native protein spectrum shows a broad shoulder in the range 250-280 nm as mentioned in chapter 2 as well. This broad shoulder indicates presence of aromatic residues in protein.⁵⁰ The changes in this broad shoulder are observed in presence of Cpx, Rfp and Amp. We find the extent of changes are different for different antibiotics and depends on their concentrations. For instance, two peaks start appearing near ~ 260 nm and ~ 280 nm as Cpx concentration increases (Figure 6.3a). These peaks are similar to Phe and Tyr fine structures reported earlier^{50, 51}. This suggests that Cpx interaction with protein occurs near Phe and Tyr residues. The perturbation of aromatic residues upon addition of Rfp to protein shows a similar trend like Cpx binding. The ellipticity values change for binding with different concentrations of Cpx (Figure 6.3a) and Rfp (Figure 6.3b) with similar peak positions. On the other hand, Amp binding to protein (Figure 6.3c) results in negligible change compared to native spectrum in near UV-CD region.

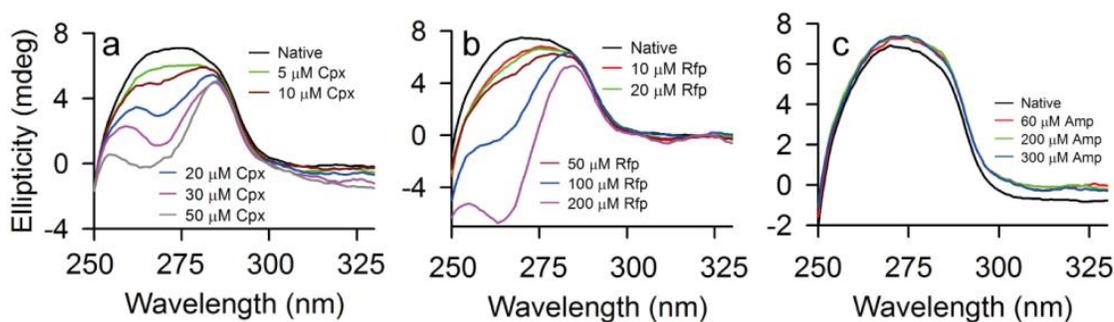


Figure 6.3. The near UV-CD spectra of antibiotic bound STY3178. (a), (b) and (c) Show the CD spectra of protein in presence of antibiotics in the near UV-range 250-330 nm. Different concentrations of (a) ciprofloxacin used are 5 μM (green), 10 μM (brown), 20 μM (blue), 30 μM (magenta) and 50 μM (grey); (b) rifampin are 10 μM (red), 20 μM (green), 50 (brown), 100 (blue), 200 μM (magenta) and (c) ampicillin are 60 μM (red), 200 μM (green) and 300 μM (blue). Spectrum of native protein is shown in black for (a)-(c).

The secondary structure of protein in far UV (200-250 nm) CD region is observed in presence of antibiotics. The CD spectra of STY3178 in presence of Cpx, Rfp and Amp remain similar to the native spectrum with two characteristic minima around 209 and 222 nm (Figure 6.4a). A slight decrease in ellipticity for Rfp and Amp bound protein is observed whereas Cpx bound protein shows almost no change. Overall helix content of the protein in presence of Cpx, Rfp and Amp are 44.5, 40.1 and 40.6 %, respectively. We compare the helical content of antibiotic bound protein with that of the native state (Table 6.3) and find only 3 % decrease in percentage of helix for Rfp and Amp bound structures. Protein in Cpx bound state has helical content almost same as the native STY3178 (Table 6.3). Thus, secondary structure of protein remains unaffected in presence of antibiotics.

Effect of antibiotics on the thermal stability of STY3178 is determined by heating it at a rate of 4°C/min. Figure 6.4b shows the plot of folded fraction (f_N) versus temperature in the range 20-70°C for native state as well as protein bound to Cpx, Rfp and Amp. All four sigmoid plots are overlapping, which indicates no substantial change in the thermal stability of protein in presence of antibiotics. We determine the enthalpy of unfolding by plotting the natural logarithm of folding constant ($\ln K$) with $1/T$, where T is the absolute temperature (Figure 6.4c). The enthalpy values are obtained by multiplying the slopes of linearly fitted plots with universal gas constant ($R = 1.98 \text{ cal mol}^{-1}$). The value of enthalpy for native protein is 131.19

cal/mol and that of Cpx, Rfp and Amp bound states are 110.78, 104.75 and 133.52 cal/mol, respectively.

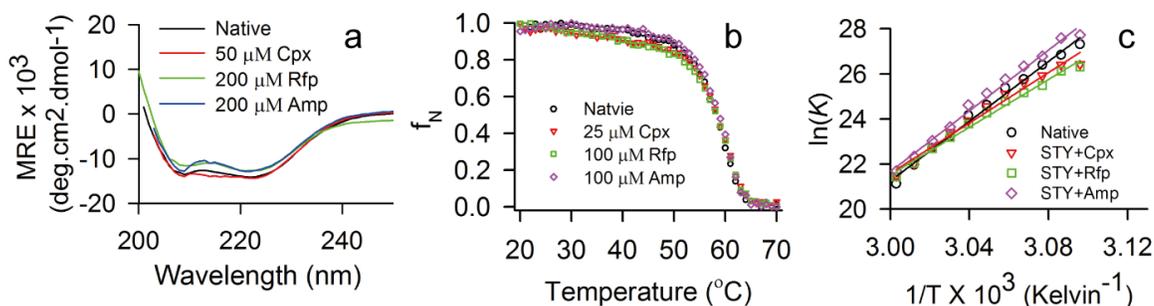


Figure 6.4. Secondary structure of STY3178 in antibiotic free and bound states. (a) CD spectra is shown in presence of 50 μM Cpx (red), 200 μM Rfp (green) and 200 μM Amp (blue) in the far UV region (200- 250 nm). The native protein spectrum is shown in black colour. **(b)** Shows the plot of folded fraction (f_N) as a function of temperature in the range 20-70°C in absence (black) and presence of antibiotics. **(c)** Shows the plot of $\ln K$ versus $1/T$ (Kelvin^{-1}) fitted linearly. The concentration of antibiotics in **(b)** and **(c)** are 25 μM Cpx (red), 100 μM Rfp (green) and 100 μM Amp (magenta).

Table 6.3. Percentage of helix for native and antibiotic bound STY3178.

State	Helix (%)
Native	43.9
Cpx bound	44.5
Rfp bound	40.1
Amp bound	40.6

6.3.3. Oligomeric state of antibiotic bound protein

The total rotation correlation time (τ_c) of STY3178 in presence of antibiotics is estimated using NMR relaxation measurements similar to that of native protein (reported in chapter 2). T_1 values for Cpx, Rfp and Amp bound protein are 1.98, 2.11 and 1.92 seconds, respectively. The respective T_2 values are 0.036, 0.033 and 0.034 seconds. τ_c is calculated using

the equation, $\tau_c = \frac{1}{4\pi V_N} \sqrt{6 \frac{T_1}{T_2} - 7}$ and the obtained values are 23.8, 25.9 and 24.2 ns for Cpx,

Rfp and Amp bound states, respectively. These values are comparable with that of native protein as shown in Table 6.4. Thus, we do not observe any substantial change in τ_c of protein in its antibiotic bound states. This suggests that the oligomeric state of protein is not perturbed upon antibiotic binding. The hydrodynamic radii (R_H) of these bound states of protein are

measured using DLS. The R_H is found to be around 3.25 nm for all the antibiotic bound protein (Table 6.4). This indicates the STY3178 interacts with the antibiotics in its trimeric state of oligomerization.

Table 6.4. Longitudinal (T_1), transverse (T_2) relaxation times, rotation correlation time (τ_c) and hydrodynamic radius (R_H) of native and antibiotic bound STY3178.

State	T_1 (s)	T_2 (s)	τ_c (ns)	R_H (nm)
Native protein	1.95	0.033	24.7	3.25
Cpx bound	1.98	0.036	23.8	3.25
Rfp bound	2.11	0.033	25.9	3.25
Amp bound	1.92	0.034	24.2	3.25

We further perform HSQC measurements of ^{15}N labeled STY3178 in presence of Cpx, Rfp and Amp. Native protein HSQC contains clustered peaks in the region 7.5-8.5 ppm along with well dispersed peaks (reported in chapter 4). The concentration of both Cpx and Rfp used during titration are 10, 50, 100, 200, 400, 800 and 1600 μM and that of Amp are 200, 400, 1100, 1600, 3200 and 6400 μM . We find HSQC spectra of bound states remain same as that of the native protein. There are only few regions in the HSQC which show either change in intensity or chemical shift value compared to the native protein spectrum. The enlarged view of these regions is shown in Figure 6.5. In region 1 (8.58-8.85 ppm), there is an appearance of new peak in presence of 1600 μM Cpx and 3200 μM Amp (Figure 6.5a1-d1). This region shows no change in presence of 1600 μM Rfp. Region 2 (9.97-10.2 ppm) corresponds to tryptophan chemical shift value (Figure 6.5a2-d2). There are two peaks in this region which are affected in presence of antibiotics. There is broadening of one peak \sim 10.2 in presence of Cpx whereas the second peak disappears in Rfp and Amp bound states. There are two peaks in region 3 (8.1-8.35 ppm) which disappears in presence of Cpx and Rfp compared to the native protein (Figure 6.5a3-c3). In presence of Amp, there is a lowering of peak intensity in region 3 (Figure 6.5d3). Region 4 (6.75-6.92 ppm) remains unaffected in presence of Cpx compared to that of native state (Figure 6.5a4,b4). Similarly, we find that peaks disappear for Rfp and Amp bound states this region (Figure 6.5c4,d4). Overall, NMR measurements confirm that the structure of protein does not change in presence of antibiotics, although few peaks are affected during the interaction.

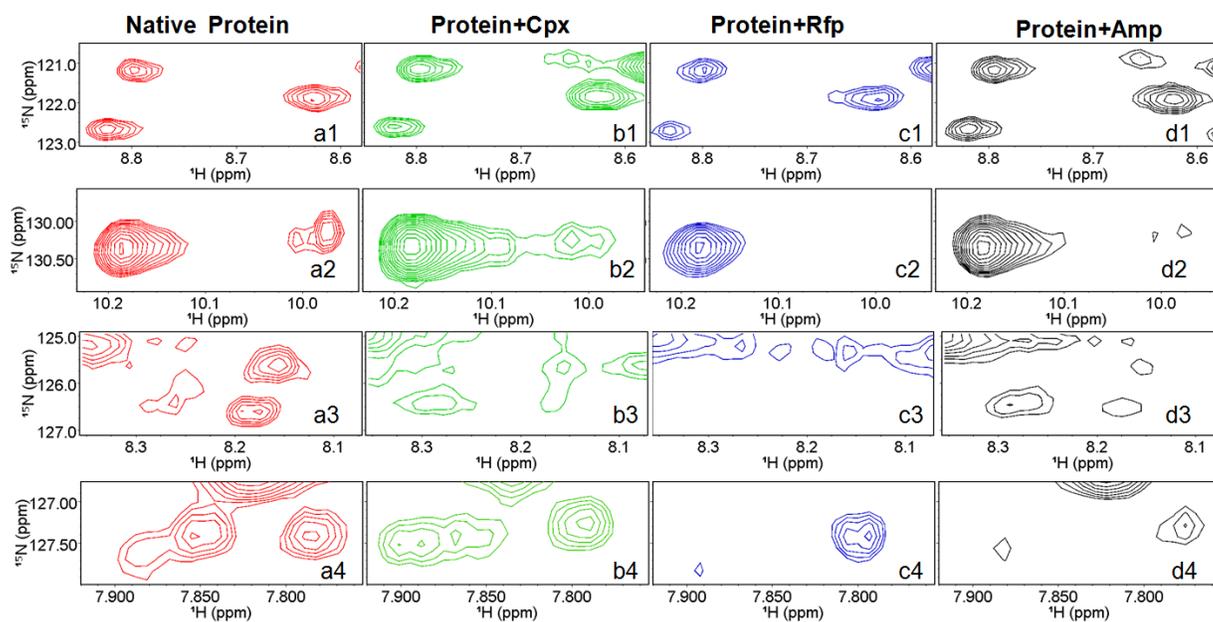


Figure 6.5. Enlarged view of HSQC regions for native and antibiotic bound STY3178. Panels a1-d1 in region 1 has 8.58-8.85 ppm range. Similarly, a2-d2, a3-d3 and a4-d4 represent regions 2 (9.97-10.2 ppm), 3 (8.1-8.35 ppm) and 4 (7.75-7.92 ppm), respectively. Native protein region is shown in a1-a4 (red). HSQC regions in presence of Cpx, Rfp and Amp are represented in b1-b4 (green), c1-c4 (blue) and d1-d4 (black), respectively.

6.3.4. Chaperone activity assay of STY3178

Protein showing stability at elevated temperature and reversibility in unfolding can act as a chaperone.^{159, 160, 173} Therefore, we first predict the function of STY3178 using the Bioinformatics tool MODexplorer⁴¹. This server relates the sequence, structure and functions of proteins with the query sequence. The output results show ~50 proteins having >50% structural similarity with STY3178 (Appendix B). Most of these proteins are helical similar to STY3178. Majority of these proteins are either chaperones or assist chaperone activity whereas some others have no reported function besides having similar structural folds. Other proteins are involved in cell cycle, transport, oxidoreductase activity, to name only a few (Appendix B). Our experimental observations suggest that STY3178 may perform chaperone activity.

We perform chaperone activity assay using insulin B-chain and alcohol dehydrogenase (ADH) as substrates in presence of STY3178 at 42°C *in vitro*. We find a steep rise in absorbance at 360 nm (A_{360}) due to formation of insulin aggregates in presence of DTT at 42°C, which saturates with time (Figure 6.6a) as reported in literature¹⁶⁵. There is a lowering of A_{360} in presence of STY3178. Figure 6.6a shows the plot of A_{360} as a function of reaction time

in absence (black) and presence of various STY3178 molar concentration ratios 0.1 (magenta), 0.25 (red) and 0.5 (green) to insulin concentration. There is decrease in A_{360} as STY3178 concentration increases in the reaction mixture. This indicates that STY3178 inhibits aggregation of insulin induced by DTT.

Similarly, we perform chaperone activity assay using ADH and observe an increase in its aggregation at 42°C (Figure 6.6b). Aggregation of ADH decreases with increases in concentration of STY3178 (Figure 6.6b). This indicates that prevention of aggregate formation is dependent on the amount of STY3178. In Figure 6.6b, the plot of A_{360} versus time is shown for STY3178 molar ratios 0.25 (red), 0.5 (green) and 1 (blue) to ADH. Similarly, we record A_{360} for STY3178 in presence of DTT only and find no aggregation at 42°C (Figure 6.6a, dark cyan). A gradual lowering of A_{360} for both insulin and ADH with increasing STY3178 concentration shows its capability to prevent the aggregation. Thus, STY3178 is capable of acting as a chaperone *in vitro*.

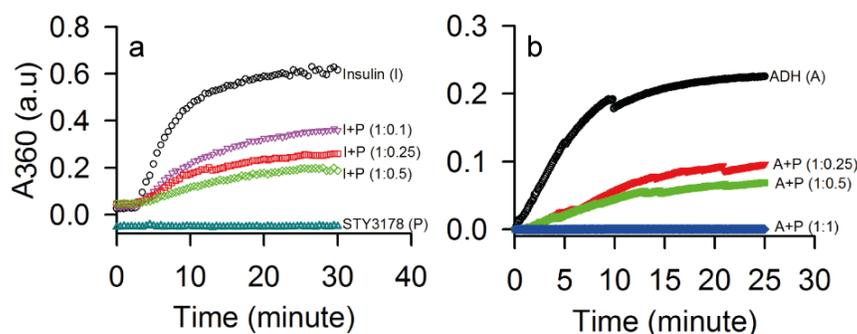


Figure 6.6. Chaperone activity assay. Plots of A_{360} with time are shown for thermal aggregation of (a) Insulin (I) induced by DTT (black) and (b) ADH (black) at 42°C. The assay is performed for insulin to STY3178 (P) molar ratios 1:0.1 (magenta), 1:0.25 (red) and 1:0.5 (green) in (a) whereas that for ADH to STY3178 are 1:0.25 (red), 1:0.5 (green) and 1:1 (blue) in (b). The plot of STY3178 in presence of DTT is shown using dark cyan colour in (a).

6.4. Discussion

6.4.1. Interaction of STY3178 with antibiotics

We demonstrate that STY3178 binds three different small antibiotic molecules. Both steady state fluorescence and ITC results indicate strong binding for Cpx, followed by Rfp and Amp. K_d values for Cpx binding measured from ITC indicate stronger binding compared to that

obtained using fluorescence. Amp binding is well captured using ITC, which is weaker in fluorescence measurements. However, for Rfp binding, K_d value is similar in both the measurements. The stoichiometries of binding events in ITC and fluorescence results also differ. These discrepancies may be due to the fact that fluorescence measurements depend on the binding of antibiotics near the fluorophores, while ITC is independent of fluorophore environment. The binding modes for all these three antibiotics are different as indicated by different models used for fitting the isotherms.

Antibiotic binding do not affect the secondary structure of STY3178 as found in far UV-CD. The oligomeric state of protein also remains unaltered as observed in DLS and NMR measurements. However, they show tertiary structural changes as observed in near UV-CD (Figures 6.3a-c) indicating involvement of aromatic residues of protein in antibiotic interaction. Near UV-CD spectra of isolated aromatic residues in presence of these antibiotic molecules are also monitored. Spectra of all three isolated amino acids Phe, Tyr and Trp are recorded in near UV-CD region (Figure 6.7a). Isolated Tyr has a shoulder around 270 nm whereas Phe and Trp have low ellipticity in near UV region. However, all these spectra changes in presence of Cpx (Figure 6.7b). Isolated Tyr and Trp show pronounced changes in presence of Rfp whereas that of Phe has minimal changes (Figure 6.7c). These changes are qualitatively similar to that observed for the protein. The nature of perturbation in presence of Amp is very small for isolated Phe and Trp (Figure 6.7d) but for isolated Tyr, it is again qualitatively similar to the change observed in STY3178 (Figure 6.3).

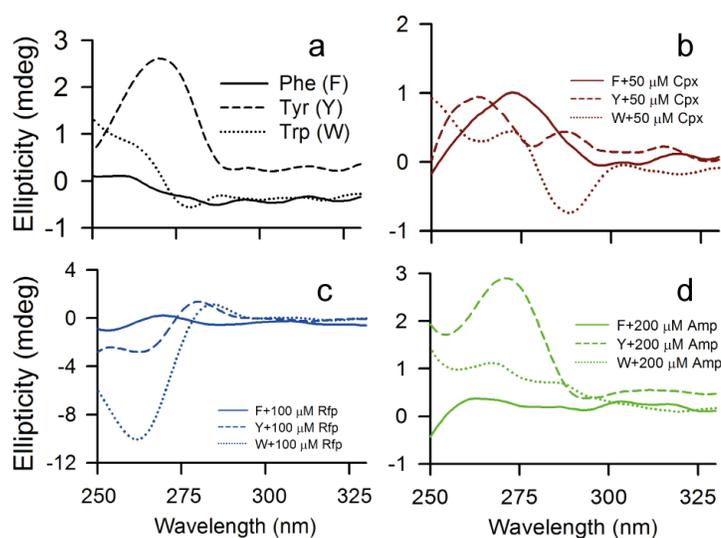


Figure 6.7. Near UV (250-330 nm) CD spectra of isolated amino acids in presence of antibiotics. (a) Shows near UV-CD spectra of isolated Phe (F, solid), Tyr (Y, dash) and Trp (W, dotted). CD spectra

in presence of **(b)** 50 μM Cpx (brown), **(c)** 100 μM Rfp (blue) and **(d)** 200 μM Amp (green) are shown for Phe (F, solid), Tyr (Y, dash) and Trp (W, dotted) .

In chapter 2, we report signature of FRET between fluorophores or aromatic residues of STY3178. This signature of FRET changes in presence of antibiotics (Figures 6.8a,b). Amp binding results in enhancement in intensity of difference spectrum whereas Cpx binding reduces that intensity (Figure 6.8a,b). Rfp bound protein does not show change in FRET intensity but the difference spectrum (275-295 nm) shows appearance of another peak around ~ 318 nm with same intensity which is not present in case of native protein. A decrease in intensity of FRET upon Amp binding and a qualitative change in the entire spectrum in presence of Rfp with a peak ~ 330 nm are observed in the difference spectra of 257-295 nm (Figure 6.8b). The difference spectrum of 257-295 nm for Cpx bound case is significantly quenched. These observations suggest involvement of aromatic residues, responsible for FRET, in antibiotic binding.

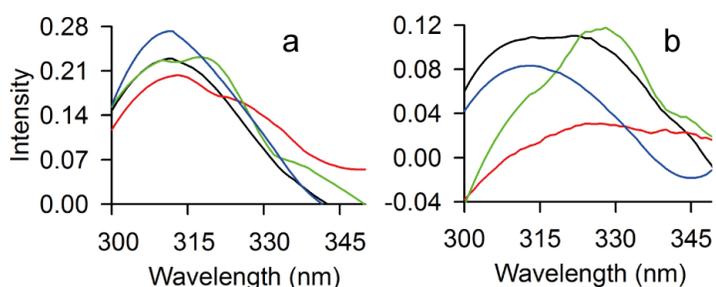


Figure 6.8. Fluorescence difference spectra of STY3178 in native and antibiotic bound states. The difference emission spectra are shown for **(a)** 275-295 nm and **(b)** 257-295 nm. The native protein is shown in black, Cpx (50 μM) bound protein in red, Rfp (200 μM) bound in green and Amp (200 μM) bound state in blue for panels **(a)** and **(b)**.

In literature^{168, 174-176}, we have observed similar binding sites, containing aromatic amino acids for proteins involved in multidrug efflux process. For instance, in BmrR¹⁷⁴ structure bound to kanamycin or tetracycline, antibiotic is stacked near Tyr and Phe residues. In antibiotic bound structure of AcrB¹⁷⁵, Phe is found in the vicinity of Rfp. Amp bound structure of OmpF¹⁷⁶ contains both Phe and Tyr in the vicinity of antibiotic. Similar observation is also

seen for participation of Phe during binding of other small molecule like rhodamine 6G to RamR¹⁶⁸.

In chapter 5, we report that, helical content of STY3178 increases in presence of 0.6 M urea and protein structure is quite stable up to 2 M urea. Therefore, we determine interaction of STY3178 with antibiotic in presence of urea using near UV-CD and steady state fluorescence measurements. In presence of 100 μ M Rfp, the broad shoulder spanning over 250-280 nm in near UV-CD, decomposes into two peaks \sim 260 and \sim 280 nm (Figure 6.3b). This decomposition of shoulder after addition of Rfp is also observed in presence of 0.6 and 1 M urea (Figure 6.9a) indicating antibiotic interaction. Near UV-CD spectrum in 2 M urea differs from that in presence of both 2 M urea and 100 μ M Rfp (Figure 6.9a). However, the broad shoulder or decomposed peaks are not observed. This indicates that 2 M urea affects the tertiary or quaternary structure of STY3178.

We also monitor the emission of tryptophan upon 280 nm excitation in presence of both urea and Rfp. Quenching of fluorescence emission is observed upon addition of 20 μ M Rfp in presence of 0.6 M and 1 M urea as found in case of native protein (Figure 6.9b). This is in agreement with near UV-CD measurements in presence of urea. We find a red shift in emission peak position in presence of 2 M urea (Figure 6.9b) due to exposure of tryptophan residues to more polar environment. However, quenching of tryptophan emission is observed after adding Rfp to protein in 2 M urea (Figure 6.9b). Thus, STY3178 interacts with the antibiotic up to 2 M urea concentration.

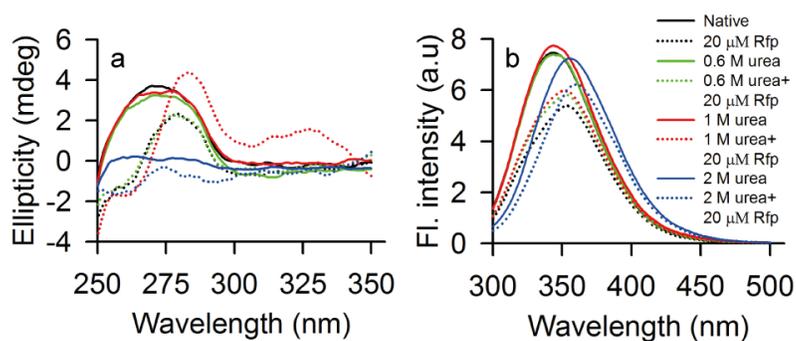


Figure 6.9. Interaction of STY3178 with rifampin (Rfp) in presence of urea. (a) Near UV (250-355 nm) CD and (b) Fluorescence emission spectra of protein for 280 nm excitation are shown in presence of 0.6 M (green), 1 M (red) and 6 M (blue) urea. Spectra in presence of antibiotics are reported for 100 μ M Rfp in (a) and 20 μ M Rfp in (b). Presence of both urea and rifampin are shown using dotted lines. Native protein spectrum is shown in black for both (a) and (b).

6.4.2. Sequence conservation of STY3178 with other yfdX proteins

STY3178 sequence is used as an input in blast¹⁷⁷ to search similar yfdX protein sequences present in different organisms. The results show high sequence identity of >92 % among all the *Salmonella* species and 40 % identity with other reported yfdX proteins from various bacteria. The sequence alignment of yfdX proteins obtained from blast¹⁷⁷ results is shown in Appendix C for MDR strains of different bacteria²²⁻³¹. Sequence similarity and conservation in these bacteria are indicative of similar structural fold among yfdX proteins. The only reported crystal structure (PDB ID 3DZA) of a yfdX protein from *K. pneumoniae* is a homotetramer. However, we find trimeric state of oligomerization for STY3178. This suggests that the primary sequence and size of protein may drive the formation of different oligomerization states in various yfdX proteins. There are several locations in the sequence alignment which have conserved Tyr and Phe residues. For instance, 3DZA structure has one Phe and two consecutive Tyr residues in close proximity. These residues are found to be conserved. This suggests that 3DZA could be involved in antibiotic interaction like STY3178. Interestingly, these proximal Phe and consecutive Tyr residues are conserved among the family of yfdX proteins, indicating a possibility of their involvement in similar antibiotic interactions as observed for STY3178.

6.4.3. Possible cellular localization of STY3178

A possible signal peptide for STY3178 is identified in signalP 4.1 server¹⁷⁸ using its amino acid sequence from Gene bank (gene ID gi|16758993:c3049965-3049366) as an input. SignalP server predicts first twelve residues of the N-terminal region of this protein to be a signal peptide. Our construct of STY3178 used for experiments does not contain nine out of twelve residues of the N-terminus. We observe experimentally that STY3178 is a soluble protein. This suggests that it is not a membrane embedded protein probably. Thus, presence of predicted signal peptide in the N-terminal region indicates a possibility of its localization in the periplasm. Its exact cellular localization is not yet established. However, servers like Cello^{179, 180} and LocTree3¹⁸¹ predicts its sub-cellular localization in the periplasm. There are many identified proteins like peptidyl-prolyl isomerases, disulphide bond isomerases, to name only a few, located in bacterial periplasm which acts as chaperones.¹⁸²⁻¹⁸⁴ These chaperones from periplasm are capable of performing their activity without ATP assistance since ATP is absent in periplasm.¹⁸² The predicted sub-cellular localization as well as chaperone activity of STY3178 without ATP suggests its possible localization in the periplasm.

6.5. Conclusion

We have reported the functional importance of a yfdX protein, STY3178 from *Salmonella* Typhi. It can interact with antibiotics ciprofloxacin, rifampin and ampicillin. We also report its ability to prevent thermal aggregation of insulin and ADH by acting as a chaperone. All these properties indicate that yfdX proteins are not non-functional and possibly perform important cellular functions. Our studies on functional characterization of yfdX protein could be relevant for understanding their role in bacterial virulence and survival.

Appendices

A. Antibiotic or small molecule binding parameters for proteins reported in literature.

S.No.	Protein	Antibiotic/ small molecule	K_d (μM)	Ref.
1	Hemocyanin	Ciprofloxacin	18.76	185
2	NorM	Ciprofloxacin	121.3 \pm 15.7	170
3	NorM	Proflavin	33.6 \pm 1.9	170
4	NorM	Norfloxacin	105.6 \pm 9.8	170
5	YdhE	Ciprofloxacin	90.9 \pm 12.4	170
6	YdhE	Norfloxacin	98.4 \pm 16.2	170
7	YdhE	Proflavin	22.1 \pm 0.9	170
8	AcrB	Proflavin	14.5 \pm 1.1	186
9	AcrB	Ciprofloxacin	74.1 \pm 2.6	186
10	HSA	Mitoxantrone	20.41	187
11	HSA	Mitoxantrone	34.48	187
12	HSA	Mitoxantrone	0.476	187
13	HSA	Maslinic acid	7.042 X 10 ⁻³	188
14	HSA	Quercetin	43.48	189
15	HSA	Oxaliplatin	268.82	190
16	HSA	Doxorubicin	90.91	191
17	HSA	Tetracycline	20.24	192
18	HSA	Cytosine β -D arabinofuranoside	416.7	193
19	HSA	Scutellarin	12.02	194
20	HSA	Betulinic acid	0.59	195
21	HSA	Catechin	82.64	196
22	HSA	cis-fac-[RuIICl ₂ (DMSO) ₃ (KTZ)]	3.03	197
23	HSA	cis-[RuIICl ₂ (bipy)(DMSO)(KTZ)]	2.63 X 10 ⁻³	197
24	HSA	[RuII(η 6-p-cymene)Cl ₂ (KTZ)]	0.526	197

S.No.	Protein	Antibiotic/ small molecule	K_d (μM)	Ref.
25	HSA	[RuII(η^6 -p-cymene)(en)(KTZ)][BF ₄] ₂	6.67	197
26	HSA	[RuII(η^6 -p-cymene)(bipy)(KTZ)][BF ₄] ₂	2.27×10^{-3}	197
27	HSA	[RuII(η^6 -p-cymene)(acac)(KTZ)][BF ₄]	1.299×10^{-3}	197
28	HSA	Hippuric acid	146.2	198
29	HSA	Camptothecin (CPT)	462.96	199
30	HSA	Virstatin	1.64	200
31	HSA	Warfarin	3.14	200
32	HSA	Ciprofloxacin	512.82	201
33	BSA	Quercetin	2.062	189
34	BSA	Streptomycin	160.51	202
35	BSA	Streptomycin	574.71	202
36	BSA	Carbenicillin	200.80	203
37	BSA	Tetracycline	16.13	192
38	BSA	Doxorubicin	128.20	191
39	BSA	Levofloxacin	11.27	204
40	BSA	Sparfloxacin	13.89	204
41	BSA	Enrofloxacin	53.76	204
42	BSA	Cefdinir	6.95	204
43	BSA	Vincristine sulphate	31.55	171
44	BSA	2-hydroxy-N'-(3-hydroxybenzylidene) benzohydrazide (HHB)	10.95	205
45	BSA	Warfarin	32.05	205
46	BSA	Ibuprofen	15.95	205
47	BSA	Olanzapine	9.73	206
	BSA	Ferulic acid	1.02	207
48	BSA	Olmesartan	10.96	208
49	BSA	Olmesartan medoxomil	1.155	208
50	BSA	Diclofenac sodium	10	209
51	Cephalosporin-binding protein from <i>Citrobacter freundii</i>	Cephalosporin	1250	210
52	HSA	Ciprofloxacin	512.82	201
53	EmrE (in SUV)	Proflavin	10.76 ± 2.8	169
54	EmrE (in SDS)	Proflavin	4.56 ± 0.8	169
55	EmrE (in DM)	Proflavin	5.26 ± 0.9	169
56	BmrR	Puromycin	17	174
57	BmrR	Tetracycline	51	174
58	BmrR	Kanamycin	28	174
59	BmrR	4-aminoqualdine	210	174

S.No.	Protein	Antibiotic/ small molecule	K_d (μM)	Ref.
60	Intestinal fatty acid-binding proteins (I-FABP)	R-Ibuprofen	51 ± 11	²¹¹
61	Intestinal fatty acid-binding proteins (I-FABP)	S-Ibuprofen	157 ± 33	²¹¹
62	Intestinal fatty acid-binding proteins (I-FABP)	Bezafibrate	26 ± 4	²¹¹
63	Intestinal fatty acid-binding proteins (I-FABP)	Nitrazepam	1200 ± 80	²¹¹
64	OmpF	Nalidixic acid	95.51	²¹²
65	OmpF	Ciprofloxacin	42.66	²¹²
66	OmpF	Grepafloxacin	26.92	²¹²
67	OmpF	Moxifloxacin	22.39	²¹²
68	OmpF	Ciprofloxacin	67.61	²¹³
69	OmpF (in DPhPC)	Ceftazidime	164.74	¹⁷²
70	Trypsin	Oxytetracycline (OTC)	13.69	²¹⁴
71	PBP2	Nitrocefin	192 ± 24	²¹⁵
72	PBP2	Cefepime	1618 ± 145	²¹⁵
73	PBP2	Ceftazidime	671 ± 116	²¹⁵
74	PBP2	Ampicillin	668 ± 124	²¹⁵
75	PBP2	Oxacillin	180 ± 25	²¹⁵
76	PBP2	Imipenem	603 ± 93	²¹⁵
77	PBP 2a	Cefepime	1620 ± 145	²¹⁶
78	PBP 2a	Ceftazidime	670 ± 115	²¹⁶
79	PBP 2a	Nitrocefin	190 ± 25	²¹⁶
80	PBP 2a	Ceftaroline	20 ± 4	²¹⁷
81	Apotransferrin	cis-fac-[RuIICl ₂ (DMSO) ₃ (KTZ)]	25	¹⁹⁷
82	Apotransferrin	cis-[RuIICl ₂ (bipy)(DMSO)(KTZ)]	1.89×10^{-3}	¹⁹⁷
83	Apotransferrin	[RuII(η 6-p-cymene)Cl ₂ (KTZ)]	30.30	¹⁹⁷
84	Apotransferrin	[RuII(η 6-p-cymene)(en)(KTZ)][BF ₄] ₂	66.67	¹⁹⁷
85	Apotransferrin	[RuII(η 6-p-cymene)(bipy)(KTZ)][BF ₄] ₂	1.69×10^{-3}	¹⁹⁷

S.No.	Protein	Antibiotic/ small molecule	K_d (μM)	Ref.
86	Apotransferrin	[RuII(η^6 -p-cymene)(acac)(KTZ)][BF ₄]	2.63×10^{-3}	¹⁹⁷
87	EhpR	Griseoluteic acid	244 ± 45	²¹⁸
88	albus G	Cephalothin	9.5×10^3	²¹⁹
89	albus G	Cephalosporin C	1.6×10^3	²¹⁹
90	R61	Cephaloglycine	0.4×10^3	²¹⁹
91	R61	Ampicillin	7.2×10^3	²¹⁹
92	R61	Carbenicillin	0.11×10^3	²¹⁹
93	R61	Cephalosporin C	$>1 \times 10^3$	²¹⁹
94	R61	Phenoxymethylpenicillin	$>1 \times 10^3$	²¹⁹
95	R61	Benzylpenicillin	13×10^3	²¹⁹
96	R39	Cephalosporin C	0.19×10^3	²¹⁹
97	FmtA	Bocillin	60	²²⁰
98	Plasma protein	Isometamidium	2.04	²²¹
99	phospholipase A ₂ (PLA ₂)	Minocycline (minoTc)	1.8×10^2	²²²
100	YndB	Flavanone	32 ± 3	²²³
101	YndB	Flavone	62 ± 9	²²³
102	YndB	Flavonol	86 ± 16	²²³
103	OprD ^{+Sias}	Piperacillin	7.60 ± 0.13	²²⁴
104	OprD ^{+Sias}	Ceftazidime	9.08 ± 0.93	²²⁴
105	CG2496	Methiothepin	54 ± 19	²²⁵
106	FPPS	Bisphosphonate 1 (pH7.4)	0.0763	²²⁶
107	FPPS	Bisphosphonate 1 (pH8.5)	0.4329	²²⁶
108	FPPS	Bisphosphonate 2 (pH7.4)	0.1972	²²⁶
109	FPPS	Bisphosphonate 2 (pH8.5)	0.6993	²²⁶
110	FPPS	Bisphosphonate 3 (pH7.4)	0.0490	²²⁶
111	FPPS	Bisphosphonate 3 (pH8.5)	0.2174	²²⁶
112	FPPS	Bisphosphonate 4 (pH7.4)	3.1949	²²⁶
113	FPPS	Bisphosphonate 4 (pH8.5)	3.1250	²²⁶
114	FPPS	Bisphosphonate 5 (pH7.4)	0.1661	²²⁶
115	FPPS	Bisphosphonate 5 (pH8.5)	0.2907	²²⁶
116	FPPS	Bisphosphonate 6 (pH7.4)	0.3205	²²⁶
117	TtgR	Phloretin	0.0476	²²⁷

B. Results of MODexplorer server for STY3178 are tabulated. Proteins present in different organisms with various function and structure, having similarity with STY3178 are shown.

S.No.	Protein (PDB)	Organism	Structure	Function	Hit coverage (%)
1	yfdX (3DZA)	<i>Klebsiella pneumoniae</i>	α -helical (tetrameric)	Unknown	≥ 80
2	N-terminal domain of Cdc27 (3KAE)	<i>Encephalitozoon cuniculi</i>	α -helical (dimeric)	Targets specific cell cycle regulatory proteins for ubiquitin-dependent degradation	≥ 80
3	Protein phosphatase 5 (PP5)- conserved serine/threonine phosphatase (4JA7)	<i>Rattus norvegicus</i>	α -helical+ β sheet	Dephosphorylation	≥ 80
4	putative kinesin (4Y6W)	<i>Podospira anserina</i>	α -helical (tetrameric)	Unknown	≥ 80
5	Ybgf (2XEV)	<i>Xanthomonas campestris</i>	α -helical (trimeric)	Metal binding	≥ 80
6	coatomer cage proteins (3MKR)	<i>Bos taurus</i>	α -helical (trimeric)	Transport protein	≥ 80
7	Adaptor protein Hop (1ELR)	<i>Homo sapiens</i>	α -helical	Chaperone	70-80
8	TRIP8b (tetraatricopeptide repeat-containing Rab8b-interacting protein) (4EQF)	<i>Mus musculus</i>	α -helical	Transport protein	70-80
9	C-terminal part of the TPR repeat-containing protein Q11TI6_CYTH3 (3U4T)	<i>Cytophaga hutchinsonii</i>	α -helical	Unknown	70-80

S.No.	Protein (PDB)	Organism	Structure	Function	Hit coverage (%)
10	TPR Domain of Arabidopsis FLU (FLU-TPR) (4YVO)	<i>Arabidopsis thaliana</i>	α -helical	Interacts with glutamyl-tRNA reductase (GluTR), in the formation of δ -aminolevulinic acid (ALA)	70-80
11	Magnetosome-associated protein, MamA (3AS5)	<i>Magnetospirillum magneticum</i>	α -helical	Protein binding	70-80
12	MalT family of ATP-dependent transcriptional activator (1HZ4)	<i>Escherichia coli</i>	α -helical	Regulates the genes for malto-oligosaccharide utilization	70-80
13	P58(IPK) TPR Domain (3IEG)	<i>Mus musculus</i>	α -helical	Chaperone	70-80
14	tetratricopeptide repeat protein (PARMER_03812) (4R7S)	<i>Parabacteroides merdae</i>	α -helical	Unknown	70-80
15	Sti1/Hop-modular protein (3UQ3)	<i>Saccharomyces cerevisiae</i>	α -helical	Chaperone	70-80
16	CHIP- dimeric U box E3 ubiquitin ligase that binds Hsp90 and/or Hsp70 via its TPR-domain (2C2L)	<i>Mus musculus / Homo sapiens</i>	α -helical	Chaperone - facilitates ubiquitylation of chaperone bound client proteins	70-80
17	N-terminal Domain of Anaphase-Promoting Complex Subunit 7 (3FFL)	<i>Homo sapiens</i>	α -helical	Cell cycle	70-80
18	yclophilin or FK506 binding protein (1IHG)	<i>Bos taurus</i>	α -helical	Isomerase	70-80
19	Type IV pili (T4P) (2HO1)	<i>Pseudomonas aeruginosa</i>	α -helical	Contribute to the virulence of bacterial pathogens	70-80

S.No.	Protein (PDB)	Organism	Structure	Function	Hit coverage (%)
20	Anaphase-promoting complex (APC/C)-multimeric RING E3 ubiquitin ligase (4UI9)	<i>Homo sapiens</i>	α -helical	Controls chromosome segregation and mitotic exit	70-80
21	Co-Chaperone P58Ipk - endoplasmic reticulum- (ER-) localised DnaJ (ERdj) protein (2Y4T)	<i>Homo sapiens</i>	α -helical	Chaperone	70-80
22	Cardiac specific histone methyltransferase SmyD1 (3N71)	<i>Mus musculus</i>	α -helical	Methylates histone H3 at lysine 4 and regulates gene transcription in early heart development	70-80
23	Domain of protein VP0806 (2R5S)	<i>Vibrio parahaemolyticus</i>	α -helical	Unknown	70-80
24	FKBP-like Protein, FKBP51 (1KT0)	<i>Homo sapiens</i>	α -helical	Isomerase-associates with Hsp90 and appears in functionally mature steroid receptor complexes	70-80
25	Active N-terminal fragment of p67 (1HH8)	<i>Homo sapiens</i>	α -helical	Cell cycle	70-80
26	Vesicular transport protein SEC17 (1QQE)	<i>Saccharomyces cerevisiae</i>	α -helical	Membrane trafficking in eukaryotic cells	60-70
27	Motor protein (3EDT)	<i>Homo sapiens</i>	α -helical	-	60-70
28	Ton1535, a hypothetical protein (3ZPJ)	<i>Thermococcus onnurineus</i>	α -helical	Unknown	60-70

S.No.	Protein (PDB)	Organism	Structure	Function	Hit coverage (%)
29	APC/C (2XPI)	<i>Schizosaccharomyces pombe</i>	α -helical	Cell cycle	60-70
30	TPR DOMAIN OF HUMAN SGT (2VYI)	<i>Homo sapiens</i>	α -helical	Chaperone	60-70
31	Uncharacterized protein (3GW4)	<i>Deinococcus radiodurans</i>	α -helical	Unknown	60-70
32	Tah1 (4CGU)	<i>Saccharomyces cerevisiae / Homo sapiens</i>	α -helical	Chaperone	60-70
33	TTC0263 (2PL2)	<i>Thermus thermophilus</i>	α -helical	Protein binding-Forms multi-subunit complexes by interacting with each other or with other subunit proteins	60-70
34	Peptide-substrate-binding domain of human type I collagen prolyl 4-hydroxylase (1TJC)	<i>Homo sapiens</i>	α -helical	Hydrolase	60-70
35	Stress-induced-phosphoprotein 1 STI1 (2LNI)	<i>Homo sapiens</i>	α -helical	Chaperone	60-70
36	Tetratricopeptide Repeat (TPR) Domain of Fluorescent (FLU)-negative regulator of chlorophyll biosynthesis in plants (4YVQ)	<i>Arabidopsis thaliana</i>	α -helical	Oxidoreductase	60-70
37	Multi-domain immunophilin (2IF4)	<i>Arabidopsis thaliana</i>	α -helical	Signaling protein	60-70
38	LapB cytoplasmic domain (4ZLH)	<i>Escherichia coli</i>	α -helical	Metal binding	60-70

S.No.	Protein (PDB)	Organism	Structure	Function	Hit coverage (%)
39	AIP TPR domain in complex with human Tomm20 peptide (4APO)	<i>Homo sapiens</i>	α -helical	Signaling protein	60-70
40	YbbN (3QOU)	<i>Escherichia coli</i>	α -helical	Proposed to act as a chaperone or co-chaperone	60-70
41	Stress-induced protein-1 (STI-1) (4GCN)	<i>Caenorhabditis elegans</i>	α -helical	Protein binding	60-70
42	O-linked GLCNAC transferase (1W3B)	<i>Homo sapiens</i>	α -helical	Transferase	60-70
43	SYCD (2VGX)	<i>Yersinia enterocolitica</i>	α -helical	Chaperone	60-70
44	EccA1 (4F3V)	<i>Mycobacterium tuberculosis</i>	α -helical	Responsible for transport of virulence factors in pathogenic mycobacteria	60-70
45	TPR1 domain of HOP in complex with a HSC70 peptide (1ELW)	<i>Homo sapiens</i>	α -helical	Chaperone	60-70
46	Peroxisomal targeting signal-1 (PTS1) (1FCH)	<i>Homo sapiens</i>	α -helical	Signaling protein	60-70
47	MamA (4XI0)	<i>Desulfovibrio magneticus</i>	α -helical	Protein binding	60-70
48	Serine chemoreceptor Tsr (2D4U)	<i>Escherichia coli</i>	α -helical	Signaling protein	60-70
49	Hsp70-interacting protein, Hip (4J8F)	<i>Homo sapiens / Rattus norvegicus</i>	α -helical	Chaperone	50-60

S.No.	Protein (PDB)	Organism	Structure	Function	Hit coverage (%)
50	Human O-GlcNAc Transferase (4GYW)	<i>Homo sapiens</i>	α -helical	Transferase	50-60
51	APC3-APC16 complex (4RG9)	<i>Homo sapiens</i>	α -helical	E3 ligase that controls mitosis by catalyzing ubiquitination of key cell cycle regulatory proteins	50-60
52	O-GlcNAc (5DJS)	<i>Thermobaculum terrenum</i>	α -helical	Transferase	50-60

C. Sequence alignment of yfdX family proteins. The sequence alignment of STY3178 protein from CT18 strain of *S. Typhi* with yfdX proteins from other virulent bacteria like *K. pneumoniae*, *S. enteritidis*, *S. Typhimurium*, *S. Paratyphi*, *S. Heidelberg*, *S. Montevideo*, *P. ananatis*, *E. tarda*, *H. alvei*, *E. coli*, and *P. Shigelloides* is shown. The secondary structural elements α -helix (cylinder) and β -strands (arrow) are indicated on the top with respect to the crystal structure of yfdX protein (PDB ID 3DZA) from *K. pneumoniae*. The conserved Tyr and Phe residues are marked by the rectangular blocks.

	H1	H1	H2	
STY3178	-----	-----	-----	DKISGQAWQAM 34
<i>K.pneumoniae</i> (3DZA)	-MVKVILASLLATMMSTSPVWATDSATAAPAAAATTQVQ---	-----	-----	KEADV LQVAVQGANAM 55
<i>S.Enteritidis</i> PT4	-----	-----	-----	DKISGQAWQAM 34
<i>S.Typhimurium</i> LT2	-----	-----	-----	DKISGQAWQAM 34
<i>S.Typhimurium</i> 798	-----	-----	-----	DKISGQAWQAM 34
<i>S.Typhimurium</i> SL1344	-----	-----	-----	DKISGQAWQAM 34
<i>S.Paratyphi</i> A	-----	-----	-----	DKISGQAWQAM 34
<i>S.Heidelberg</i> 41579	-----	-----	-----	DKISGQAWQAM 34
<i>S.Heidelberg</i> SL476	-----	-----	-----	DKISGQAWQAM 34
<i>S.Montevideo</i>	-----	-----	-----	DKISGQAWQAM 34
<i>P.ananatis</i>	MKKMIVSSMIVAVLVFTGGAAAESFKDSHTLNN-----	-----	-----	EVVSAQAAQAM 45
<i>E.tarda</i>	-MKRIILTATLSLAMFSTFTWAADAAPAPASATSTAAG-	LSAQQVQDVAKVAQQGFTAM 58		
<i>H.alvei</i>	-----	MMSTFTWAADSAATVPAATTTVSTNVLTAQQAKEVAKVAQQGFNAM 46		
<i>E.coli</i>	-MKRLIMATMVTAI LASSTVWAADNAPVAAQQQTQQVQQ--	TQKTAAAERI SEQGLYAM 57		
<i>P.shigelloides</i>	-MKHLLSAVALSVALISGGVAAEQFATATVAATQQQK-----	DLQVAQQGYNAM 51		
	: : * :		: : * **	

Chapter 7 | Homology modeling and molecular dynamics simulation of STY3178

7.1. Introduction

In previous chapters, we have biophysically characterized the yfdX protein, STY3178 from *Salmonella* Typhi. We find it to be capable of undergoing reversible thermal unfolding reported in chapter 4. In order to understand these results, we generate a homology model for STY3178. There is only one crystal structure of yfdX protein from *Klebsiella pneumoniae*, deposited in PDB with an ID 3DZA. The deposited structure has four identical chains A, B, C and D forming a homotetramer. We observe metal ion coordination between these chains.

Homology modeling is a tool to generate the three dimensional model of proteins, which lack experimentally determined structure.⁵³ We build the three dimensional model of STY3178 using different modeling servers with respect to the only available crystal structure (PDB ID 3DZA) of yfdX protein from *Klebsiella pneumoniae*. Energy minimization is performed using NAMD followed by molecular dynamics (MD) simulation at 310 K. The equilibrated model structure at 310 K is used to perform MD simulations at elevated temperatures. We report in chapter 3 that STY3178 is a trimer in solution. Thus, we generate the model for the trimeric assembly of the protein by docking three chains of STY3178 monomer structure.

This chapter is organized as follows: Methods and results related to model building and molecular dynamics simulation are discussed in sections 7.2 and 7.3, respectively. Finally, we conclude in section 7.4 stating the overall finding of this work.

7.2. Methodology

7.2.1. Molecular model of STY3178

We perform homology modeling of STY3178 using various servers namely Swiss model workspace²²⁸, Phyre2⁵⁴ and RaptorX⁵⁵⁻⁵⁸. All these servers generate a tertiary structure for STY3178 with respect to the same template yfdX protein (PDB ID 3DZA) from *Klebsiella pneumoniae*. Swiss model workspace and RaptorX use chain A of the template protein whereas Phyre2 build the model using the chain B. Model structures obtained using

RaptorX and Phyre2 have 173 amino acid (A¹¹-Q¹⁸³) residues. Swiss model workspace generated structure contains only 165 residues (L²¹-Q¹⁸⁶). This shows 10 residues (A¹¹-T²⁰) are missing in the N-terminus of the model obtained from Swiss model workspace. We therefore, work further with the model structure generated using RaptorX and Phyre2 servers having 173 residues (A¹¹-Q¹⁸³).

Amino acid sequence of the STY3178 protein construct used for the experiments contains total number of 210 residues. We compare the construct amino acid sequence with that of the model. There are 21 residues at the N-terminus of the construct including the 6 histidine (His) tags and few others inserted for the purpose of cloning. The model structure (A¹¹-Q¹⁸³) lacks these 21 residues in the N-terminus. Similarly, it does not contain 16 residues at the C-terminus compared to the construct protein sequence. Therefore, we include only the C-terminal residues (S¹⁸⁴-H¹⁹⁹) of the construct in our model leaving N-terminal ones. We add these 16 C-terminal residues De-novo one by one followed by energy minimization of 100 steepest decent steps after addition of each amino acid in Swiss-PdbViewer⁵⁹. The final model has 189 amino acid (A¹¹-H¹⁹⁹) residues. Hydrogen atoms are added to this model structure having 189 residues followed by solvation using explicit solvent water model in a rectangular parallelepiped box with dimension 69.4 X 62.2 X 73 Å³. The solvated system is neutralized with counter ions and minimized using NAMD⁶⁰.

7.2.2. Molecular dynamics simulation of the model

All atom molecular dynamics (MD) simulation of the minimized monomer structure is performed using NAMD⁶⁰. The force field and water model used are CHARMM27⁶¹ and TIP3P⁶², respectively. MD is performed at 1 atm pressure in isothermal-isobaric (NPT) ensemble with periodic boundary condition and 1 femtosecond (fs) time-step. Electrostatic interactions are treated using particle-mesh Ewald method. Energy minimization of 10,000 steps is carried out with total number of 29848 atoms prior to 300 nanosecond (ns) MD simulation at 310 K temperature. MD simulations at elevated temperatures 350 K and 400 K are performed using an equilibrated structure from the 310 K ensemble. The root mean square deviation (RMSD) is calculated with respect to the C_α-atoms over all the trajectories up to 300 ns for each of the temperatures. The root means square fluctuation (RMSF) of C_α-atoms of each residue is also estimated over the equilibrated trajectories from 150 to 300 ns at temperatures 310 K, 350 K and 400 K. The histogram distribution, Ramachandran plot and standard deviation (σ) of backbone dihedral angles ϕ and ψ is calculated over 150-300 ns trajectories for

each residue. The relative change in σ at 350 K ($r_{350} = \sigma_{350}/\sigma_{310}$) and 400 K ($r_{400} = \sigma_{400}/\sigma_{310}$) is then estimated with respect to that at 310 K for each residue. The propensity (percentage) of residues forming α -helix and β -strands are calculated over 150-300 nm trajectories from the distribution of backbone dihedral angles ϕ and ψ .

Similarly, MD simulation is performed for the protein lysozyme (PDB ID 193L) for 100 ns using the total number of 21002 atoms at 310 K and 400 K following the aforementioned protocol.

7.2.3. Molecular model of trimeric assembly

Different structures for the trimeric assemblies of STY3178 are generated using 100 ns equilibrated monomer structure at 310 K ensemble in ZDOCK⁶³ docking server. The energetically most favourable trimeric assembly among the best possible five structures is determined by energy minimization using NAMD⁶⁰ with CHARMM27 force field⁶¹ and TIP3P⁶² model of water. We then perform MD simulation with the energetically favourable trimer assembly at 310 K following the similar protocol used for the monomer. The total number of atoms during MD simulation of the trimer is 61860.

The interface residues of the STY3178 trimeric structure are identified by measuring the distance within 5 Å cut off between the C_α atoms of each residue belonging to different chains of the trimer.

The radius of gyration (R_g) of the simulated trimeric assembly of STY3178 is calculated as the square of the distance of C_α -atoms from the centre of mass (\vec{R}_{CM}) defined by

$\vec{R}_{CM} = \sum_i m_i \vec{r}_i / \sum_i m_i$, where m_i and \vec{r}_i are the respective mass and position vectors of the i^{th}

C_α -atom. R_g is calculated using the expression⁶⁴, $R_g^2 = \sum_i m_i \left(\vec{r}_i - \vec{R}_{CM} \right)^2 / \sum_i m_i$. The

hydrodynamic radius (R_H) of the trimer model is then estimated from the R_g value, following

the relation⁶⁵, $\frac{R_g}{R_H} = 0.77$.

helical secondary structural elements of STY3178 model are H1 (A12-D24), H2 (G28-N46), H3 (T49-N64), H4 (S67-Y73), H5 (A103-E119), H6 (K122-A132), H7 (L145-E161), H8 (K163-G177), H9 (S182-T190) and H10 (Q194-A198). The β -sheet comprises of two β -strands E1 (Y87-P102) and E2 (V136-P144). The structure is composed of two helical bundles. One bundle comprising of helices H2, H3, H7 and H8 and the other with helices H5, H6 and H10. These two helical bundles are separated by the single β -sheet. The α -helix and β -sheet contents of the model are 68.25 % and 13.23 %, respectively. We report that STY3178 has 44 % α -helix content from CD measurement (in chapter 2). Dichroweb server provided 50 % α -helix and 20 % β -sheet content for the CD data. This discrepancy in the percentage of secondary structural elements might be due to the difference in the total number of amino acid residues of the construct and that of the model structure.

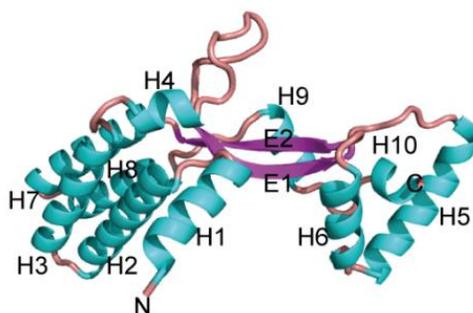


Figure 7.2. Cartoon representation of homology model of STY3178.

Ten α -helices are H1, H2, H3, H4, H5, H6, H7, H8, H9 and H10. Two β -strands are marked as E1 and E2 along with the N- and C-terminals. The α -helices and β -strands are shown in cyan and magenta colours, respectively. Loop residues are shown in pink colour.

7.3.2. Molecular dynamic simulations at elevated temperatures

We report in chapter 4 that STY3178 is thermally stable without change in its trimeric state of oligomer upon heating. It also refolds back from thermally unfolded state when cooled. This leads us to determine the thermal stability of monomeric protein structure. We perform MD simulations with the monomer structure at elevated temperatures 350 K and 400 K. A simulated structure from the ensemble of 310 K is used to perform MD at 350 K and 400 K. The root means square deviation (RMSD) plot show saturation for the time scale 150-300 ns at all the three temperatures (Figure 7.3a). The root means square fluctuation (RMSF) is found to increase with temperature (Figure 7.3b). There are two regions showing enhanced RMSF with

increase in temperature. The first region forms the part of H3 and H4 with the adjacent loop residues and the other region composed of E1, H7 and adjacent loops.

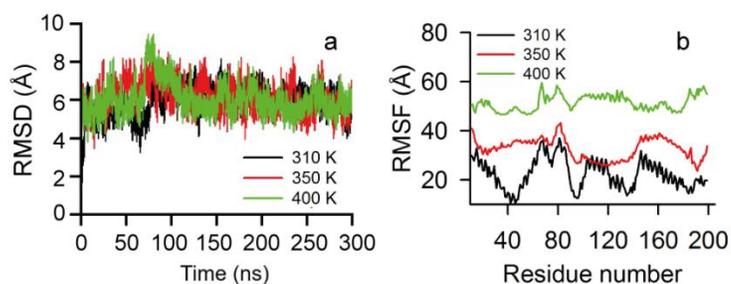


Figure 7.3. Molecular dynamics simulation of STY3178 monomer model. (a) RMSD and (b) RMSF plots of the molecular model, simulated for 300 ns at temperatures 310 K (black), 350 K (red) and 400 K (green).

The backbone dihedral angle (ϕ and ψ) distribution for majority of residues does not change with increase in temperature. For example, the ϕ and ψ distribution of residues N91, R41 and K112 at temperature 310 K, 350 K and 400 K is shown in Figures 7.4a-f. The histogram plots are overlapping with each other for all the three temperatures indicating no major change in the structural parts to which these three residues belong.

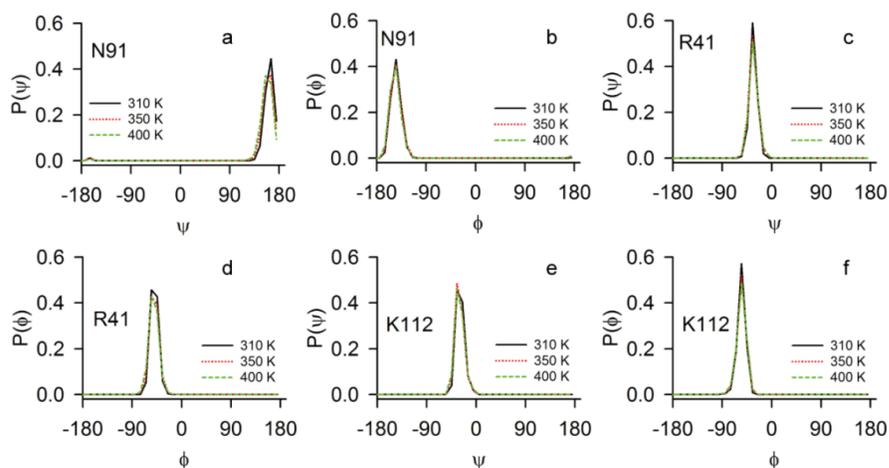


Figure 7.4. Backbone dihedral angle distribution of residues belonging to STY3178 model. The respective ϕ and ψ distributions are shown for N91 in (a) and (b); R41 in (c) and (d); K112 in (e) and (f). The temperatures are 310 K (black), 350 K (red) and 400 K (green) for (a) to (f).

There are few residues which show sensitivity in the distribution of backbone dihedral angles ϕ and ψ at elevated temperatures. These residues are N14, D17, N18 of H1; D69, W70 and N71 belonging to H4; A132 from H6; S182, Q183, S184 and V185 of H9; V193, H195, A197 and A198 of H10. An example of such change is shown for residue W70 (Figures 7.5a,b) and S184 (Figures 7.5c,d). The distribution for W70 and S184 shows deviation at 350 K and 400 K compared to that at 310 K. This indicates structural changes for these residues at elevated temperatures. These structural changes are clearly reflected in the Ramachandran plots of these residues which show transition from helix to loop and β -sheet (Figures 7.6a-f).

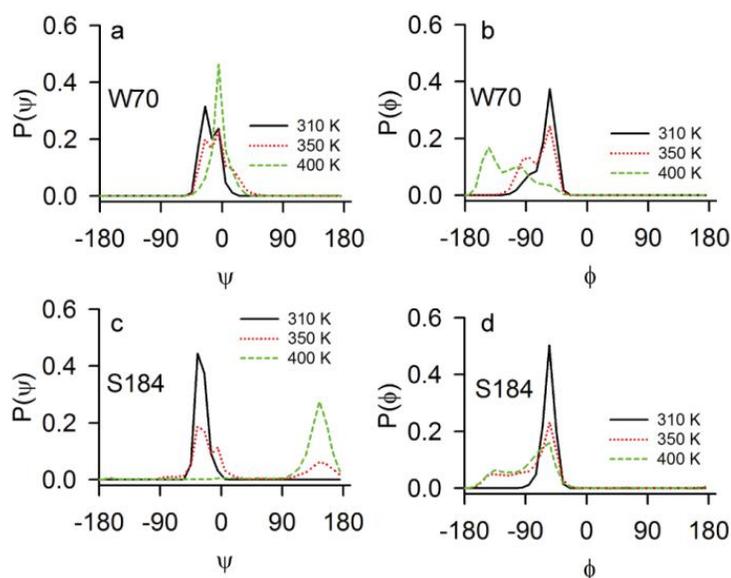


Figure 7.5. Backbone dihedral angle distribution of residues showing sensitivity. The respective ϕ and ψ distributions are represented for W70 in (a) and (b); S184 in (c) and (d) at temperatures 310 K (black), 350 K (red) and 400 K (green).

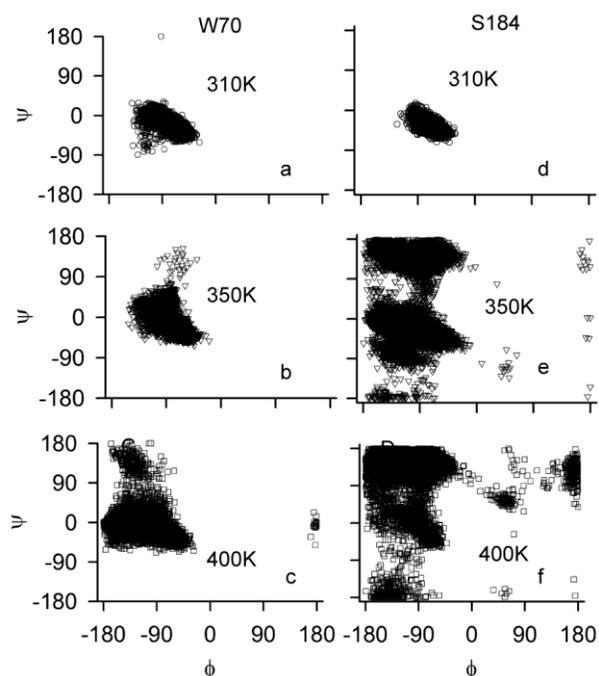


Figure 7.6. Ramachandran plot of affected residues. (a), (b) and (c) Show the Ramachandran plot of W70 at 310 K (circle), 350 K (inverted triangle) and 400K (square), respectively. The plots for S184 at 310 K (circle), 350 K (inverted triangle) and 400 K (square) are shown in (d), (e) and (f), respectively.

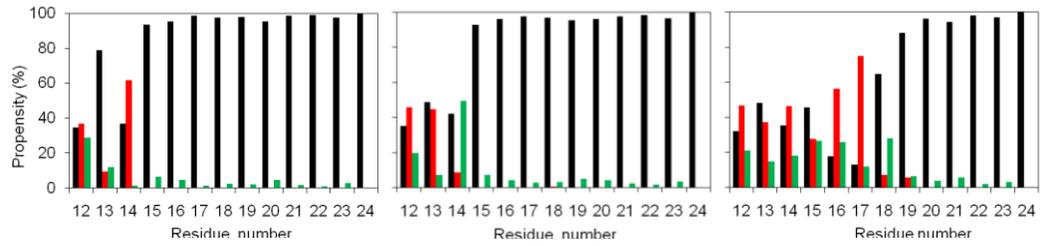
We further calculate the propensity of different residues forming various secondary structural elements (helices and strands) of the model to get a better understanding of the effect of elevated temperatures on STY3178 monomer structure. Figure 7.7 shows the propensity of various residues being a part of α -helix or β -strand for three temperatures 310 K, 350 K and 400 K. The N-terminal residues of H1 shows change in α -helical propensity with increase in temperature. We find that the propensity of residues forming helices H2, H3, H5, H6, H7 and H8 does not change even at 400 K. Residues of H4 are affected at 400 K. The C-terminal residues forming H9 and H10 shows change in their α -helical percentage at 350 K and 400 K compared to that at 310 K (Figure 7.7). The helical nature of H9 and H10 reduces substantially at elevated temperatures 350 K and 400 K. We find the propensity of residues belonging to β -strands also remains unchanged during simulation at 350 K and 400 K (Figure 7.7). This shows that the model structure is quite stable at elevated temperature except for the N- and C-terminal secondary structural elements.

310 K

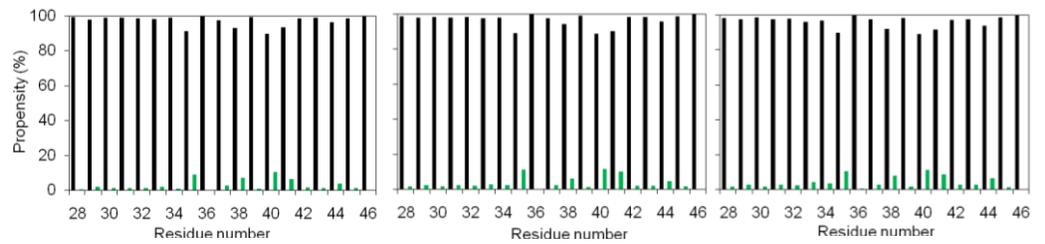
350 K

400 K

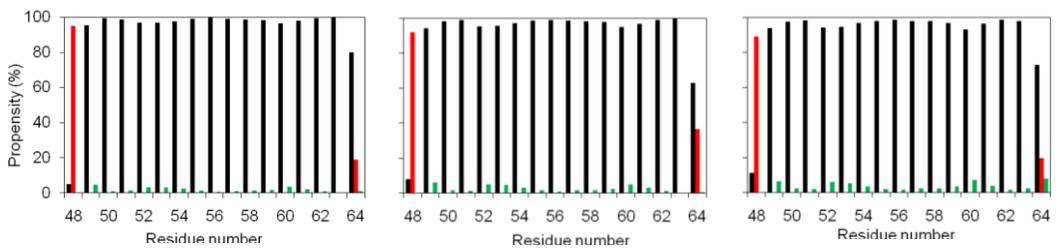
Helix 1 (H1)



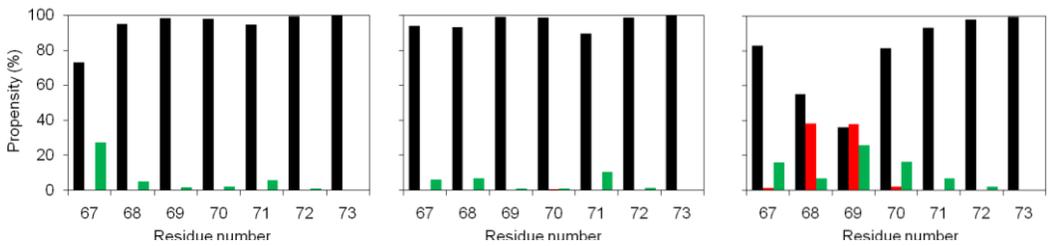
Helix 2 (H2)



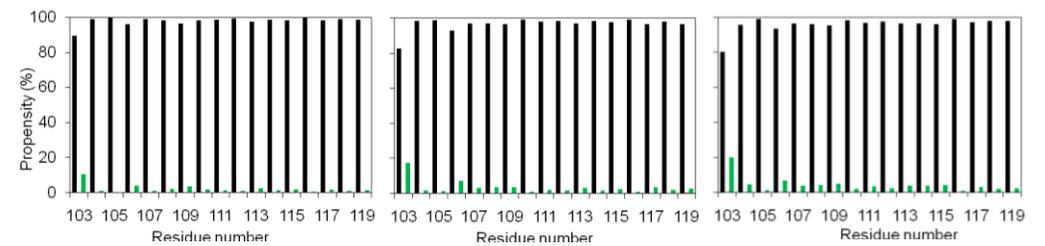
Helix 3 (H3)



Helix 4 (H4)



Helix 5 (H5)

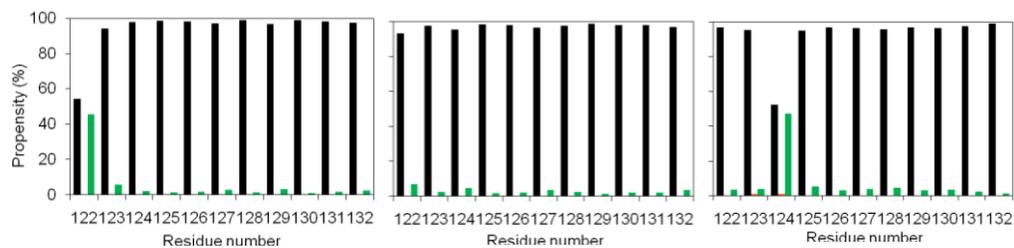


310 K

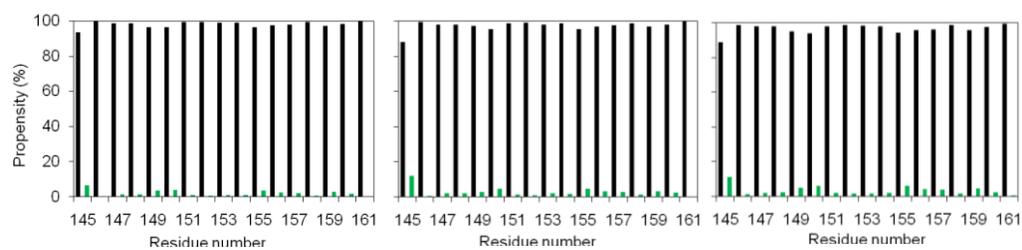
350 K

400 K

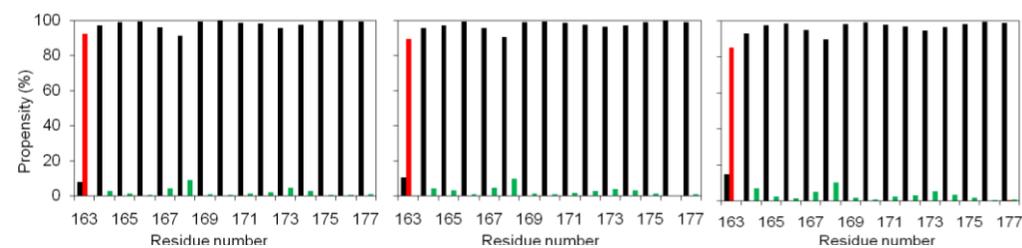
Helix 6 (H6)



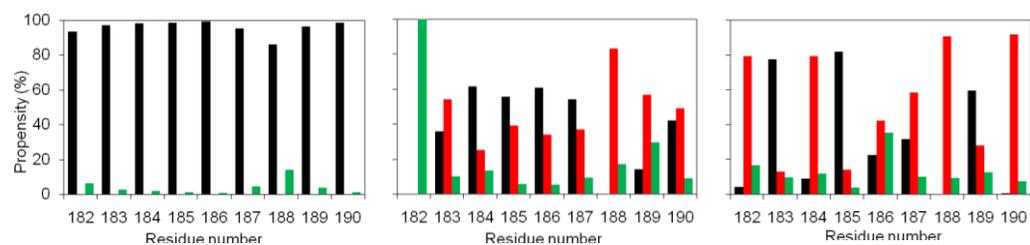
Helix 7 (H7)



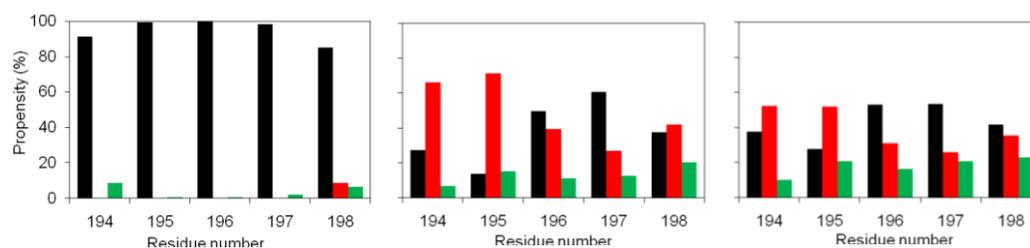
Helix 8 (H8)



Helix 9 (H9)



Helix 10 (H10)



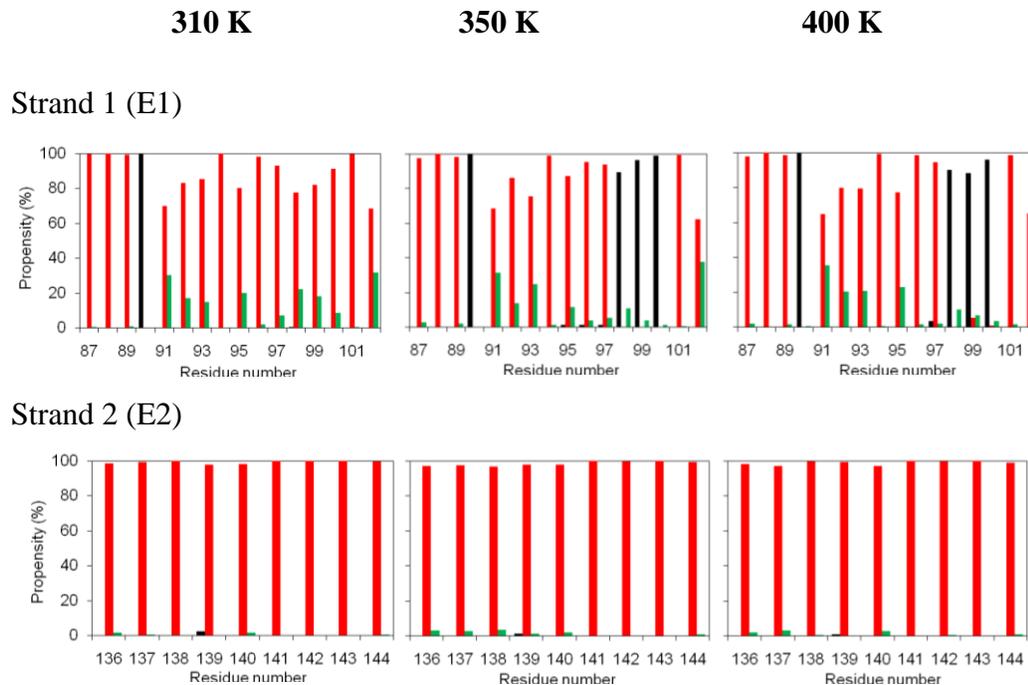


Figure 7.7. The propensity of different residues belonging to various secondary structural elements (α -helices and β -strands) is shown for 310 K, 350 K and 400 K over 150 ns to 300 ns simulation trajectories. The bar chart includes ten helices (H1-H10) and two strands (E1 and E2) of the STY3178 model. The propensity of being in helix, strand and loop are represented in black, red and green, respectively.

We find the RMSF increases at 350 K and 400 K compared to that at 310 K (Figure 7.3b) which indicates enhanced thermal fluctuations at elevated temperatures. Therefore, we calculate the change in backbone dihedral angle fluctuations of all the residues of the model at 350 K and 400 K with respect to that at 310 K. This fluctuation in ϕ and ψ is represented by calculating the relative standard deviations (r_{350} and r_{400}) in backbone dihedral angles of various residues at 350 K and 400 K compared to those at 310 K. We define the relative standard deviations at 350 K (r_{350}) and 400 K (r_{400}) as $r_{350} = \sigma_{350} / \sigma_{310}$ and $r_{400} = \sigma_{400} / \sigma_{310}$, respectively, where σ denotes the standard deviation at a particular temperature. The 300 ns simulated structure at 310 K are coloured according to the r_{350} and r_{400} values at 350 K and 400 K, respectively. Residues with $r_{350}, r_{400} \leq 1$ for both ϕ and ψ are coloured green (Figures 7.8a,b). Maximum residues belong to the region $1 < r_{350}, r_{400} < 4$ having moderate fluctuations in either ϕ or ψ or both at 350 K and 400 K are shown in orange (Figures 7.8a,b). Residues with maximum fluctuation has $r_{350}, r_{400} > 4$ and marked in red. These residues belong to the terminal helices H1, H9 and H10. This is in agreement with the changes observed in the propensity of

the terminal residues (Figure 7.7). The overall structure of STY3178 remains stable besides enhancement in fluctuations in some region of the model. Figures 7.8c,d show the average structure at temperatures 350 K and 400 K over the equilibrated trajectories 150-300 ns. Both the average structures contain a third β -strand (E3) comprising of residues V179-D181, which is not present in structure at 310 K (Figure 7.2). Helices H9 and H10 are affected at both the temperatures whereas H1 and H4 unwind at 400 K only. These structural changes commensurate with that of propensity changes of these residues shown in Figure 7.7. Rest of the structural elements does not show substantial change.

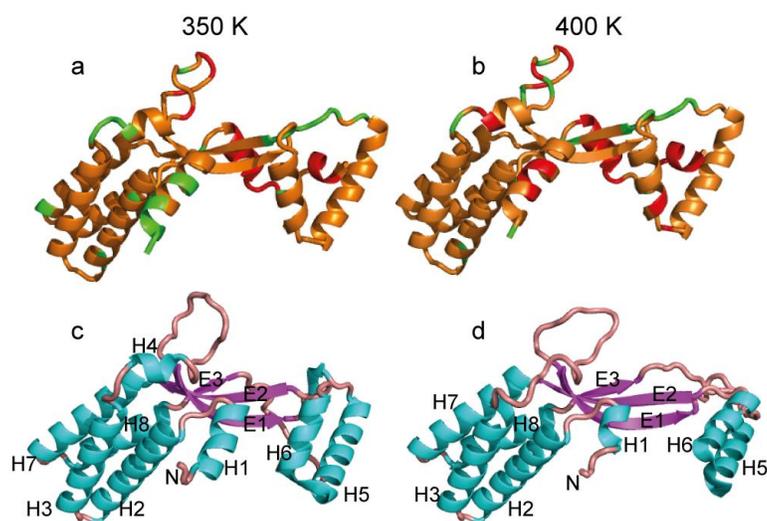


Figure 7.8. Cartoon representation of the molecular model (at 310 K) with residues coloured according to their degree of fluctuations at (a) 350 K (r_{350}) and (b) 400 K (r_{400}). Green indicates residues with $r_{350}, r_{400} \leq 1$, orange shows residues with r_{350} and r_{400} in the range 1-4 and the residues with $r_{350}, r_{400} > 4$ are shown in red. The average structures over the trajectories 150-300 ns are shown for 350 K in (c) and 400 K in (d). The α -helices (cyan) and β -strands (magenta) are marked by H and E, respectively. Loops are shown in pink.

We compare our simulation results at elevated temperature with that of a standard protein lysozyme. We perform MD simulation using the crystal structure of lysozyme (PDB ID 193L) at 310 K and 400 K. We observe enhanced RMSF for lysozyme structure simulated at 400 K compared to 310 K within 100 ns using the same force field (Figure 7.9a). There is melting of secondary structural element at 400 K (Figure 7.9b). This loss of secondary structure

of lysozyme at elevated temperature is in agreement with earlier report²²⁹. This further supports that our model of STY3178 structure is thermally stable.

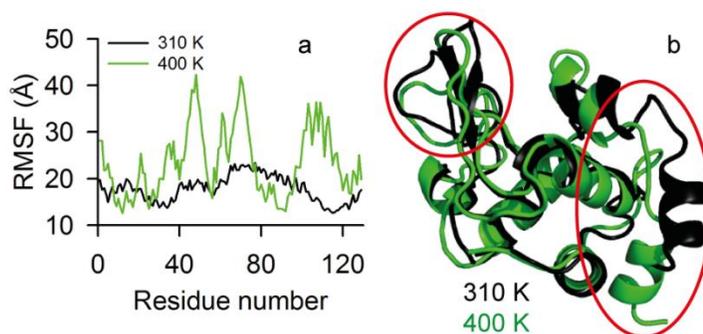


Figure 7.9. MD simulation of lysozyme structure (PDB ID 193L). (a) Represents the root mean square fluctuation (RMSF) of lysozyme residues over 50 ns to 100 ns at temperatures 310 K (black) and 400 K (green). (b) Superimposed cartoon representation of 100 ns simulated lysozyme structure at 310 K (black) and 400 (green). The structural elements showing changes at 400 K with respect to 310 K are circled in red.

7.3.3. Trimer model

We report in chapter 3 that STY3178 forms trimer in solution. Thus, we generate the trimer structures of the protein using ZDOCK⁶³ docking server. Three identical STY3178 structures of the 100 ns equilibrated monomer at 310 K, are docked to obtain the trimeric arrangement. The best five possible trimer structures obtained from the server are minimized using NAMD⁶⁰ with CHARMM27 force field⁶¹ and TIP3P⁶² water model to determine the energetically most favourable assembly. Figure 7.10a shows the energetically stable trimeric assembly after minimization among the best five structures. It consists of three monomer chains (A, B and C) arranged symmetrically. There is a central pore composed of H2 and H8 (Figure 7.10a). The interface residues of the trimer are tabulated in Table 7.1. The trimer interface has hydrophobic, acidic, basic as well as polar residues of both the adjacent chains. These residues are shown in Figure 7.10b. The interaction at the interface would be both hydrophobic and electrostatic. There are three salt bridge pairs existing in the interface of trimer. These are K118-D85, K163-D24 and K122-E161. Along with this, there are intra-chain and inter-chain FRET pairs. The intra-chain pair is W70 and Y73, Y87 whereas W71 and Y165, Y164 forms the inter-chain FRET pairs (Figure 7.10c) of the trimer assembly. The presence of these pairs can explain the observed phenomenon of FRET in steady state fluorescence of STY3178

reported in chapter 2. The electrostatic surface diagram of STY3178 trimer is represented in Figure 7.10d. It shows presence of large number of exposed basic as well as acidic residues on the surface of assembly (Figure 7.10d). We report in chapter 5 that the secondary, tertiary structure and hydrodynamic size of STY3178 is stable over a pH range of 4.5 to 10. The pH stability could be due to the uniform distribution of both acidic and basic residues on the surface of the protein. Our proposed trimer model has evenly distributed acidic and basic residue which can explain the pH stability of STY3178 reasonably.

We next determine the radius of gyration (R_g) of the trimeric assembly following the relation⁶⁴, $R_g^2 = \sum_i m_i \left(\vec{r}_i - \vec{R}_{CM} \right)^2 / \sum_i m_i$, described in method section. \vec{R}_{CM} is the distance of C_α -atoms from the centre of mass and can be estimated by using the relation, $\vec{R}_{CM} = \sum_i m_i \vec{r}_i / \sum_i m_i$. The respective mass and position vectors of the i^{th} C_α -atom are m_i and \vec{r}_i . The value of R_g obtained for the trimeric assembly is 2.69 nm. We further calculate the hydrodynamic radius (R_H) for this R_g value of the trimeric assembly using the relation⁶⁵, $\frac{R_g}{R_H} = 0.77$. The R_H obtained for the trimer model is 3.49 nm and the corresponding hydrodynamic diameter is 6.98 nm. This is in good agreement with hydrodynamic diameter ~ 6.5 nm measured from DLS in chapter 2. All these suggest that the trimer model can explain the experimental observations quite well and is a reasonable structure.

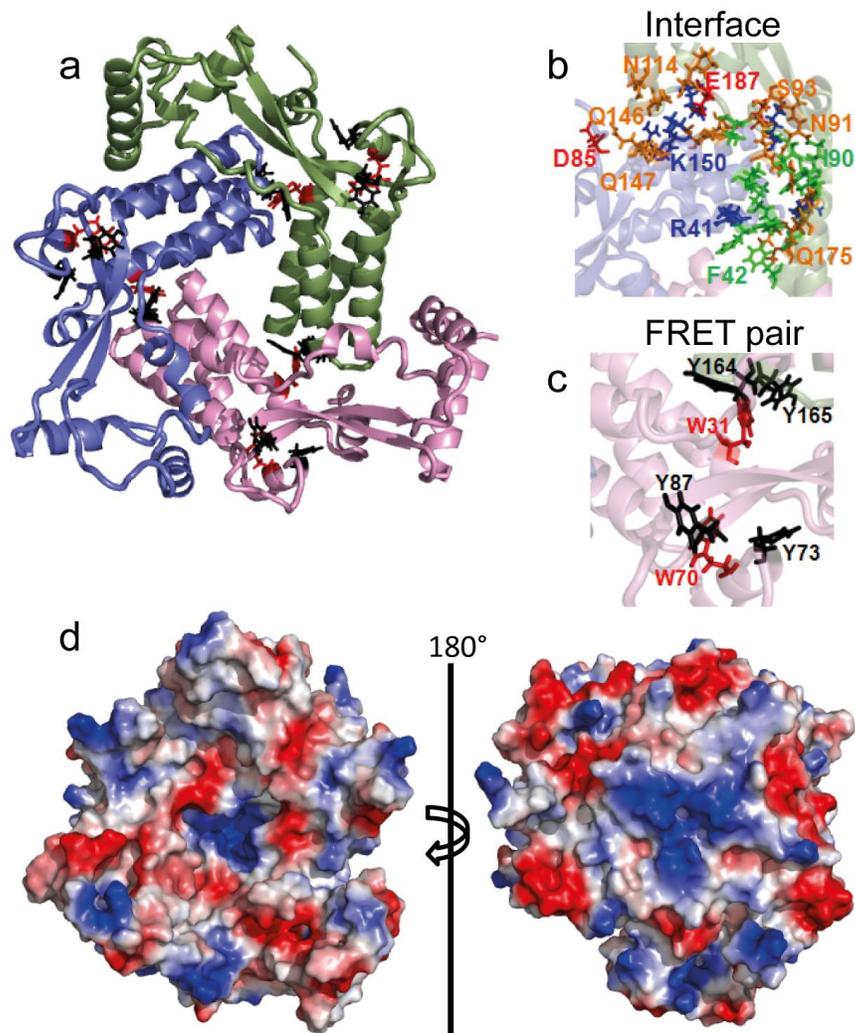


Figure 7.10. Trimeric assembly of STY3178. (a) Three monomer chains (green, blue and pink) are arranged in symmetrical manner. Tyrosine (Y) and tryptophan (W) are coloured black and red, respectively and shown in stick representation. (b) Shows the residues at the interface of monomer chains of trimer. Hydrophobic, acidic, basic and polar residues are shown in green, red, blue and orange, respectively. (c) Inter and intra chain tryptophan (red)-tyrosine (black) FRET pairs present in the trimer model are shown. (d) Represents the front and rear view of electrostatic surface of STY3178 trimer. The basic, acidic and neutral surfaces are coloured blue, red and white, respectively.

Table 7.1. Interface residues of STY3178 trimeric assembly

Residue Type	Hydrophobic	Acidic	Basic	Polar
Chain 1	F42, F45, G162, L169	D85, E161	R41, R149, K150, K163	N46, Q146, Q147, T153, T154, S157, Y164, Y165, Q166
Chain 2	W31, G38, F42, I90, I178, V180, V185, P189	E187	R35, K122	M39, N91, S92, S93, N114, M117, Q175, S184, S188

7.4. Conclusion

We report a MD simulated structure of STY3178 monomer. MD simulations of the model at elevated temperature show its thermal stability. Finally, we propose a simulated structure for the trimeric assembly of STY3178 using the equilibrated monomer structure. The proposed trimer structure can explain our experimental observations reasonably well.

References

1. Branden C & Tooze, J. Introduction to Protein Structure, Edn. 2nd. (Garland, New York; 1999).
2. Valastyan, J.S. & Lindquist, S. Mechanisms of protein-folding diseases at a glance. *Dis Model Mech* **7**, 9-14 (2014).
3. Putman, M., van Veen, H.W. & Konings, W.N. Molecular properties of bacterial multidrug transporters. *Microbiol Mol Biol Rev* **64**, 672-693 (2000).
4. Laursen, B.S., Sorensen, H.P., Mortensen, K.K. & Sperling-Petersen, H.U. Initiation of protein synthesis in bacteria. *Microbiol Mol Biol Rev* **69**, 101-123 (2005).
5. Creighton, T.E. Protein folding. *Biochem J* **270**, 1-16 (1990).
6. Dill, K.A., Ozkan, S.B., Shell, M.S. & Weikl, T.R. The protein folding problem. *Annual review of biophysics* **37**, 289-316 (2008).
7. England, J.L. & Haran, G. Role of solvation effects in protein denaturation: from thermodynamics to single molecules and back. *Annu Rev Phys Chem* **62**, 257-277 (2011).
8. Anfinsen, C.B. & Scheraga, H.A. Experimental and theoretical aspects of protein folding. *Adv Protein Chem* **29**, 205-300 (1975).
9. Kishore, D., Kundu, S. & Kayastha, A.M. Thermal, chemical and pH induced denaturation of a multimeric beta-galactosidase reveals multiple unfolding pathways. *PLoS One* **7**, e50380 (2012).
10. Strucksberg, K.H., Rosenkranz, T. & Fitter, J. Reversible and irreversible unfolding of multi-domain proteins. *Biochim Biophys Acta* **1774**, 1591-1603 (2007).
11. Duy, C. & Fitter, J. Thermostability of irreversible unfolding alpha-amylases analyzed by unfolding kinetics. *J Biol Chem* **280**, 37360-37365 (2005).
12. Findlay, H.E., Rutherford, N.G., Henderson, P.J. & Booth, P.J. Unfolding free energy of a two-domain transmembrane sugar transport protein. *Proc Natl Acad Sci U S A* **107**, 18451-18456 (2010).
13. Tanford, C. Protein denaturation. *Adv Protein Chem* **23**, 121-282 (1968).
14. Iwakura, M., Nakamura, D., Takenawa, T. & Mitsuishi, Y. An approach for protein to be completely reversible to thermal denaturation even at autoclave temperatures. *Protein Eng* **14**, 583-589 (2001).
15. Galperin, M.Y. & Koonin, E.V. Who's your neighbor? New computational approaches for functional genomics. *Nature biotechnology* **18**, 609-613 (2000).
16. Brun, C. et al. Functional classification of proteins for the prediction of cellular function from a protein-protein interaction network. *Genome biology* **5**, R6 (2003).
17. Bateman, A., Coghill, P. & Finn, R.D. DUFs: families in search of function. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **66**, 1148-1152 (2010).
18. Goodacre, N.F., Gerloff, D.L. & Uetz, P. Protein domains of unknown function are essential in bacteria. *mBio* **5**, e00744-00713 (2014).
19. Nishino, K. & Yamaguchi, A. Overexpression of the response regulator evgA of the two-component signal transduction system modulates multidrug resistance conferred by multidrug resistance transporters. *J Bacteriol* **183**, 1455-1458 (2001).
20. Nishino, K., Inazumi, Y. & Yamaguchi, A. Global analysis of genes regulated by EvgA of the two-component regulatory system in Escherichia coli. *J Bacteriol* **185**, 2667-2672 (2003).
21. Nishino, K. & Yamaguchi, A. EvgA of the two-component signal transduction system modulates production of the yhiUV multidrug transporter in Escherichia coli. *J Bacteriol* **184**, 2319-2323 (2002).

22. Blattner, F.R. et al. The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453-1462 (1997).
23. Parkhill, J. et al. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature* **413**, 848-852 (2001).
24. Fricke, W.F. et al. Comparative Genomics of 28 *Salmonella enterica* Isolates: Evidence for CRISPR-Mediated Adaptive Sublineage Evolution. *Journal of Bacteriology* **193**, 3556-3568 (2011).
25. Deng, W. et al. Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2 and CT18. *J Bacteriol* **185**, 2330-2337 (2003).
26. McClelland, M. et al. Complete genome sequence of *Salmonella enterica* serovar typhimurium LT2. *Nature* **413**, 852-856 (2001).
27. Holt, K.E. et al. Multidrug-resistant *Salmonella enterica* serovar Paratyphi A harbors IncHI1 plasmids similar to those found in serovar Typhi. *Journal of Bacteriology* **189**, 4257-4264 (2007).
28. Mulvey, M.R., Boyd, D., Cloeckert, A., Ahmed, R. & Ng, L.K. Emergence of multidrug-resistant *Salmonella* Paratyphi B dT+, Canada. *Emerg Infect Dis* **10**, 1307-1310 (2004).
29. Holt, K.E. et al. Pseudogene accumulation in the evolutionary histories of *Salmonella enterica* serovars Paratyphi A and Typhi. *BMC Genomics* **10**, 36 (2009).
30. Jiang, Y. et al. Complete nucleotide sequence of *Klebsiella pneumoniae* multidrug resistance plasmid pKP048, carrying blaKPC-2, blaDHA-1, qnrB4, and armA. *Antimicrob Agents Chemother* **54**, 3967-3969 (2010).
31. Hirsch, E.B. & Tam, V.H. Detection and treatment options for *Klebsiella pneumoniae* carbapenemases (KPCs): an emerging cause of multidrug-resistant infection. *J Antimicrob Chemother* **65**, 1119-1125 (2010).
32. Saha, P., Manna, C., Das, S. & Ghosh, M. Antibiotic binding of STY3178, a yfdX protein from *Salmonella* Typhi. *Sci Rep* **6**, 21305 (2016).
33. Masuda, N. & Church, G.M. *Escherichia coli* gene expression responsive to levels of the response regulator EvgA. *J Bacteriol* **184**, 6225-6234 (2002).
34. Mullins, E.A., Sullivan, K.L. & Kappock, T.J. Function and X-ray crystal structure of *Escherichia coli* YfdE. *PLoS One* **8**, e67901 (2013).
35. Fontenot, E.M. et al. YfdW and YfdU are required for oxalate-induced acid tolerance in *Escherichia coli* K-12. *J Bacteriol* **195**, 1446-1455 (2013).
36. Parry, C.M. et al. Randomized controlled comparison of ofloxacin, azithromycin, and an ofloxacin-azithromycin combination for treatment of multidrug-resistant and nalidixic acid-resistant typhoid fever. *Antimicrob Agents Chemother* **51**, 819-825 (2007).
37. Rowe, B., Ward, L.R. & Threlfall, E.J. Multidrug-resistant *Salmonella typhi*: A worldwide epidemic. *Clinical Infectious Diseases* **24**, S106-S109 (1997).
38. Kapil, A., Sood, S., Dash, N.R., Das, B.K. & Seth, P. Ciprofloxacin in typhoid fever. *Lancet* **354**, 164 (1999).
39. Umasankar, S., Wall, R.A. & Berger, J. A case of ciprofloxacin-resistant typhoid fever. *Commun Dis Rep CDR Rev* **2**, R139-140 (1992).
40. Senthilkumar, B. & Prabakaran, G. Multidrug resistant *Salmonella typhi* in asymptomatic typhoid carriers among food handlers in Namakkal district, Tamil Nadu. *Indian J Med Microbiol* **23**, 92-94 (2005).
41. Kosinski, J., Barbato, A. & Tramontano, A. MODexplorer: an integrated tool for exploring protein sequence, structure and function relationships. *Bioinformatics* **29**, 953-954 (2013).

42. Kay, L.E., Torchia, D.A. & Bax, A. Backbone dynamics of proteins as studied by ¹⁵N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972-8979 (1989).
43. Farrow, N.A. et al. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ¹⁵N NMR relaxation. *Biochemistry* **33**, 5984-6003 (1994).
44. Saha, P., Manna, C., Chakrabarti, J. & Ghosh, M. Reversible thermal unfolding of a yfdX protein with chaperone-like activity. *Sci Rep* **6**, 29541 (2016).
45. Saha, P., Sikdar, S., Chakrabarti, J. & Ghosh, M. Response to chemical induced changes and their implication in yfdX protein. *Preprint* (2016).
46. Saha, P., Sikdar, S., Manna, C., Chakrabarti, J. & Ghosh, M. SDS induced dissociation of STY3178 oligomer. *Preprint* (2016).
47. Lakowicz, J.R. *Principles of Fluorescence Spectroscopy, 3rd ed.* Springer, Ch.8, 277-290 (2006).
48. Patel, S. & Datta, A. Steady state and time-resolved fluorescence investigation of the specific binding of two chlorin derivatives with human serum albumin. *J Phys Chem B* **111**, 10557-10562 (2007).
49. Anand, U., Kurup, L. & Mukherjee, S. Deciphering the role of pH in the binding of ciprofloxacin hydrochloride to bovine serum albumin. *Phys Chem Chem Phys* **14**, 4250-4258 (2012).
50. Kelly, S.M., Jess, T.J. & Price, N.C. How to study proteins by circular dichroism. *Bba-Proteins Proteom* **1751**, 119-139 (2005).
51. Correa, D.H.A. & Ramos, C.H.I. The use of Circular Dichroism spectroscopy to study protein folding, form and function. *Afr. J. Biochem. Res.* **3**, 164-173 (2009).
52. Cavasotto, C.N. & Phatak, S.S. Homology modeling in drug discovery: current trends and applications. *Drug discovery today* **14**, 676-683 (2009).
53. Vyas, V.K., Ukawala, R.D., Ghate, M. & Chintha, C. Homology modeling a fast tool for drug discovery: current perspectives. *Indian journal of pharmaceutical sciences* **74**, 1-17 (2012).
54. Kelley, L.A. & Sternberg, M.J.E. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**, 363-371 (2009).
55. Peng, J. & Xu, J.B. RaptorX: Exploiting structure information for protein alignment by statistical inference. *Proteins* **79**, 161-171 (2011).
56. Kallberg, M. et al. Template-based protein structure modeling using the RaptorX web server. *Nat Protoc* **7**, 1511-1522 (2012).
57. Ma, J.Z., Wang, S., Zhao, F. & Xu, J.B. Protein threading using context-specific alignment potential. *Bioinformatics* **29**, 257-265 (2013).
58. Peng, J. & Xu, J. A multiple-template approach to protein threading. *Proteins* **79**, 1930-1939 (2011).
59. Guex, N. & Peitsch, M.C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714-2723 (1997).
60. Phillips, J.C. et al. Scalable molecular dynamics with NAMD. *J Comput Chem* **26**, 1781-1802 (2005).
61. Brooks, B.R. et al. CHARMM: the biomolecular simulation program. *J Comput Chem* **30**, 1545-1614 (2009).
62. Mahoney, M.W. & Jorgensen, W.L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *Journal of Chemical Physics* **112**, 8910-8922 (2000).

63. Pierce, B.G. et al. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics* **30**(12), 1771-1773 (2014).
64. Sikdar, S., Chakrabarti, J. & Ghosh, M. A microscopic insight from conformational thermodynamics to functional ligand binding in proteins. *Mol Biosyst* **10**, 3280-3289 (2014).
65. Tande, B.M., Wagner, N.J., Mackay, M.E., Hawker, C.J. & Jeong, M. Viscosimetric, hydrodynamic, and conformational properties of dendrimers and dendrons. *Macromolecules* **34**, 8580-8585 (2001).
66. UniProt: a hub for protein information. *Nucleic Acids Res* **43**, D204-212 (2015).
67. Wilson, K. & Walker, J. Principles and Techniques of Biochemistry and Molecular Biology, Edn. 7. (Cambridge university press, New York; 2010).
68. Gasteiger, E. et al. Protein Identification and Analysis Tools on the ExPASy Server;. (In) *John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press, 571-607* (2005).
69. Morrisett, J.D., David, J.S.K., Pownall, H.J. & Gotto, A.M., Jr. Interaction of an apolipoprotein (apoLP-alanine) with phosphatidylcholine. *Biochemistry* **12**, 1290-1299 (1973).
70. Whitmore, L. & Wallace, B.A. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32**, W668-W673 (2004).
71. Lobley, A., Whitmore, L. & Wallace, B.A. DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* **18**, 211-212 (2002).
72. Whitmore, L. & Wallace, B.A. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89**, 392-400 (2008).
73. Andrade, M.A., Chacon, P., Merelo, J.J. & Moran, F. Evaluation of Secondary Structure of Proteins from UV Circular-Dichroism Spectra Using an Unsupervised Learning Neural-Network. *Protein Eng* **6**, 383-390 (1993).
74. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30 (2000).
75. Franceschini, A. et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* **41**, D808-815 (2013).
76. Ellrott, K. et al. TOPSAN: a dynamic web database for structural genomics. *Nucleic Acids Res* **39**, D494-496 (2011).
77. Claes P., D.M., Kenney A., and Vardy P. An on-line dynamic light scattering instrument for macromolecular characterization" in "Laser Light scattering in biochemistry". *SE Harding, DB Sattelle, and VA Bloomfield (eds)*, 66-76 (1992).
78. Hereld, D., Krakow, J.L., Bangs, J.D., Hart, G.W. & Englund, P.T. A phospholipase C from *Trypanosoma brucei* which selectively cleaves the glycolipid on the variant surface glycoprotein. *J Biol Chem* **261**, 13813-13819 (1986).
79. Blake-Hodek, K.A., Cassimeris, L. & Huffaker, T.C. Regulation of microtubule dynamics by Bim1 and Bik1, the budding yeast members of the EB1 and CLIP-170 families of plus-end tracking proteins. *Mol Biol Cell* **21**, 2013-2023 (2010).
80. Van-Quynh, A., Willson, S. & Bryant, R.G. Protein reorientation and bound water molecules measured by H-1 magnetic spin-lattice relaxation. *Biophysical Journal* **84**, 558-563 (2003).
81. Sun, Y. & MacRae, T.H. Small heat shock proteins: molecular structure and chaperone function. *Cell Mol Life Sci* **62**, 2460-2476 (2005).
82. Chien, A., Edgar, D.B. & Trela, J.M. Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J Bacteriol* **127**, 1550-1557 (1976).

83. Ishino, S. & Ishino, Y. DNA polymerases as useful reagents for biotechnology - the history of developmental research in the field. *Frontiers in microbiology* **5**, 465 (2014).
84. Blumlein, A. & McManus, J.J. Reversible and non-reversible thermal denaturation of lysozyme with varying pH at low ionic strength. *Bba-Proteins Proteom* **1834**, 2064-2070 (2013).
85. Meersman, F. et al. Consistent picture of the reversible thermal unfolding of hen egg-white lysozyme from experiment and molecular dynamics. *Biophys J* **99**, 2255-2263 (2010).
86. Eisenberg, M.A. & Schwert, G.W. The reversible heat denaturation of chymotrypsinogen. *J Gen Physiol* **34**, 583-606 (1951).
87. Roychoudhuri, R., Sarath, G., Zeece, M. & Markwell, J. Reversible denaturation of the soybean Kunitz trypsin inhibitor. *Arch Biochem Biophys* **412**, 20-26 (2003).
88. Yu, M.H., Weissman, J.S. & Kim, P.S. Contribution of individual side-chains to the stability of BPTI examined by alanine-scanning mutagenesis. *J Mol Biol* **249**, 388-397 (1995).
89. Wu, Y.V. & Scheraga, H.A. Studies of soybean trypsin inhibitor. II. Conformational properties. *Biochemistry* **1**, 905-911 (1962).
90. Garcia, A.F., Garcia, W., Nonato, M.C. & Araujo, A.P. Structural stability and reversible unfolding of recombinant porcine S100A12. *Biophys Chem* **134**, 246-253 (2008).
91. Wallevik, K. Reversible denaturation of human serum albumin by pH, temperature, and guanidine hydrochloride followed by optical rotation. *J Biol Chem* **248**, 2650-2655 (1973).
92. Pico, G.A. Thermodynamic features of the thermal unfolding of human serum albumin. *Int J Biol Macromol* **20**, 63-73 (1997).
93. Benitez-Cardoza, C.G., Rojo-Dominguez, A. & Hernandez-Arana, A. Temperature-induced denaturation and renaturation of triosephosphate isomerase from *Saccharomyces cerevisiae*: evidence of dimerization coupled to refolding of the thermally unfolded protein. *Biochemistry* **40**, 9049-9058 (2001).
94. Camarillo-Cadena, M., Garza-Ramos, G., Peimbert, M., Perez-Hernandez, G. & Zubillaga, R.A. Thermal denaturation of beta-glucosidase B from *Paenibacillus polymyxa* proceeds through a Lumry-Eyring mechanism. *Protein J* **30**, 318-323 (2011).
95. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).
96. Johnson, B.A. & Blevins, R.A. NMR View: A computer program for the visualization and analysis of NMR data. *J Biomol NMR* **4**, 603-614 (1994).
97. Freire, E., van Osdol, W.W., Mayorga, O.L. & Sanchez-Ruiz, J.M. Calorimetrically determined dynamics of complex unfolding transitions in proteins. *Annu Rev Biophys Biophys Chem* **19**, 159-188 (1990).
98. Tello-Solis, S.R. & Hernandez-Arana, A. Effect of irreversibility on the thermodynamic characterization of the thermal denaturation of *Aspergillus saitoi* acid proteinase. *Biochem J* **311**, 969-974 (1995).
99. Kajitani, S., Fukuoka, M. & Sakai, N. Kinetics of thermal denaturation of protein in cured pork meat. *Japan Journal of Food Engineering* **12**, 19-26 (2011).
100. Zoldak, G., Zubrik, A., Musatov, A., Stupak, M. & Sedlak, E. Irreversible thermal denaturation of glucose oxidase from *Aspergillus niger* is the transition to the denatured state with residual structure. *Journal of Biological Chemistry* **279**, 47601-47609 (2004).
101. Banerjee, T. & Kishore, N. A differential scanning calorimetric study on the irreversible thermal unfolding of concanavalin A. *Thermochim Acta* **411**, 195-201 (2004).

102. Di Russo, N.V., Estrin, D.A., Marti, M.A. & Roitberg, A.E. pH-Dependent conformational changes in proteins and their effect on experimental pK(a)s: the case of Nitrophorin 4. *PLoS computational biology* **8**, e1002761 (2012).
103. Srivastava, J., Barber, D.L. & Jacobson, M.P. Intracellular pH sensors: design principles and functional significance. *Physiology* **22**, 30-39 (2007).
104. Talley, K. & Alexov, E. On the pH-optimum of activity and stability of proteins. *Proteins* **78**, 2699-2706 (2010).
105. Antosiewicz, J., McCammon, J.A. & Gilson, M.K. Prediction of pH-dependent properties of proteins. *J Mol Biol* **238**, 415-436 (1994).
106. Privalov, P.L. Stability of proteins: small globular proteins. *Adv Protein Chem* **33**, 167-241 (1979).
107. Catanzano, F., Graziano, G., Capasso, S. & Barone, G. Thermodynamic analysis of the effect of selective monodeamidation at asparagine 67 in ribonuclease A. *Protein Sci* **6**, 1682-1693 (1997).
108. Catanzano, F., Graziano, G., Fusi, P., Tortora, P. & Barone, G. Differential scanning calorimetry study of the thermodynamic stability of some mutants of Sso7d from *Sulfolobus solfataricus*. *Biochemistry* **37**, 10493-10498 (1998).
109. Dill, K.A. Dominant forces in protein folding. *Biochemistry* **29**, 7133-7155 (1990).
110. Graziano, G., Catanzano, F. & Barone, G. On the pH dependence of thermodynamic stability of α -amylase inhibitor tendamistat. *Thermochim Acta* **345**, 59-66 (2000).
111. O'Brien, E.P., Brooks, B.R. & Thirumalai, D. Effects of pH on proteins: predictions for ensemble and single-molecule pulling experiments. *J Am Chem Soc* **134**, 979-987 (2012).
112. Rosemann, M., Jencks, W. Interaction of urea and other polar compounds in water. *Journal of american chemical society* **97**, 631-640 (1975).
113. Caballero-Herrera, A., Nordstrand, K., Berndt, K.D. & Nilsson, L. Effect of urea on peptide conformation in water: molecular dynamics and experimental characterization. *Biophys J* **89**, 842-857 (2005).
114. Hedwig, G.R., Lilley, T.H. & Linsdell, H. Calorimetric and Volumetric Studies of the Interactions of Some Amides in Water and in 6 Mol Dm(-3) Aqueous Guanidinium Chloride. *J Chem Soc Faraday T* **87**, 2975-2982 (1991).
115. Greene, R.F., Jr. & Pace, C.N. Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin, and beta-lactoglobulin. *J Biol Chem* **249**, 5388-5393 (1974).
116. Montserret, R., McLeish, M.J., Bockmann, A., Geourjon, C. & Penin, F. Involvement of electrostatic interactions in the mechanism of peptide folding induced by sodium dodecyl sulfate binding. *Biochemistry* **39**, 8362-8373 (2000).
117. Reynolds, J.A. & Tanford, C. The gross conformation of protein-sodium dodecyl sulfate complexes. *J Biol Chem* **245**, 5161-5165 (1970).
118. Mattice, W.L., Riser, J.M. & Clark, D.S. Conformational properties of the complexes formed by proteins and sodium dodecyl sulfate. *Biochemistry* **15**, 4264-4272 (1976).
119. Wu, C.S. & Yang, J.T. Conformation of naturally-occurring peptides in surfactant solution: its relation to the structure-forming potential of amino acid sequence. *Biochem Biophys Res Commun* **82**, 85-91 (1978).
120. Yu, M.H., Weissman, J.S. & Kim, P.S. Contribution of Individual Side-Chains to the Stability of Bpti Examined by Alanine-Scanning Mutagenesis. *Journal of Molecular Biology* **249**, 388-397 (1995).

121. Haque, M.A. et al. In vitro and in silico studies of urea-induced denaturation of yeast iso-1-cytochrome c and its deletants at pH 6.0 and 25 degrees C. *J Biomol Struct Dyn*, 1-10 (2014).
122. Anand, U., Jash, C. & Mukherjee, S. Protein unfolding and subsequent refolding: a spectroscopic investigation. *Phys Chem Chem Phys* **13**, 20418-20426 (2011).
123. El Kadi, N. et al. Unfolding and refolding of bovine serum albumin at acid pH: ultrasound and structural studies. *Biophys J* **91**, 3397-3404 (2006).
124. Patra, M., Mukhopadhyay, C. & Chakrabarti, A. Probing conformational stability and dynamics of erythroid and nonerythroid spectrin: effects of urea and guanidine hydrochloride. *PLoS One* **10**, e0116991 (2015).
125. Guidry, J.J., Moczygemba, C.K., Steede, N.K., Landry, S.J. & Wittung-Stafshede, P. Reversible denaturation of oligomeric human chaperonin 10: denatured state depends on chemical denaturant. *Protein Sci* **9**, 2109-2117 (2000).
126. Faham, S. et al. Side-chain contributions to membrane protein structure and stability. *J Mol Biol* **335**, 297-305 (2004).
127. Curnow, P. & Booth, P.J. Combined kinetic and thermodynamic analysis of alpha-helical membrane protein unfolding. *Proc Natl Acad Sci U S A* **104**, 18970-18975 (2007).
128. Lau, F.W. & Bowie, J.U. A method for assessing the stability of a membrane protein. *Biochemistry* **36**, 5884-5892 (1997).
129. Barrera, F.N. et al. Unfolding and refolding in vitro of a tetrameric, alpha-helical membrane protein: the prokaryotic potassium channel KcsA. *Biochemistry* **44**, 14344-14352 (2005).
130. Otzen, D.E. Folding of DsbB in mixed micelles: a kinetic analysis of the stability of a bacterial membrane protein. *J Mol Biol* **330**, 641-649 (2003).
131. Roman, E.A., Arguello, J.M. & Gonzalez Flecha, F.L. Reversible unfolding of a thermophilic membrane protein in phospholipid/detergent mixed micelles. *J Mol Biol* **397**, 550-559 (2010).
132. Harris, N.J. & Booth, P.J. Folding and stability of membrane transport proteins in vitro. *Biochim Biophys Acta* **1818**, 1055-1066 (2012).
133. Miller, D. et al. In vitro unfolding and refolding of the small multidrug transporter EmrE. *J Mol Biol* **393**, 815-832 (2009).
134. Nielsen, M.M., Andersen, K.K., Westh, P. & Otzen, D.E. Unfolding of beta-sheet proteins in SDS. *Biophys J* **92**, 3674-3685 (2007).
135. Mandal, B., Ghosh, S. & Moulik, S.P. Detailed characterization of lysozyme (Lyz)-surfactant (SDDS) interaction and the structural transitions. *New J Chem* **40**, 4617-4624 (2016).
136. Anand, U. & Mukherjee, S. Reversibility in protein folding: effect of beta-cyclodextrin on bovine serum albumin unfolded by sodium dodecyl sulphate. *Phys Chem Chem Phys* **15**, 9375-9383 (2013).
137. Breyton, C., Tribet, C., Olive, J., Dubacq, J.P. & Popot, J.L. Dimer to monomer conversion of the cytochrome b6 f complex. Causes and consequences. *J Biol Chem* **272**, 21892-21900 (1997).
138. Bay, D.C., Budiman, R.A., Nieh, M.P. & Turner, R.J. Multimeric forms of the small multidrug resistance protein EmrE in anionic detergent. *Biochim Biophys Acta* **1798**, 526-535 (2010).
139. Biverstal, H. et al. Dissociation of a BRICHOS trimer into monomers leads to increased inhibitory effect on Abeta42 fibril formation. *Biochim Biophys Acta* **1854**, 835-843 (2015).
140. Gelamo, E.L., Itri, R. & Tabak, M. Small angle x-ray scattering (SAXS) study of the extracellular Hemoglobin of *Glossoscolex paulistus*: effect of pH, iron oxidation state, and interaction with anionic SDS surfactant. *J Biol Chem* **279**, 33298-33305 (2004).

141. Kwon, Y.M., Baudys, M., Knutson, K. & Kim, S.W. In situ study of insulin aggregation induced by water-organic solvent interface. *Pharmaceutical research* **18**, 1754-1759 (2001).
142. Mammi, S. & Peggion, E. Conformational studies of human [15-2-aminohexanoic acid]little gastrin in sodium dodecyl sulfate micelles by ¹H NMR. *Biochemistry* **29**, 5265-5269 (1990).
143. Rizo, J., Blanco, F.J., Kobe, B., Bruch, M.D. & Gierasch, L.M. Conformational behavior of Escherichia coli OmpA signal peptides in membrane mimetic environments. *Biochemistry* **32**, 4881-4894 (1993).
144. Xu, Q. & Keiderling, T.A. Effect of sodium dodecyl sulfate on folding and thermal stability of acid-denatured cytochrome c: a spectroscopic approach. *Protein Sci* **13**, 2949-2959 (2004).
145. Zhong, L. & Johnson, W.C., Jr. Environment affects amino acid preference for secondary structure. *Proc Natl Acad Sci U S A* **89**, 4462-4465 (1992).
146. Waterhous, D.V. & Johnson, W.C., Jr. Importance of environment in determining secondary structure in proteins. *Biochemistry* **33**, 2121-2128 (1994).
147. Hiramatsu, K. & Yang, J.T. Cooperative binding of hexadecyltrimethylammonium chloride and sodium dodecyl sulfate to cytochrome c and the resultant change in protein conformation. *Biochim Biophys Acta* **743**, 106-114 (1983).
148. Gelamo, E.L. & Tabak, M. Spectroscopic studies on the interaction of bovine (BSA) and human (HSA) serum albumins with ionic surfactants. *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy* **56A**, 2255-2271 (2000).
149. Gelamo, E.L., Silva, C.H., Imasato, H. & Tabak, M. Interaction of bovine (BSA) and human (HSA) serum albumins with ionic surfactants: spectroscopy and modelling. *Biochim Biophys Acta* **1594**, 84-99 (2002).
150. Gebicka, L. & Gebicki, J.L. Kinetic studies on the interaction of ferricytochrome c with anionic surfactants. *J Protein Chem* **18**, 165-172 (1999).
151. Das, T.K., Mazumdar, S. & Mitra, S. Characterization of a partially unfolded structure of cytochrome c induced by sodium dodecyl sulphate and the kinetics of its refolding. *Eur J Biochem* **254**, 662-670 (1998).
152. Muga, A., Mantsch, H.H. & Surewicz, W.K. Membrane binding induces destabilization of cytochrome c structure. *Biochemistry* **30**, 7219-7224 (1991).
153. Takeda, K., Takahashi, K. & Batra, P.P. Kinetic aspects of the interaction of horse heart cytochrome c with sodium dodecyl sulfate. *Arch Biochem Biophys* **236**, 411-417 (1985).
154. Gutierrez, C. & Devedjian, J.C. Osmotic induction of gene *osmC* expression in Escherichia coli K12. *J Mol Biol* **220**, 959-973 (1991).
155. Oh, J.T. et al. Osmotic stress in viable Escherichia coli as the basis for the antibiotic response by polymyxin B. *Biochem Biophys Res Commun* **246**, 619-623 (1998).
156. Wang, Y. The function of OmpA in Escherichia coli. *Biochem Biophys Res Commun* **292**, 396-401 (2002).
157. Greenfield, N.J. Determination of the folding of proteins as a function of denaturants, osmolytes or ligands using circular dichroism. *Nat Protoc* **1**, 2733-2741 (2006).
158. Ahmad, S. & Rao, N.M. Thermally denatured state determines refolding in lipase: mutational analysis. *Protein Sci* **18**, 1183-1196 (2009).
159. Shearstone, J.R. & Baneyx, F. Biochemical characterization of the small heat shock protein IbpB from Escherichia coli. *J Biol Chem* **274**, 9937-9945 (1999).
160. Boudker, O., Todd, M.J. & Freire, E. The structural stability of the co-chaperonin GroES. *Journal of Molecular Biology* **272**, 770-779 (1997).

161. Schlapschy, M. et al. The periplasmic E. coli chaperone Skp is a trimer in solution: biophysical and preliminary crystallographic characterization. *Biological chemistry* **385**, 137-143 (2004).
162. Scholz, C. et al. SlyD proteins from different species exhibit high prolyl isomerase and chaperone activities. *Biochemistry* **45**, 20-33 (2006).
163. Panse, V.G., Swaminathan, C.P., Aloor, J.J., Surolia, A. & Varadarajan, R. Unfolding thermodynamics of the tetrameric chaperone, SecB. *Biochemistry* **39**, 2362-2369 (2000).
164. Greenfield, N.J. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat Protoc* **1**, 2527-2535 (2006).
165. Weinreb, O., van Boekel, M.A., Dovrat, A. & Bloemendal, H. Effect of UV-A light on the chaperone-like properties of young and old lens alpha-crystallin. *Investigative ophthalmology & visual science* **41**, 191-198 (2000).
166. Lewis, K. Multidrug resistance: Versatile drug sensors of bacterial cells. *Curr Biol* **9**, R403-407 (1999).
167. Sikora, C.W. & Turner, R.J. SMR proteins SugE and EmrE bind ligand with similar affinity and stoichiometry. *Biochem Biophys Res Commun* **335**, 105-111 (2005).
168. Yamasaki, S. et al. The crystal structure of multidrug-resistance regulator RamR with multiple drugs. *Nat Commun* **4**, 2078 (2013).
169. Sikora, C.W. & Turner, R.J. Investigation of ligand binding to the multidrug resistance protein EmrE by isothermal titration calorimetry. *Biophys J* **88**, 475-482 (2005).
170. Long, F., Rouquette-Loughlin, C., Shafer, W.M. & Yu, E.W. Functional cloning and characterization of the multidrug efflux pumps NorM from *Neisseria gonorrhoeae* and YdhE from *Escherichia coli*. *Antimicrob Agents Chemother* **52**, 3052-3060 (2008).
171. Kamat, B.P. & Seetharamappa, J. Mechanism of interaction of vincristine sulphate and rifampicin with bovine serum albumin: A spectroscopic study. *J Chem Sci* **117**, 649-655 (2005).
172. Mihaela Bacalum, H.W., M. Radu Interaction between ceftazidime and bacterial porin OmpF analyzed by fluorescence. *Romanian Journal of Biophysics* **19**, 105-116 (2009).
173. Basha, E., Jones, C., Wysocki, V. & Vierling, E. Mechanistic Differences between Two Conserved Classes of Small Heat Shock Proteins Found in the Plant Cytosol. *Journal of Biological Chemistry* **285**, 11489-11497 (2010).
174. Bachas, S., Eginton, C., Gunio, D. & Wade, H. Structural contributions to multidrug recognition in the multidrug resistance (MDR) gene regulator, BmrR. *Proc Natl Acad Sci U S A* **108**, 11046-11051 (2011).
175. Nakashima, R., Sakurai, K., Yamasaki, S., Nishino, K. & Yamaguchi, A. Structures of the multidrug exporter AcrB reveal a proximal multisite drug-binding pocket. *Nature* **480**, 565-569 (2011).
176. Ziervogel, B.K. & Roux, B. The binding of antibiotics in OmpF porin. *Structure* **21**, 76-87 (2013).
177. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J Mol Biol* **215**, 403-410 (1990).
178. Petersen, T.N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods* **8**, 785-786 (2011).
179. Yu, C.S., Lin, C.J. & Hwang, J.K. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci* **13**, 1402-1406 (2004).

180. Yu, C.S., Chen, Y.C., Lu, C.H. & Hwang, J.K. Prediction of protein subcellular localization. *Proteins* **64**, 643-651 (2006).
181. Goldberg, T. et al. LocTree3 prediction of localization. *Nucleic Acids Res* **42**, W350-355 (2014).
182. Wulfing, C. & Pluckthun, A. Protein folding in the periplasm of Escherichia coli. *Mol Microbiol* **12**, 685-692 (1994).
183. Bodelon, G., Marin, E. & Fernandez, L.A. Role of periplasmic chaperones and BamA (YaeT/Omp85) in folding and secretion of intimin from enteropathogenic Escherichia coli strains. *J Bacteriol* **191**, 5169-5179 (2009).
184. Sklar, J.G., Wu, T., Kahne, D. & Silhavy, T.J. Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli. *Genes Dev* **21**, 2473-2484 (2007).
185. Zhang, H.M., Tang, B.P., Wang, Y.Q., Zhang, D.Z. & Zhang, H.B. Studies of the interaction between ciprofloxacin and the hemocyanin from chinese mitten crab Eriocheir Japonica Sinensis. *Anal Lett* **44**, 2094-2106 (2011).
186. Su, C.C. & Yu, E.W. Ligand-transporter interaction in the AcrB multidrug efflux pump determined by fluorescence polarization assay. *FEBS Lett* **581**, 4972-4976 (2007).
187. Keswani, N. & Kishore, N. Calorimetric and spectroscopic studies on the interaction of anticancer drug mitoxantrone with human serum albumin. *J Chem Thermodyn* **43**, 1406-1413 (2011).
188. Subramanyam, R. et al. Novel binding studies of human serum albumin with trans-feruloyl maslinic acid. *J Photoch Photobio B* **95**, 81-88 (2009).
189. Mishra, B., Barik, A., Priyadarsini, K.I. & Mohan, H. Fluorescence spectroscopic studies on binding of a flavonoid antioxidant quercetin to serum albumins. *J Chem Sci* **117**, 641-647 (2005).
190. Yue, Y., Chen, X., Qin, J. & Yao, X. Spectroscopic investigation on the binding of antineoplastic drug oxaliplatin to human serum albumin and molecular modeling. *Colloids Surf B Biointerfaces* **69**, 51-57 (2009).
191. Agudelo, D. et al. Probing the binding sites of antibiotic drugs doxorubicin and N-(trifluoroacetyl) doxorubicin with human and bovine serum albumins. *PLoS One* **7**, e43814 (2012).
192. Mukherjee, M. et al. A Comparative Study of Interaction of Tetracycline with Several Proteins Using Time Resolved Anisotropy, Phosphorescence, Docking and FRET. *Plos One* **8** (2013).
193. Alam, P. et al. Biophysical and molecular docking insight into the interaction of cytosine beta-D arabinofuranoside with human serum albumin. *J Lumin* **164**, 123-130 (2015).
194. Tian, J.N. et al. Probing the binding of scutellarin to human serum albumin by circular dichroism, fluorescence spectroscopy, FTIR, and molecular modeling method. *Biomacromolecules* **5**, 1956-1961 (2004).
195. Subramanyam, R., Gollapudi, A., Bonigala, P., Chinnaboina, M. & Amooru, D.G. Betulinic acid binding to human serum albumin: A study of protein conformation and binding affinity. *J Photoch Photobio B* **94**, 8-12 (2009).
196. Li, X.R. & Wang, S. Study on the interaction of (+)-catechin with human serum albumin using isothermal titration calorimetry and spectroscopic techniques. *New J Chem* **39**, 386-395 (2015).
197. Estrada, J.G. & Sanchez-Delgado, R.A. Spectroscopic Study of the Interactions of Ruthenium-Ketoconazole Complexes of Known Antiparasitic Activity with Human Serum Albumin and Apotransferrin. *Journal of the Mexican Chemical Society* **57**, 169-174 (2013).

198. Zaidi, N., Ajmal, M.R., Rabbani, G., Ahmad, E. & Khan, R.H. A Comprehensive Insight into Binding of Hippuric Acid to Human Serum Albumin: A Study to Uncover Its Impaired Elimination through Hemodialysis. *Plos One* **8** (2013).
199. Ishtikhar, M. et al. Biophysical and molecular docking insight into interaction mechanism and thermal stability of human serum albumin isoforms with a semi-synthetic water-soluble camptothecin analog irinotecan hydrochloride. *J Biomol Struct Dyn*, 1-49 (2015).
200. Chatterjee, T., Pal, A., Dey, S., Chatterjee, B.K. & Chakrabarti, P. Interaction of virstatin with human serum albumin: spectroscopic analysis and molecular modeling. *PLoS One* **7**, e37468 (2012).
201. Iranfar, H., Rajabi, O., Salari, R. & Chamani, J. Probing the interaction of human serum albumin with ciprofloxacin in the presence of silver nanoparticles of three sizes: multispectroscopic and zeta potential investigation. *J Phys Chem B* **116**, 1951-1964 (2012).
202. Jha, N.S. & Kishore, N. Binding of streptomycin with bovine serum albumin: Energetics and conformational aspects. *Thermochim Acta* **482**, 21-29 (2009).
203. Thoppil, A.A., Sharma, R. & Kishore, N. Complexation of beta-lactam antibiotic drug carbenicillin to bovine serum albumin: energetics and conformational studies. *Biopolymers* **89**, 831-840 (2008).
204. Seedher, N. & Agarwal, P. Mechanism of interaction of some antimicrobial agents with bovine serum albumin. *Journal of Global Pharma Technology* **2(3)**, 77-86 (2010).
205. Tong, J.Q. et al. Probing the adverse temperature dependence in the static fluorescence quenching of BSA induced by a novel anticancer hydrazone. *Photoch Photobio Sci* **11**, 1868-1879 (2012).
206. Mohammad A Rashid¹, S.N.I.R., Tania Sultana, Md. Zamil Sultan and Md. Zakir Sultan Fluorescence Spectroscopic Study of Interaction between Olanzapine and Bovine Serum Albumin. *Pharmaceutica Analytica Acta* **6** (2015).
207. Ojha, H., Mishra, K., Hassan, M.I. & Chaudhury, N.K. Spectroscopic and isothermal titration calorimetry studies of binding interaction of ferulic acid with bovine serum albumin. *Thermochim Acta* **548**, 56-64 (2012).
208. SHARMA, R.N., PANCHOLI, SHYAM S. Protein binding interaction study of olmesartan medoxomil and its metabolite olmesartan by fluorescence spectroscopy. *International Journal of Pharmacy and Pharmaceutical Sciences* **6**, 726-729 (2014).
209. Dutta, S.K., Basu, S.K. & Sen, K.K. Binding of diclofenac sodium with bovine serum albumin at different temperatures, pH and ionic strengths. *Indian J Exp Biol* **44**, 123-127 (2006).
210. H, O. A specific cephalosporin-binding protein of *Citobacter Freundii*. *Biochim Biophys Acta* **420**, 155-164 (1976).
211. Velkov, T. et al. The interaction of lipophilic drugs with intestinal fatty acid-binding protein. *J Biol Chem* **280**, 17769-17776 (2005).
212. Neves, P., Berkane, E., Gameiro, P., Winterhalter, M. & de Castro, B. Interaction between quinolones antibiotics and bacterial outer membrane porin OmpF. *Biophys Chem* **113**, 123-128 (2005).
213. Fernandes, F., Neves, P., Gameiro, P., Loura, L.M. & Prieto, M. Ciprofloxacin interactions with bacterial protein OmpF: modelling of FRET from a multi-tryptophan protein trimer. *Biochim Biophys Acta* **1768**, 2822-2830 (2007).
214. Chi, Z., Liu, R. & Zhang, H. Noncovalent interaction of oxytetracycline with the enzyme trypsin. *Biomacromolecules* **11**, 2454-2459 (2010).

215. Fuda, C., Suvorov, M., Vakulenko, S.B. & Mobashery, S. The basis for resistance to beta-lactam antibiotics by penicillin-binding protein 2a of methicillin-resistant *Staphylococcus aureus*. *J Biol Chem* **279**, 40802-40806 (2004).
216. Fuda, C. et al. Mechanistic basis for the action of new cephalosporin antibiotics effective against methicillin- and vancomycin-resistant *Staphylococcus aureus*. *J Biol Chem* **281**, 10035-10041 (2006).
217. Fishovitz, J. et al. Disruption of Allosteric Response as an Unprecedented Mechanism of Resistance to Antibiotics. *Journal of the American Chemical Society* **136**, 9814-9817 (2014).
218. Yu, S. et al. Atomic resolution structure of EhpR: phenazine resistance in *Enterobacter agglomerans* Eh1087 follows principles of bleomycin/mitomycin C resistance in other bacteria. *BMC Struct Biol* **11**, 33 (2011).
219. Frere, J.M., Geurts, F. & Ghuysen, J.M. The exocellular DD-carboxypeptidase-endopeptidase of *Streptomyces albus* G. Interaction with beta-lactam antibiotics. *Biochem J* **175**, 801-805 (1978).
220. Fan, X. et al. Diversity of penicillin-binding proteins. Resistance factor FmtA of *Staphylococcus aureus*. *J Biol Chem* **282**, 35143-35152 (2007).
221. Suprita Sinha, S.A.a.T.K.M. Study of plasma protein binding activity of isometamidium and its impact on anthelmintic activity using trypanosoma induced calf model. *Vet World* **6** 7, 444-448 (2013).
222. Dalm, D., Palm, G.J., Aleksandrov, A., Simonson, T. & Hinrichs, W. Nonantibiotic properties of tetracyclines: structural basis for inhibition of secretory phospholipase A2. *J Mol Biol* **398**, 83-96 (2010).
223. Stark, J.L. et al. Solution structure and function of YndB, an AHSA1 protein from *Bacillus subtilis*. *Proteins* **78**, 3328-3340 (2010).
224. Khatua, B., Van Vleet, J., Choudhury, B.P., Chaudhry, R. & Mandal, C. Sialylation of outer membrane porin protein D: a mechanistic basis of antibiotic uptake in *Pseudomonas aeruginosa*. *Molecular & cellular proteomics : MCP* **13**, 1412-1428 (2014).
225. Stark, J.L. et al. Identification of low-molecular-weight compounds inhibiting growth of corynebacteria: potential lead compounds for antibiotics. *ChemMedChem* **9**, 282-285 (2014).
226. Yin, F., Cao, R., Goddard, A., Zhang, Y. & Oldfield, E. Enthalpy versus entropy-driven binding of bisphosphonates to farnesyl diphosphate synthase. *J Am Chem Soc* **128**, 3524-3525 (2006).
227. Alguel, Y. et al. Crystal structures of multidrug binding protein TtgR in complex with antibiotics and plant antimicrobials. *J Mol Biol* **369**, 829-840 (2007).
228. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195-201 (2006).
229. Eleftheriou, M., Germain, R.S., Royyuru, A.K. & Zhou, R.H. Thermal denaturing of mutant lysozyme with both the OPLSAA and the CHARMM force fields. *Journal of the American Chemical Society* **128**, 13388-13395 (2006).

List of Publications

1. **Saha, P.**, Manna, C., Das, S. & Ghosh, M. Antibiotic binding of STY3178, a yfdX protein from *Salmonella Typhi*. *Scientific Reports* **6**, 21305 (2016).
2. **Saha, P.**, Manna, C., Chakrabarti, J. & Ghosh, M. Reversible thermal unfolding of a yfdX protein with chaperone-like activity. *Scientific Reports* **6**, 29541 (2016).
3. **Saha, P.**, Sikdar, S., Chakrabarti, J. & Ghosh, M. Response to chemical induced changes and their implication in yfdX protein. *Submitted* (2016).
4. **Saha, P.**, Sikdar, S., Manna, C., Chakrabarti, J. & Ghosh, M. SDS induced dissociation of STY3178 oligomer. *Preprint* (2016).