

STRUCTURAL MODIFICATIONS IN AN ARCHAEAL SMALL HEAT
SHOCK PROTEIN TO REVEAL MOLECULAR BASIS OF SUBSTRATE
TARGETING AND BINDING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

AZRA RAFIQ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
BIOCHEMISTRY

MAY 2022

Approval of the thesis:

STRUCTURAL MODIFICATIONS IN AN ARCHAEL SMALL HEAT SHOCK PROTEIN TO REVEAL MOLECULAR BASIS OF SUBSTRATE TARGETING AND BINDING

Submitted by **AZRA RAFIQ** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Biochemistry, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences**

Assoc. Prof. Dr. Özgül Persil Çetinkol
Head of the Department, **Biochemistry**

Prof. Dr. Semra Kocabıyık
Supervisor, **Biology, METU**

Assoc. Prof. Dr. Erhan Bat
Co-Supervisor, **Chemical Eng., METU**

Examining Committee Members:

Assoc. Prof. Dr. Erkan Kiriş
Biology, METU

Prof. Dr. Semra Kocabıyık
Biology, METU.

Prof. Dr. Özlem Darcansoy İşeri
Mol. Biology and Gen., Başkent University

Assoc. Prof. Dr. Tülin Yanık
Biology, METU.

Assoc. Prof. Dr. Sezer Okay
Vaccine Institute, Hacettepe University.

Date: 12.05.2022

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name Last name: Azra Rafiq

Signature:

ABSTRACT

STRUCTURAL MODIFICATIONS IN AN ARCHAEAL SMALL HEAT SHOCK PROTEIN TO REVEAL MOLECULAR BASIS OF SUBSTRATE TARGETING AND BINDING

Rafiq, Azra
Doctor of Philosophy, Biochemistry
Supervisor: Prof. Dr. Semra Kocabıyık
Co-Supervisor: Assoc. Prof. Dr. Erhan Bat

May 2022, 292 pages

The N-terminal domain (NTD) of small heat shock protein (sHSP) from archaea, *Thermoplasma volcanium*, is highly hydrophobic in the proximal and distal end. It is 32 amino acids long and has two highly conserved glutamic acid residues at positions 11 and 22. The 3-D model of the dimer generated by homology modelling predicted the NTD to be running away from each other and the dimer interface being formed mainly between $\beta 2$ and $\beta 6$ sheets of the Alpha Crystallin Domain (ACD). The decrease in hydrophobicity in the proximal region of the NTD decreased the chaperone activity of *Tpv* sHSP 14.3 while increasing hydrophilicity at the distal end increased the chaperone activity of the sHSP as compared to wild-type (WT). However, abolishing the negative charge at position 11 and 22 not only decreased the thermal stability of the resultant mutants but also decreased the chaperone activity. In addition, these mutations increased the propensity of the NTD to form fibril like structures as well. In the NTD, at position 12, when hydrophobicity was decreased by M12T mutation, not only the chaperone activity was decreased, but also, at high concentration the sHSP protein itself aggregated, at high temperature, along with the substrate. Moreover, M12T mutation also led to inclination of the

NTD towards fibril structure formation. The length of the disordered region increased with the decrease in hydrophobicity in the proximal, as well as the distal NTD. However, abolishing the negative charge resulted in spatial rearrangements of the NTD, without affecting the disordered region. *Tpv* sHSP 14.3 WT protein possessed an oligomeric profile composed of 24-mer, in addition to 60-mer, dodecamer and the active dimeric forms. The I5T mutation alone, or in combination with F7SF8Y mutations, resulted in disappearance of 24-mer oligomer and a new species 36-meric form is formed. However, F7SF8Y double mutation resulted in disappearance of all species larger than 24-mer. In addition, for the first time, in this study it has been found that the whole NTD of *Tpv* sHSP 14.3, an archaeal sHSP, also possessed chaperone activity. Moreover, it was shown that a section of NTD, the initial 12 amino acids, possessed protection capacity for substrate against heat induced aggregation, equivalent to the intact NTD.

Keywords: *Tpv* sHSP 14.3, N-terminal domain (NTD), *Thermoplasma volcanium*, Hydrophobicity, Mutagenesis, Chaperone activity, 3-D model structure analysis.

ÖZ

ARKEA KÖKENLİ BİR KÜÇÜK ISI ŞOKU PROTEİNİNDE SUBSTRAT TANINMASI VE BAĞLANMASININ MOLEKÜLER TEMELİNİ BELİRLEMEK AMAÇLI YAPISAL MODİFİKASYONLAR

Rafiq, Azra
Doktora, Biyokimya
Tez Yöneticisi: Prof. Dr. Semra Kocabıyık
Ortak Tez Yöneticisi: Doç. Dr. Erhan Bat

Mayıs 2022, 292 sayfa

Archaea *Thermoplasma volcanium*'un küçük ısı şoku proteinin (*Tpv* sHSP 14.3) N-terminal domaininin (NTD), N-terminal yakın ve N-terminal uzak ucu oldukça hidrofobiktir. Uzunluğu 32 amino asit olup ve 11 ve 22 konumlarında yüksek oranda korunmuş iki glutamik asit kalıntısına sahiptir. Homoloji modellemesi ile oluşturulan dimerin 3D modeli, NTD'nin birbirinden ayrılacak şekilde konumlandığını ve dimer ara yüzünün esas olarak Alfa Kristalin Domain (ACD)'nin $\beta 2$ ile $\beta 6$ yaprakları arasında oluştuğunu ön görmektedir. NTD'nin N-terminal yakın bölgesindeki hidrofobiklikteki azalma, *Tpv* sHSP 14.3'ün şaperon aktivitesini azaltırken, N-terminal uzak uçta hidrofilikliğin artması, şaperon aktivitesini yabanıl tiple (WT) kıyasla arttırmıştır. Ancak, 11. ve 22. pozisyonlarda negatif yok edilmesi sadece oluşan mutantların ısıl stabilitesini azaltmakla kalmamış, aynı zamanda şaperon aktivitesini de azaltmıştır. Ek olarak, bu mutasyonlar NTD'nin fibril benzeri yapılar oluşturma eğilimini de arttırmıştır. NTD'de, 12. pozisyonda, hidrofobiklik M12T mutasyonu ile azaltıldığından, sadece şaperon aktivitesi azalmamış, aynı zamanda yüksek sıcaklıkta ve konsantrasyonda sHSP proteininin kendisi, substrat ile birlikte kümelenmiştir. Ayrıca, M12T mutasyonu, NTD'nin fibril yapı

oluşturmaya doğru eğilim göstermesine de yol açmıştır. Hidrofobikliğin N-terminal yakın ve uzak uçlarda azalması düzensiz bölgenin uzunluğunu arttırmıştır. Fakat, negatif yükün kalkması düzensiz bölgeyi etkilemezken, NTD'nin üç boyutta yeniden düzenlenmesi ile sonuçlanmıştır. *Tpv* sHSP 14.3 WT protein, 60-mer, dodekamer (12 alt üniteli) ve aktif dimerik formlara ek olarak 24-mer'den oluşan bir oligomerik profile sahiptir. I5T mutasyonu tek başına veya F7SF8Y mutasyonları ile kombinasyon halinde 24-mer'in kaybolmasına yol açmış ve yeni bir 36-merik form oluşmuştur. Ancak, F7SF8Y çift mutasyonu, 24-mer'den büyük olan tüm formların yok olmasına sebep olmuştur. Ayrıca, bu çalışmada ilk kez, bir arke sHSP olan *Tpv* sHSP 14.3'ün NTD'sinin bütünü de şaperon aktivitesine sahip olduğu bulunmuştur. Buna ek olarak, NTD'nin ilk 12 amino asitlik bir bölümünün, ısı indüklenmeli kümelenmeye karşı tüm NTD'ye eşdeğer substrat koruma kapasitesine sahip olduğu gösterilmiştir.

Anahtar Kelimeler: *Thermoplasma volcanium*, *Tpv* sHSP 14.3, N-terminal Domain (NTD), Hidrofobisite, Mutagenез, Şaperon Aktivitesi, 3-D model Yapı Analizi.

To My Parents and my brother Naeem

For their struggles, sacrifices and for giving me the little they had, ensuring I would
have the opportunity for good education.

I owe my life to them.

ACKNOWLEDGMENTS

First and foremost, I thank Almighty Allah for bestowing me with strength, intelligence and opportunity to embark on this task and complete it satisfactorily.

The personality that deserves my maximum acknowledgment, sincerest gratitude and warmest thanks is my esteemed supervisor, Prof. Dr. Semra KOCABIYIK for her vigilant guidance, invaluable and expert advice, strong but creative criticism, warm encouragements and insight throughout the research. In a million times, I would have never imagined or hoped to have any other person as my supervisor and mentor in my Ph.D. Her immense knowledge and abundant experience always served as a source of light and I will, forever, be in debt to her services.

Besides my supervisor, I would also like to extend my gratitude to Assoc. Prof. Dr. Erhan Bat for being my co-supervisor during my research.

I would also like to thank all the jury members for their insightful suggestions and comments.

A heartfelt thanks to my lab mate Sema ZABCI. Words will never be enough to express my feelings, but I am grateful for the sleepless nights, petty fights, walk to and from lab, energetic and stimulating discussions, motivation sessions, funny talks, language classes and for being with me, through best and worst. I have found a sister (from Turkey) in her and (although, she threatened many a times that she won't be my friend), I am sure, our bond will be life long. I will always cherish the time I spent with her and I am also grateful to her family, particularly my (her) mother Sevil ZABCI, for all the love and prayers and delicious food (Turkish sarma and cakes) during this period.

My appreciation and thanks to the friends I made in Turkey, for their help and support, Merve Akkulak, Sibel Öztürk, Özge Demirdoğan and Yağmur Kaçer.

My gratitude also extends to my roommates throughout this period; Sonum CHAUDHARY, for all the prayers, support and love (ofcourse food, Pakistani tea and motivating advices), especially during the last few months, where she took care of me like a sister; Hira TARIQ, for being a good listener and advisor in my difficult times, Ümran AŞCI, for chit chat and venting out sessions and Buket ALTINÇELEP for the positive energy, always.

A big, warm and full of love, gratitude to my family who have loved me unconditionally and have been my backbone, throughout my life. No words will be enough to express my feelings of appreciation and gratitude to my dearest father, Muhammad RAFIQ for his unshakeable trust and belief in me and for encouraging me always to pursue my dreams no matter what risk it included. Huge thanks to my love, my mother, Khursheed RAFIQ, for being my biggest critique and yet loving me deep down, for her prayers and advice. Sincerest thanks to my brother Naeem for the love and for making me the person I am, today. I extend my thanks to my brother Haroon for supporting me morally, when I was about to leave study and go, to my sister-in-law Rabia for her prayers and love, to my sister Zahra for guiding me and praying I have good friends, to my brother Abdul Hafeez for his prayers, to my dear little brother Umer for all the technical help and for keeping my secrets, to my dear little sister, Farah, my go to, in times I cried and needed attention and love and my secret-keeper, and my nephews Muhammad and Ahmad for entertaining me when I was down. I owe them all, big time.

I thank my country mates for being there whenever I needed them; my bestie Mahjabeen Khalid for being a good listener and providing unbiased advices, Zahra Ansab for an awesome company and good food especially during pandemic times, Ishfaq Bashir for support and good company and Namra Kiran for always being ready to go around.

The author wishes to express her deepest gratitude to a few people who have been a firm support, both economically and mentally, throughout this period. This piece would not have been possible without their valuable input, in many ways. Their names are confidential, but they have a special place in my heart and my prayers, today and always.

A debt of gratitude is also owed to a bunch of people who have tried to push me down, and back in time, time when I was at my lowest. It was also, because of them that I wished to and worked to get up and move forward. It was because of them that I valued myself and they made me want to fly higher and shine brighter.

In the end, my heartfelt gratitude to the Higher Education Commission, Government of Pakistan for funding my stay in Turkey.

I would also like to thank METU-BAP (DKT-108-2018-3561) for grants to support this work.

TABLE OF CONTENTS

ABSTRACT.....	v
ÖZ.....	vii
ACKNOWLEDGMENTS.....	x
TABLE OF CONTENTS.....	xiii
LIST OF TABLES.....	xviii
LIST OF FIGURES.....	xx
LIST OF ABBREVIATIONS.....	xxx
CHAPTERS	
1 INTRODUCTION.....	1
1.1 Function of sHSPs.....	7
1.2 Activation of sHSPs.....	8
1.3 Substrate – sHSP sites and Interaction.....	9
1.4 sHSP mechanism of action.....	10
1.5 Structural and functional role of NTD.....	10
1.5.1 Studies in Eukaryotes.....	11
1.5.2 Studies in Eubacteria.....	19
1.5.3 Studies in Archaea.....	19
1.6 Mini-Chaperones/Micro-Peptides as chaperones.....	21
1.7 Why study sHSPs/Significance of sHSPs in life.....	22
1.8 Scope of this Study.....	24
2 MATERIAL AND METHODS.....	29
2.1 Materials.....	29

2.1.1	Chemicals.....	29
2.1.2	Enzymes.....	30
2.1.3	Kits.....	30
2.1.4	Competent Cells.....	30
2.2	Methods.....	31
2.2.1	Subcloning of TVNO775 gene for expression	31
2.2.2	Site-Directed Mutagenesis.....	32
2.2.3	Preparation of BL21 competent cells.....	34
2.2.4	Small scale lysate preparation and heat-treatment.....	35
2.2.5	Recombinant Protein Expression and Purification (Large scale)	36
2.2.6	Citrate Synthase enzyme aggregation assay	37
2.2.7	Citrate Synthase enzyme activity protection assay.....	38
2.2.8	Native - Polyacrylamide Gel Electrophoresis.....	38
2.2.9	Blue Native - Polyacrylamide Gel Electrophoresis	39
2.2.10	3-Dimensional molecular modeling of <i>Tpv</i> sHSP 14.3 and its bioinformatics.....	40
2.2.11	Procedure for Mini peptides	41
3	RESULTS.....	43
3.1	Subcloning.....	43
3.1.1	Amplification and purification of the <i>Tpv</i> sHSP 14.3 gene.....	44
3.1.2	Digestion of the amplified TVN0775 gene and the vector pET23a(+) and purification of the digested products	44
3.1.3	Ligation and transformation.....	45
3.1.4	Verification of the cloning in pET23a(+) expression vector	46

3.2	Transformation of the cloned gene in NovaBlue cells and making stocks...	50
3.3	Expression of <i>Tpv</i> sHSP 14.3 gene	52
3.4	Selection of residues for mutation	54
3.5	Mutagenesis	59
3.5.1	Analysis of Sequencing Results for Mutagenesis	61
3.6	Transformation of <i>E.coli</i> BL21(DE3) and <i>E.coli</i> T7-Express cells for expression of WT and mutant <i>Tpv</i> sHSP 14.3 protein.....	65
3.7	Expression of <i>Tpv</i> sHSP 14.3 N-Terminal Domain Mutants and heat treatment	70
3.7.1	NTD Proximal Part Mutations	70
3.7.2	Combined NTD Proximal and Distal Point Mutation.....	72
3.7.3	NTD Middle Part Mutations	73
3.7.4	NTD Distal Part Mutations	75
3.8	Preparation of Cell Lysate on Large Scale and Purification by HPLC.....	78
3.8.1	Sample Preparation.....	78
3.8.2	Anion Exchange Chromatography	79
3.8.3	Concentration of the eluted protein by Ultrafiltration Method	81
3.9	Citrate Synthase Enzyme Activity Assay	82
3.10	Citrate Synthase Aggregation Assay	88
3.10.1	NTD Proximal Mutants	88
3.10.2	NTD Middle Part Mutants.....	94
3.10.3	NTD Distal Part Mutants.....	98
3.11	Analysis of Oligomeric States by Native PAGE	103
3.12	Analysis of Oligomeric States by BN-PAGE	104

3.12.1	Buffer Exchange by Ultrafiltration	104
3.12.2	Result of BN-PAGE.....	105
3.13	Bioinformatics Section (3-D Molecular model Study)	107
3.13.1	Generation of Monomer and Dimer structure.....	107
3.13.2	Inter and Intra-molecular interactions in a dimer	109
3.13.3	ZipperDB and Rosetta Energy	131
3.13.4	Intrinsic Disorder Region (IDR) Analysis	141
3.13.5	Analysis of Structure and Surface Hydrophobicity of WT and Mutants by Chimera	153
3.13.6	Thermodynamic Stability of Mutants	162
3.14	<i>Tpv</i> sHSP 14.3 Mini-Peptides.....	163
4	DISCUSSION.....	169
4.1	Group 1: Decreasing hydrophobicity in the proximal NTD.....	178
4.2	Group 2: Negatively charged residues in the middle domain of NTD (E11V and E22G).....	189
4.3	Group 3: Hydrophobic residue in the middle domain of NTD (M12T).....	193
4.4	Group 4: Mutations in Distal domain of NTD (V23GF26YI27T, F26YI27T, V31G and V31IL33I)	196
4.5	Group 5 F8YV31G	200
4.6	Mini-peptides.....	203
5	CONCLUSION	205
	REFERENCES	209
	APPENDICES	
A.	Buffers and Solutions	233
B.	Preparation of Gels	237

C. Primer sequence	239
D. Figures for Chapter 3	240
E. Tables	253
CURRICULUM VITAE	291

LIST OF TABLES

TABLES

Table 2.1. Forward and reverse complementary primer for each mutant generated. Exchanged codons are underlined	33
Table 3.1 Transformation Efficiency of various ligation mixes in <i>E.coli</i> BL21(DE3) cells plated directly and with serial dilutions.....	46
Table 3.2 Positions of cut sites of various restriction enzymes in the restriction map of recombinant plasmid pET23a+TVN0775 (4077 bp)	48
Table 3.3 List of conserved residues in the NTD of <i>Thermoplasma volcanium</i> sHSP 14.3, after running MSA with most similar sHSPs from BLAST.....	57
Table 3.4 Important residues in different sHSPs and their functions	58
Table 3.5 Equivalent residues for the amino acids chosen for site specific mutagenesis in the NTD of <i>Tpv</i> sHSP 14.3.	58
Table 3.6 Transformation Efficiency of mutagenesis with various samples	60
Table 3.7 Summary of alignment results of mutation samples (present in XL1-Blue and DH5a competent cells) after re-sequencing.	64
Table 3.8 Transformation efficiency of mutants in BL21(DE3) competent cells.	66
Table 3.9 Name given to selected colonies for each mutant	68
Table 3.10 Transformation efficiency of mutants in T7 Express competent cells	69
Table 3.11 List of <i>Tpv</i> sHSP 14.3 NTD variants generated in this study.	70
Table 3.12 Comparison of Rosetta Energy, Shape Complementarity and C-Score of <i>Tpv</i> sHSP 14.3 WT and Mutants involving initial NTD residues	138
Table 3.13 Comparison of Rosetta Energy, Shape Complementarity and C-Score of <i>Tpv</i> sHSP 14.3 WT and Mutants involving central NTD residues ...	139
Table 3.14 Comparison of Rosetta Energy, Shape Complementarity and C-Score of <i>Tpv</i> sHSP 14.3 WT and Mutants involving NTD residues towards the carboxyl end central NTD residues	140

Table 3.15 Thermodynamic values by I-Mutant2.0 and MUpro of single mutants involving the proximal part of NTD, studied.	163
---	------------

LIST OF FIGURES

FIGURES

Figure 1.1. Schematic primary structure of sHSP indicating the three main components.....	2
Figure 1.2. a) Interactions leading to formation of a dimer and tetramer in αB-crystallin and b) dimer formation in Hsp21.	4
Figure 1.3. sHSP arrangement in resolved crystal structure as an oligomer. ...	6
Figure 1.4. Function of sHSP (Reinle et al., 2021).	7
Figure 1.5. MSA of N-terminal domain of the sHSPs of different organism showing the conserved residues.....	18
Figure 2.1. Flow chart for cloning.....	32
Figure 3.1 Restriction enzyme digestion of several replicates of cloning vector	43
Figure 3.2 Amplified (non-purified) gene (a) and purified gene (b), TVN0775.	44
Figure 3.3 Double digested and purified insert, TVN0775, and vector, pET23a(+) on 1% agarose gel	45
Figure 3.4 Isolated plasmids of successful clones after double digestion by restriction enzymes <i>NdeI</i> and <i>BamHI</i>.....	47
Figure 3.5 Characterization of plasmid sample 6 by double digestion with various enzymes	47
Figure 3.6 Characterization of plasmid samples 1 - 5 by single and double digestion with various enzymes.	48
Figure 3.7 Several replicas of recombinant plasmids of pET23a(+) containing the cloned TVN0775 gene.	49
Figure 3.8 MSA of the gene sequence of TVN0775 obtained from National Centre for Biotechnology Information (NCBI)	50
Figure 3.9 Circular view of Recombinant plasmid map of pET23a + TVN0775.. ..	51

Figure 3.10 Recombinant plasmids isolated from <i>E.coli</i> (NovaBlue) cells and double digested with <i>Xba</i>I and <i>Hind</i>III Restriction enzymes	51
Figure 3.11 <i>Tpv</i> sHSP 14.3 (protein of TVN0775) viewed on SDS-PAGE	52
Figure 3.12 <i>Tpv</i> sHSP 14.3 (protein of TVN0775) viewed on SDS-PAGE after heat treatment.	53
Figure 3.13 <i>Tpv</i> sHSP 14.3 (protein of TVN0775), viewed on SDS-PAGE	54
Figure 3.14 MSA of (NTD) of sHSP of <i>T.volcanium</i> with sHSPs of top (17) Archaea from BLAST.....	55
Figure 3.15 MSA of (NTD) of sHSP of <i>T.volcanium</i> with sHSPs of top (14) eubacteria species from BLAST	55
Figure 3.16 MSA of (NTD) of sHSP of <i>T.volcanium</i> with sHSPs of top (8) eukaryotes from BLAST	56
Figure 3.17 The blue labelled part is the N-terminal region of the sHSP.....	59
Figure 3.18 Recombinant plasmid pET23_tvsHSP15 isolated from <i>E.coli</i> (NovaBlue) cells.....	59
Figure 3.19 Double digestion of mutant plasmid with <i>Hind</i>III/<i>Bgl</i>III restriction enzymes	60
Figure 3.20 Double digestion of mutant plasmid with <i>Hind</i>III/<i>Bgl</i>III restriction enzymes	61
Figure 3.21 Double digestion of mutant plasmid with <i>Hind</i>III/<i>Bgl</i>III restriction enzymes	61
Figure 3.22 Chromatogram images of the mutants (a) I5T and (b) F8Y.....	62
Figure 3.23 Chromatogram images of the mutants (a) E11V and (b) M12T ..	62
Figure 3.24 Chromatogram images of the mutants (a) E22G and (b) V31G ..	62
Figure 3.25 Chromatogram images of the mutants (a) I5TF8Y and (b) V31IL33I	63
Figure 3.26 Chromatogram images of the mutants (a) F7SF8Y and (b) F26YI27T	63
Figure 3.27 Chromatogram images of the mutant F8YV31G	63
Figure 3.28 Chromatogram image of the mutant I5TF7SF8Y	63

Figure 3.29 Chromatogram image of the mutant V23GF26YI27T	64
Figure 3.30 Mutated plasmids after double digestion	65
Figure 3.31 Mutated plasmids isolated from <i>E.coli</i> BL21(DE3) cells after double digestion with <i>Bgl</i>II and <i>Hind</i>III	67
Figure 3.32 Mutated plasmids isolated from <i>E.coli</i> BL21(DE3) cells after double digestion with <i>Bgl</i>II and <i>Hind</i>III	67
Figure 3.33 Mutated plasmids from transformed successfully in <i>E.coli</i> T7-Express competent cells after double digestion with <i>Xba</i>I and <i>Hind</i>III RE	69
Figure 3.34 Schematic diagram for virtual division (into proximal, middle and distal sections) of N-terminal domain of <i>Tpv</i> sHSP 14.3.	69
Figure 3.35 Cell lysate of BL21(DE3) <i>E.coli</i> cells possessing I5T mutation plasmid viewed on SDS-PAGE	71
Figure 3.36 Cell free lysate of BL21(DE3) <i>E.coli</i> cells possessing mutant plasmid viewed on SDS-PAGE	72
Figure 3.37 Cell lysate of T7-Express <i>E.coli</i> cells possessing F8YV31G (14.308 KDa) mutation plasmid viewed on SDS-PAGE	73
Figure 3.38 Cell lysate of BL21(DE3) <i>E.coli</i> cells possessing M12T (13.304 KDa) mutation plasmid	73
Figure 3.39 Cell lysate of T7-Express <i>E.coli</i> cells possessing E11V (14.304 KDa) mutation plasmid	74
Figure 3.40 Cell lysate of T7-Express <i>E.coli</i> cells possessing E22G (14.262 KDa) mutation plasmid	75
Figure 3.41 Cell lysate of BL21(DE3) <i>E.coli</i> cells possessing F26YI27T (14.338 KDa) mutation plasmid	75
Figure 3.42 Cell lysate of BL21(DE3) <i>E.coli</i> cells possessing V23GF26YI27T (14.386 KDa) mutation plasmid	76
Figure 3.43 Cell lysate of T7-Express <i>E.coli</i> cells possessing V31G (14.292 KDa) mutation plasmid	77
Figure 3.44 Cell lysate of BL21(DE3) <i>E.coli</i> cells possessing V31IL33I (14.348 KDa) mutation plasmid	77































Figure 3.45 Cell lysate of BL21(DE3) <i>E.coli</i> cells possessing <i>Tpv</i> sHSP 14.3 WT I5T mutation plasmid, F8Y mutation plasmid, I5TF8Y mutation plasmid, F7SF8Y mutation plasmid and I5TF7SF8Y mutation plasmid.....	78
Figure 3.46 Production of multiple volumes of concentrated protein of (a) <i>Tpv</i> sHSP 14.3 WT, (b) I5T variant, (c) F8Y variant and (d) I5TF8Y variant	79
Figure 3.47 Chromatogram of (a) <i>Tpv</i> sHSP 14.3 WT, (b) I5T variant and (c) F8Y variant protein.....	80
Figure 3.48 SDS-PAGE of fractions containing (a) <i>Tpv</i> sHSP 14.3 WT (fractions), (b) <i>Tpv</i> sHSP 14.3 I5T (fractions), (c) <i>Tpv</i> sHSP 14.3 F8Y (fractions).	81
Figure 3.49 SDS-PAGE of (Lane 1) <i>Tpv</i> sHSP 14.3 WT, (Lane 2) F8Y variant and (Lane 3) I5T variant, after concentration of the eluted fractions into SEC buffer	81
Figure 3.50 Enzyme activity plot and bar chart (representing the slopes) of the percentage protection of citrate synthase enzyme activity, at 47°C, by <i>Tpv</i> sHSP 14.3 WT	82
Figure 3.51 Enzyme activity plot () positive control. () negative control. () WT 1/1700, () WT 1/850, () I5T 1/1700 and () I5T 1/850 (a) and bar chart (b).....	83
Figure 3.52 Enzyme activity plot () positive control. () negative control. () WT 1/1700, () WT 1/850, () F8Y 1/1700 and () F8Y 1/850 (a) and bar chart (b).	84
Figure 3.53 Enzyme activity plot () positive control. () negative control. () WT 1/1700, () WT 1/850, () I5TF8Y 1/1700 and () I5TF8Y 1/850 (a) and bar chart (b).....	85
Figure 3.54 Enzyme activity plot () positive control. () negative control. () WT 1/1700, () WT 1/850, () F7SF8Y 1/1700 and () F7SF8Y 1/850 (a) and bar chart (b).....	86
Figure 3.55 Enzyme activity plot () positive control. () negative control. () WT 1/1700, () WT 1/850, () I5TF7SF8Y 1/1700 and () I5TF7SF8Y 1/850 (a) and bar chart (b)	87

Figure 3.56 Light scattering plot and bar chart of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Trpv* sHSP 14.3 WT, at 45°C..... 88

Figure 3.57 Light scattering plot (—◆—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✖—) WT 1:35, (—✖—) I5T 1:7 and (—○—) I5T 1:35 (a) and bar chart (b) 89

Figure 3.58 Light scattering plot (—◆—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✖—) WT 1:35, (—✖—) F8Y 1:7 and (—○—) F8Y 1:35 (a) and bar chart (b)..... 90

Figure 3.59 Light scattering plot (—◆—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✖—) WT 1:35, (—✖—) I5TF8Y 1:7 and (—○—) I5TF8Y 1:35 (a) and bar chart (b) 91

Figure 3.60 Light scattering plot (—◆—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✖—) WT 1:35, (—✖—) F7SF8Y 1:7 and (—○—) F7SF8Y 1:35 (a) and bar chart (b) 92

Figure 3.61 Light scattering plot (—◆—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✖—) WT 1:35, (—✖—) I5TF7SF8Y 1:7 and (—○—) I5TF7SF8Y 1:35 (a) and bar chart (b) 93

Figure 3.62 Light scattering plot (—◆—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✖—) WT 1:35, (—✖—) M12T 1:7 and (—○—) M12T 1:35 (a) and bar chart (b)..... 95

Figure 3.63 Light scattering plot (—◆—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✖—) WT 1:35, (—✖—) E11V 1:7 and (—○—) E11V 1:35 (a) and bar chart (b)..... 96

Figure 3.64 Light scattering plot (—◆—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✖—) WT 1:35, (—✖—) E22G 1:7 and (—○—) E22G 1:35 (a) and bar chart (b) 97

Figure 3.65 Light scattering plot (—◆—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✖—) WT 1:35, (—✖—) V23GF26YI27T 1:7 and (—○—) V23GF26YI27T 1:35 (a) and bar chart (b) 98

Figure 3.66 Light scattering plot (—◆—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✖—) WT 1:35, (—✖—) F26YI27T 1:7 and (—○—) F26YI27T 1:35 (a) and bar chart (b) 99

Figure 3.67 Light scattering plot (—◆—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✖—) WT 1:35, (—✖—) V31G 1:7 and (—○—) V31G 1:35 (a) and bar chart (b).... 100

Figure 3.68 Light scattering plot (—♦—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✱—) WT 1:35, (—✱—) V31IL33I 1:7 and (—○—) V31IL33I 1:35 (a) and bar chart (b).....	101
Figure 3.69 Light scattering plot (—♦—) Blank, (—□—) CS control, (—▲—) WT 1:7 (—✱—) WT 1:35, (—✱—) F8YV31G 1:7 and (—○—) F8YV31G 1:35 (a) and bar chart (b).....	102
Figure 3.70 Native Gel Electrophoresis of <i>Tpv</i> HSP 14.3 WT and its I5T variant	104
Figure 3.71 SDS-PAGE image of <i>Tpv</i> sHSP 14.3 WT and its NTD mutant proteins (I5T, F8Y, I5TF8Y, F7SF8Y, I5TF7SF8Y and F8YV31G).....	105
Figure 3.72 Resolution of the oligomeric forms of the <i>Tpv</i> sHSP 14.3 WT (a) and its various NTD variants; <i>Tpv</i> sHSP 14.3 I5T (b); <i>Tpv</i> sHSP 14.3 F8Y (c) and <i>Tpv</i> sHSP 14.3 I5TF8Y (d).....	106
Figure 3.73 Resolution of the oligomeric forms of the <i>Tpv</i> sHSP 14.3 WT (a) and its various NTD variants; <i>Tpv</i> sHSP 14.3 F7SF8Y (b); <i>Tpv</i> sHSP 14.3 I5TF7SF8Y (c) and <i>Tpv</i> sHSP 14.3 F8YV31G (d).....	107
Figure 3.74 Monomer structure developed using Easy Modeller 4.0 of <i>Tpv</i> sHSP 14.3 WT	108
Figure 3.75 Top view of a dimer of <i>Tpv</i> sHSP 14.3 WT indicating the monomer-monomer interaction between the $\beta 2$ and $\beta 6$ of each monomer.....	109
Figure 3.76 Mutation points in the NTD of <i>Tpv</i> sHSP 14.3 WT shown as stick model on a ribbon structure.....	109
Figure 3.77 3-D image of dimer of <i>Tpv</i> sHSP 14.3 WT.....	110
Figure 3.78 3-D image of dimer of <i>Tpv</i> sHSP 14.3 I5T	111
Figure 3.79 3-D images of dimer of <i>Tpv</i> sHSP 14.3 WT	112
Figure 3.80 3-D images of dimer of <i>Tpv</i> sHSP 14.3 F8Y	113
Figure 3.81 3-D images of dimer of <i>Tpv</i> sHSP 14.3 WT.....	114
Figure 3.82 3-D images of dimer of <i>Tpv</i> sHSP 14.3 I5TF8Y	115
Figure 3.83 3-D images of dimer of <i>Tpv</i> sHSP 14.3 WT.....	116
Figure 3.84 3-D images of dimer of <i>Tpv</i> sHSP 14.3 F7SF8Y	117

Figure 3.85 3-D images of dimer of <i>Tpv</i> sHSP 14.3 WT.....	118
Figure 3.86 3-D images of dimer of <i>Tpv</i> sHSP 14.3 I5TF7SF8Y.....	119
Figure 3.87 3-D images of dimer of <i>Tpv</i> sHSP 14.3 WT.....	120
Figure 3.88 3-D image of dimer of <i>Tpv</i> sHSP 14.3 E11V.....	120
Figure 3.89 3-D images of dimer of <i>Tpv</i> sHSP 14.3 WT.....	121
Figure 3.90 3-D images of dimer of <i>Tpv</i> sHSP 14.3 E22G.....	121
Figure 3.91 3-D images of dimer of <i>Tpv</i> sHSP 14.3 WT.....	122
Figure 3.92 3-D image of dimer of <i>Tpv</i> sHSP 14.3 M12T.....	123
Figure 3.93 3-D images of dimer of <i>Tpv</i> sHSP 14.3 WT.....	124
Figure 3.94 3-D images of dimer of <i>Tpv</i> sHSP 14.3 F26YI27T.....	124
Figure 3.95 3-D images of dimer of <i>Tpv</i> sHSP 14.3 WT.....	125
Figure 3.96 3-D images of dimer of <i>Tpv</i> sHSP 14.3 V23GF26YI27T.....	126
Figure 3.97 3-D images of dimer of <i>Tpv</i> sHSP 14.3 WT.....	127
Figure 3.98 3-D image of dimer of <i>Tpv</i> sHSP 14.3 V31G.....	127
Figure 3.99 3-D images of dimer of <i>Tpv</i> sHSP 14.3 WT.....	128
Figure 3.100 3-D images of dimer of <i>Tpv</i> sHSP 14.3 V31IL33I.....	129
Figure 3.101 3-D images of dimer of <i>Tpv</i> sHSP 14.3 WT.....	130
Figure 3.102 3-D images of dimer of <i>Tpv</i> sHSP 14.3 F8YV31G.....	131
Figure 3.103 Graphical representation of Rosetta energy of <i>Tpv</i> sHSP 14.3 WT protein.....	132
Figure 3.104 Rosetta Energy graph of each of <i>Tpv</i> HSP 14.3 WT and mutants as analyzed by ZipperDB for (a) WT, (b) I5T, (c) F8Y, (d) I5TF8Y, (e) F7SF8Y, (f) I5TF7SF8Y, (g) E11V and (h) E22G.....	136
Figure 3.105 Rosetta Energy graph of each of <i>Tpv</i> HSP 14.3 WT and mutants as analyzed by ZipperDB for (a) WT, (b) M12T, (c) V31G, (d) F8YV31G, (e) V23GF26YI27T, (f) F26YI27T and (g) V31IL33I.....	137
Figure 3.106 PONDR output of <i>Tpv</i> HSP 14.3 protein.....	142
Figure 3.107 PONDR score plots of NTD of <i>Tpv</i> HSP 14.3 WT (a), I5T (b), F8Y (c), I5TF8Y (d), F7SF8Y (e), and I5TF7SF8Y (f).....	144

Figure 3.108 PONDR score plots of NTD of <i>Tpv</i> HSP 14.3 WT (a), E11V (b), M12T (c), and E22G (d)	145
Figure 3.109 PONDR score plots of NTD of <i>Tpv</i> HSP 14.3 WT (a), V23GF26YI27T (b), F26YI27T (c), V31G (d), V31IL33I (e), and F8YV31G (f)	147
Figure 3.110 Graphical output from PrDOS of <i>Tpv</i> sHSP 14.3 protein	148
Figure 3.111 Disorder probability score plots of variants of <i>Tpv</i> sHSP 14.3; WT (a), I5T (b), F8Y (c), I5TF8Y (d), F7SF8Y (e), I5TF7SF8Y (f)	150
Figure 3.112 Disorder probability score plots of variants of <i>Tpv</i> sHSP 14.3; E11V (a), M12T (b) and E22G (c)	151
Figure 3.113 Disorder probability score plots of variants of <i>Tpv</i> sHSP 14.3; WT (a), V23GF26YI27T (b), F26YI27T (c), V31G (d), V31IL33I and (e), F8YV31G (f)	152
Figure 3.114 Three dimensional dimer structure of <i>Tpv</i> sHSP 14.3 WT showing the surface hydrophobicity	153
Figure 3.115 Effect of I5T mutation on 3-D structure of <i>Tpv</i> sHSP 14.3	154
Figure 3.116 Effect of F8Y mutation on 3-D structure of <i>Tpv</i> sHSP 14.3	155
Figure 3.117 Effect of I5TF8Y mutation on 3-D structure of <i>Tpv</i> sHSP 14.3	155
Figure 3.118 Effect of F7SF8Y mutation on 3-D structure of <i>Tpv</i> sHSP 14.3	156
Figure 3.119 Effect of I5TF7SF8Y mutation on 3-D structure of <i>Tpv</i> sHSP 14.3	156
Figure 3.120 Effect of M12T mutation on 3-D structure of <i>Tpv</i> sHSP 14.3 ...	157
Figure 3.121 Effect of E11V mutation on 3-D structure of <i>Tpv</i> sHSP 14.3	158
Figure 3.122 Effect of E22G mutation on 3-D structure of <i>Tpv</i> sHSP 14.3 ...	158
Figure 3.123 Effect of F26YI27T mutation on 3-D structure of <i>Tpv</i> sHSP 14.3	159
Figure 3.124 Effect of V23GF26YI27T mutation on 3-D structure of <i>Tpv</i> sHSP 14.3	159
Figure 3.125 Effect of V31G mutation on 3-D structure of <i>Tpv</i> sHSP 14.3 ...	160

Figure 3.126 Effect of V31IL33I mutation on 3-D structure of <i>Tpv</i> sHSP 14.3.	160
.....	
Figure 3.127 Effect of F8YV31G mutation on 3-D structure of <i>Tpv</i> sHSP 14.3.	161
.....	
Figure 3.128 Light scattering plot (—◆—) Blank, (—□—) Control, (—▲—) Peptide1 and (—✦—) Peptide2 and bar chart (representing the slopes) of the percentage protection of γADH against heat induced aggregation in the presence of 10 μg/ml of peptide1 and peptide2	165
.....	
Figure 3.129 Light scattering plot (—◆—) Blank, (—□—) Control, (—▲—) Peptide1 and (—✦—) Peptide2 and bar chart (representing the slopes) of the percentage protection of γADH against heat induced aggregation in the presence of 20 μg/ml of peptide1 and peptide2	165
.....	
Figure 3.130 Light scattering plot (—◆—) Blank, (—□—) Control, (—▲—) Peptide1 and (—✦—) Peptide2 and bar chart (representing the slopes) of the percentage protection of γADH against heat induced aggregation in the presence of 30 μg/ml of peptide1 and peptide2	166
.....	
Figure 3.131 Light scattering plot (—◆—) Blank, (—□—) Control, (—▲—) Peptide1 and (—✦—) Peptide2 and bar chart (representing the slopes) of the percentage protection of γADH against heat induced aggregation in the presence of 40 μg/ml of peptide1 and peptide2	166
.....	
Figure 3.132 Light scattering plot (—◆—) Blank, (—□—) Control, (—▲—) Peptide1 and (—✦—) Peptide2 and bar chart (representing the slopes) of the percentage protection of γADH against heat induced aggregation in the presence of 50 μg/ml of peptide1 and peptide2	167
.....	
Figure 3.133 Light scattering plot (—◆—) Blank, (—□—) Control, (—▲—), Peptide1 (55 μg/ml) (—✦—) Peptide1 (60 μg/ml), (—➤—) Peptide 2 (65 μg/ml) and bar chart	167
.....	
Figure 4.1 Side view of the dimer of WT (a) and I5T mutant variant (b). Black arrows point towards NTD, yellow arrows towards CTD.	180

Figure 4.2 Enzyme activity plot (a) and bar chart (b) of the percentage protection of citrate synthase.....	187
Figure 4.3 Enzyme activity plot (a) and bar chart (b) of the percentage protection of citrate synthase.....	188
Figure 4.4 3-D structure of WT (a) and F26YI27T (b). Black arrow showing widening of the gap between the helix.....	199

LIST OF ABBREVIATIONS

ABBREVIATIONS

sHSP	Small Heat Shock Protein
<i>Tpv</i>	<i>Thermoplasma volcanium</i>
NTD	Amino terminal domain
ACD	Alpha crystallin domain
CTD	Carboxyl terminal domain
CS	Citrate Synthase
IDR	Intrinsic Disordered Region
ADH	Alcohol Dehydrogenase
3-D	Three Dimensional
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
BN-PAGE	Blue Native polyacrylamide gel electrophoresis
KDa	Kilo Dalton
AGE	Agarose Gel Electrophoresis
O/N	Overnight

CHAPTER 1

INTRODUCTION

Small heat shock proteins have a unique mechanism of activation. Unlike other chaperones which get activated by some form of ATP-dependent mechanisms, the sHSPs are activated by stress related processes. This in turn results in conformational changes in the sHSPs, leading to exposure of the hydrophobic surfaces/hydrophobicity rich areas which are by default hidden/buried in the higher order complex multimers in a no stressful environment. Thus exposure to stress leads to conversion of sHSPs from a low affinity binding state to a high affinity binding state. In the high affinity binding state, the exposed hydrophobic residues interact and perhaps bind to the exposed hydrophobic residues of a partially folded or denatured protein, which are thus the substrate of sHSPs (Haslbeck *et al.*, 2005; Haslbeck, Weinkauff and Buchner, 2019). It was initially found that stress, for example heat stress, can lead to changes in the conformation of the sHSPs. Later, the fact was established, through various experiments that sHSPs protect other cellular proteins against heat or chemical induced degradations (Maiti, Kono and Chakrabarti, 1988; Horwitz, 1992).

Despite low sequence homology, sHSPs present in different domains of life have structural similarities. The monomeric mass of sHSPs varies between 12 KDa to 43 KDa (Haslbeck *et al.*, 2005). Besides a few, almost all the known sHSPs' primary structure can be distinctively divided into three regions. The most extensively studied element is the alpha crystallin domain (ACD), which is located closer to the carboxyl end of the protein, comprising of an average of 90 amino acids and is also the trade mark of sHSPs (Basha, O'Neill and Vierling, 2012). Of the three elements, the ACD is the most conserved, in terms of sequence, structure as well as function (Poulain, Gelly and Flatters, 2010; Bondino, Valle and ten Have, 2012; Klevit, 2020). This region accounts for the largest part of the sHSPs, that is, approximately 58% of the

primary sequence. The genomic records show that the primary sequence forming the ACD has an abundance of charged amino acids and a scarcity of aromatic amino acids (Poulain, Gelly and Flatters, 2010; Alderson *et al.*, 2020). The ACD is flanked by amino terminal domain (NTD) and carboxyl terminal domain (CTD) on its left and right, respectively. The NTD is rich in hydrophobic amino acids and is, therefore, not only packed within the quaternary assembly of a non-activated sHSP, but also critical for substrate recognition and binding, once exposed to the environment, after the sHSP is activated. The length of the NTD varies from specie to specie and also within a specie, for example humans. This variation in length and sequence is thought to dictate the needs of that specific specie possessing it. On the other hand, CTD is mainly composed of hydrophilic residues, but also possess a highly conserved three aminoacid long sequence I/VXI/V motif, (Figure 1.1). This motif is found to interact with the ACD groove in the process of oligomer formation (Treweek *et al.*, 2015; Carra *et al.*, 2017; Haslbeck, Weinkauff and Buchner, 2019) (Hilton *et al.*, 2012).

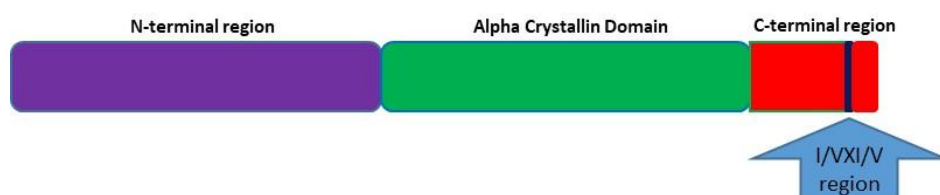


Figure 1.1. Schematic primary structure of sHSP indicating the three main components, NTD in Violet color, ACD in green color and CTD in red. The blue section corresponds to highly conserved IXI region.

The well resolved structure, by X-ray diffraction and NMR studies, is that of highly organized and, therefore, easily crystallized and studied structure of the middle portion of sHSPs, which is the ACD. It forms β -sheets that are arranged into a sandwich-like structure. The topology of ACD is to a great extent conserved during evolution while the N and C terminal structure and sequence vary. All metazoans including humans have seven and invertebrates/non-metazoans have eight β -sheets in the ACD due to amino acid arrangements in metazoans sHSP at β_6 and β_7 sheet making it to appear as one single beta strand called β_{6+7} (Treweek *et al.*, 2015). The NTD and CTD occasionally possess a well-defined secondary structure. Solution-

phase NMR studies reveal the CTD to be a flexible coil which helps in its function of interaction with the ACD grooves to stabilize the quaternary structure. The NTD consists of a mixture of helix and coils or disordered regions, which are generally in the amino terminal of it (Boelens, 2020).

Studies have shown that sHSPs are rarely found as single separate entities within the cell. Instead, they show various degrees of multimeric organizations. In order to achieve this, the foremost step is the formation of dimers. There are two distinct modes of dimerization formations in sHSPs. In metazoans, *i.e.*, humans and other eukaryotes, the β_6+7 sheet of each of the monomer arranges in an antiparallel mode to each other (Haslbeck, Weinkauff and Buchner, 2019). This positioning helps in the formation of salt bridges between them, which are the base on which the whole mulimerization structure hinges on (Wen *et al.*, 2010; Santhanagopalan *et al.*, 2018). In non-metazoans, the dimer is formed by virtue of interactions between the β_6 strand of the ACD of one monomer with β_2/β_3 strand of the ACD of the other (Treweek *et al.*, 2015). Figure 1.2a shows a dimer and a tetramer formation by virtue of salt bridge, *i.e.*, interaction between arginine and aspartic acid of β_6+7 sheet in ACD of two α B-crystallin. In Figure 1.2b, the dimerization mode of non-metazoans is seen where the interaction is between $\beta_2 - \beta_6$ of adjacent monomers. Although β_6 sheet is absent in Hsp20.1 of *Sulfolobus solfataricus*, still it forms stable dimers where the charged and polar residues of the ACD, interact by virtue of salt bridges, to form the dimer interface. This suggests that β_6 might not be only responsible, but may provide extra support, for dimerization in non-metazoans (Liu, J. Chen, *et al.*, 2015).

Dimerization is followed by the dimers' interaction into building of an oligomer. These oligomers, by virtue of interaction of IXI motif with a groove like structure at the C-terminus are formed by the β_4 and β_8 strands. sHSP oligomers may be homo-oligomer (composed of single type of subunit), which is the general trend found in all sHSPs besides human, or hetero-oligomer (composed of different subunits), largely established in human sHSPs (Delbecq *et al.*, 2013).

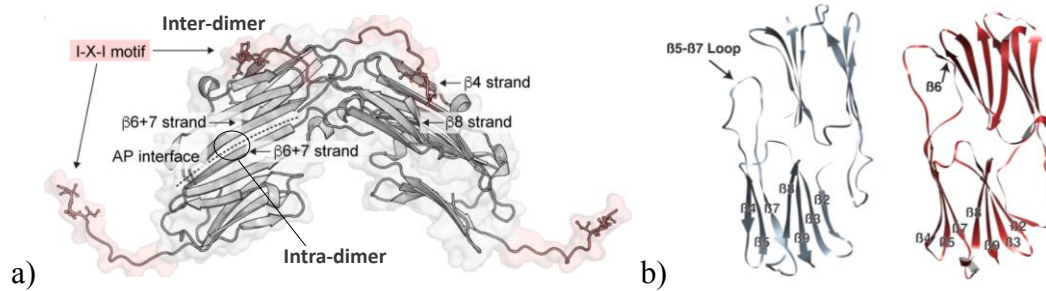


Figure 1.2. a) Interactions leading to formation of a dimer and tetramer in α B-crystallin (Collier and Benesch, 2020) and b) dimer formation in Hsp21 of *Arabidopsis thaliana* and TaHsp16.9 from *Triticum aestivum* (Yu *et al.*, 2021).

The oligomerization of sHSPs is a complex mechanism and varies in orthologues. Some sHSPs exhibit a distinct oligomer structure with a specified and fixed number of monomer subunits making it up. For example, the sHSP of an archaea *Methanococcus jannaschii*, MjHsp16.5 is a ball-like structure, (Figure 1.3), hollow in the centre and is composed of 24 subunits (Kim *et al.*, 1998). On the other hand, there are merely 12 subunits/monomers assembled in the oligomer of wheat TaHsp16.9, organized in the form of two hexameric discs, (Figure 1.3), having a hole of 25 Å in the centre, despite the fact that the mode of dimerization is similar in both the sHSPs. Although the ACD is structurally very similar in the two sHSPs, the sequence identity (of both ACDs) is only 23% (Van Montfort *et al.*, 2001). Although the NTD is not well resolved in the crystal structures of many sHSPs, six of the twelve NTD arms in the dodecamer of TaHsp16.9 show the propensity to form helix. In the crystal structure of TaHsp16.9 NTD is shown to form contacts (among top and bottom hexameric discs) for the stabilization of the oligomers.

A set of studies have shown that formation of oligomers is by virtue of the interaction of terminal regions of sHSPs in a dimer with neighboring dimers. It has been established that the highly conserved IXI motif in the CTD of a dimer interacts with the $\beta 4 - \beta 8$ groove of another dimer, while the NTD of a dimer interacts with the groove formed by virtue of monomer – monomer interaction, also called the dimer interface (Reinle, Mogk and Bukau, 2021). Another sHSP which is reported to be present as a dodecamer, which is stabilized by interactions of CTD, NTD and

dimerization loop is M1 (PDB: 5ZS3) from *Mycobacterium marinum* M and Acr1 from *M.tuberculosis*. (Kennaway *et al.*, 2005; Bhandari *et al.*, 2019). However, not always the NTD is essential in oligomerization. For example, similar to its WT, the NTD truncated MjHsp16.5, is present as a 24-mer oligomer, but, with a low density. Moreover, a recently resolved structure by Cryo-microscopy, the dodecamer of Hsp21 from *A.thaliana* also shows that its NTD accumulates in the central cavity forming interactions and providing strength to the oligomer. However, the foundation of the oligomer, the dimer, is formed by interaction of $\beta 2$ and $\beta 5 - \beta 7$ coil of two monomers (since $\beta 6$ is absent in this sHSP). The IXI motif in the CTD of a dimer then exhibits interactions with adjacent dimer $\beta 4 - \beta 8$ groove which is occupied by the substrate when the sHSP is activated and form smaller species (Yu *et al.*, 2021). Another well studied oligomer is the crystal structure of *Xanthomonas citri* sHSP, XaHspA. By size exclusion chromatography, it was found that this sHSP exists as 36-mer. The mode of dimerization, by virtue of charged and polar interactions, is similar to non-metazoans ($\beta 2 - \beta 6$ interaction). Three dimers than form a hexamer and the 36-mer is formed by either of the two configurations of these hexamer discs; 6-6-12-6-6 or 12-6-6-12 where the numerals indicate the number of monomers in the disc (Hilario *et al.*, 2011).

HSPBs, in humans, e.g., HSPB1 – HSPB10, are known mammalian sHSPs, that occur in nature as polydisperse entities and show phenomenon of heterooligomerization by means of N-terminal arm (Shashidharamurthy *et al.*, 2005), in opposition to archaeal sHSPs, that occur as monodisperse particles with a specified, constant number of subunit in each oligomer when present in inactive form (Mishra *et al.*, 2018). Studies involving Hsp27, (or human HSPB1), have shown that, when inactive, it occurs in equilibrium between two oligomeric states, 530 KDa and 120 KDa species. However, the activation mimicking mutant, Hsp27D3, shows presence of only smaller sized oligomers in SEC profile and 4-fold increase in denatured substrate binding than WT, suggesting that phosphorylation/activation leads to dissociation of oligomers and subunit exchange (Shashidharamurthy *et al.*, 2005).

Apart from the usual oligomer formation, there are a few exceptions of sHSPs which do not obey the customary principle of quaternary assembly. A crystal structure of a sHSP from a parasitic worm, Tsp36 (PDB: 2BOL) is an excellent example to it. The sHSP exists as dimers in reducing environments and forms merely tetramers when exposed to oxidizing surroundings. This unique metazoan sHSP owns two distinct features, which are typical to other sHSPs of flatworms that are parasites. Firstly, it lacks CTD, an important structure known to stabilize the oligomers, and secondly, it possesses two ACDs, a rare but possible phenomenon in sHSP world. In this sHSP, the dimer and tetramer formation is by virtue of interaction between NTD and ACD of adjacent monomers. The NTD structure is well resolved in this sHSP, and is arranged into a succession of helices (Kappé *et al.*, 2004; Stamler *et al.*, 2005).

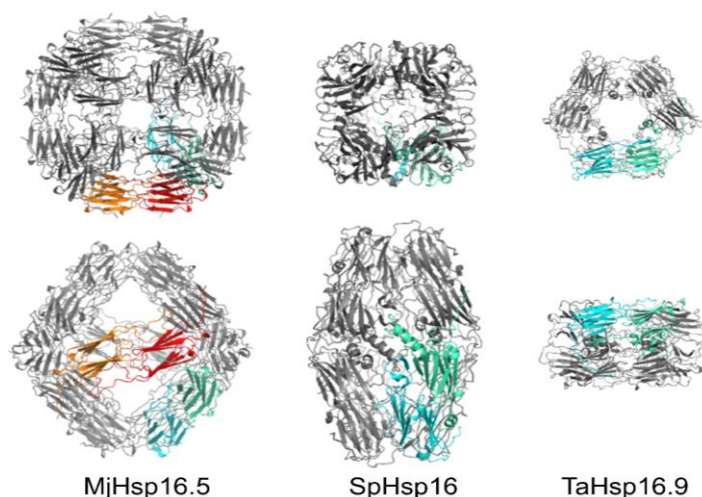


Figure 1.3. sHSP arrangement in resolved crystal structure as an oligomer in an archaea *Methanococcus jannaschii* MjHsp16.5 (a 24-mer), a yeast *Schizosaccharomyces pombe* SpHsp16.0 (a 16-mer) and wheat, *Triticum aestivum* TaHsp16.9 (a 12-mer) (Haslbeck & Vierling, 2015).

Not all NTR regions contribute to the core density of the oligomer. In the model structure generated by online Tools (MODELLER) and cryo-EM difference maps of Hsp21 from *Arabidopsis Thaliana* (using Hsp16.9 as a template), a chloroplast localized sHSP crucial for plant stress resistance, it was noticed that in addition to the major positive densities in the inside of the dodecamer, there were minor positive densities towards the outside of the oligomer which reflected the location of relatively more flexible NTD arms of at least six of the monomers. In addition,

limited proteolysis of the Hsp21 resulted in immediate degradation of the NTR without disturbing the dodecameric assembly, suggesting that NTR is not primarily responsible for maintaining the oligomerization (Rutsdottir *et al.*, 2017).

1.1 Function of sHSPs

In a single statement, the sHSPs protect the partially folded or misfolded proteins in the cell preventing them from making insoluble aggregates. This process, which is not completely understood, contains a whole set of steps where the sHSPs are converted to active form, followed by recognition and interaction with the substrate, to form a stabilized sHSP-substrate complex and then presenting this substrate to Heat Shock Proteins for refolding or for digestion and releasing itself to be recycled over again. This whole process is presented by cartoon graphics in the Figure 1.4.

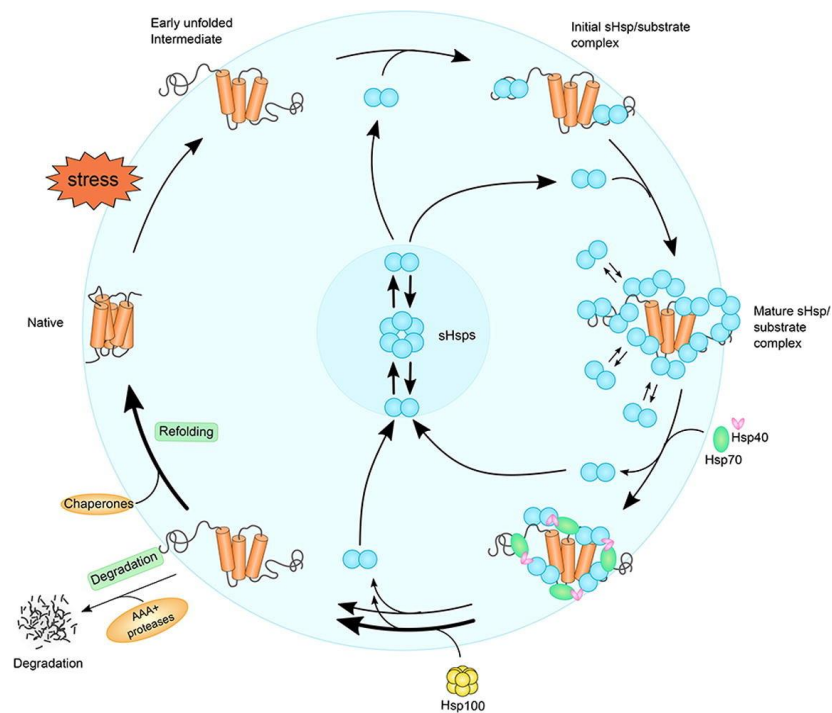


Figure 1.4. Function of sHSP (Reinle et al., 2021).

1.2 Activation of sHSPs

It is known that sHSPs occur in nature in at least two states, one a comparatively dormant and inactive form, the oligomers, and another in which the sHSPs are highly active and sensitive towards their substrate, the partially folded protein. In order to be prompt in action and minimize the disastrous effects of insoluble protein aggregates on the cell, the sHSPs are expected to be highly sensitive to any kind of stress and therefore, perceive smallest changes in the environment in terms of any kind of stress, *i.e.*, temperature, oxidative and/or pH. For heat activation of sHSPs, temperature of activation varies from specie to specie, where sHSPs from hyperthermophiles require comparatively higher temperature for activation (MjHsp16.5 activation temperature = 60°C), while mammalian HSPBs are expected to show action in response to slight changes in physiological temperatures (Janowska *et al.*, 2019). Structural changes in α -crystallins when exposed to higher temperatures was observed long before the protective nature of these proteins was established (Spector and Zorn, 1967; Horwitz, 1992).

Another notable stress in the intra and extra cellular environment is the pH, which gets altered in several diseases like cancer, Alzheimer's and Parkinson's where the pH of the affected tissue decreases. Similar pH changes occur when a tissue is exposed to ischemic circumstances. Studies show activation of sHSPs in ischemia tissue. This most likely occurs by virtue of highly conserved histidine residues (their side chain has a pKa of 6.4 – 7.5) present in abundance in the sHSPs. Many studies show that pH when decreased below the physiological working pH, results in expansion of sHSPs oligomers or decrease in the size of oligomers, and this can be due to formation of smaller oligomers (Janowska *et al.*, 2019).

Reactive oxygen species (ROS) formed as a result of normal and pathological processes in the cells, if not properly neutralized, leads to oxidative stress. The exact process of sHSP in neutralizing ROS is not clearly known, but it is proposed that the levels of glutathione are modulated in the cells by sHSPs, like HSPB1 and HSPB5 (Janowska *et al.*, 2019).

Besides these external environmental stresses and internal oxidative stress, the activation and associated changes in the structure are also regulated by phosphorylation, at specific sites in the NTD of the sHSP in some species. Phosphomimicking variants, *in vitro*, have been shown to have reduced oligomer size, preferably dimers, as compared to their WT counterparts in HSPBs and phosphorylation has shown to increase the chaperone activity (Jovcevski *et al.*, 2015) with a few exceptions where increasing phosphorylation resulted in a decrease in chaperone activity (Morrison *et al.*, 2003; Janowska *et al.*, 2019).

These stresses lead to subunit exchange which has been previously seen to contribute to effective chaperone activity (Bova *et al.*, 1997).

1.3 Substrate – sHSP sites and Interaction

It has been proposed that hydrophobic sites, hidden away from the environment in an oligomeric state and exposed during the breakdown of oligomer are the substrate binding sites, for example, the NTD, the CTD and $\beta 4 - \beta 8$ groove (Van Montfort *et al.*, 2001). However, recently it is shown that different regions of the same sHSP has varied affinity towards variable substrates, for example high affinity of ACD is seen towards amyloid protein and NTD towards lysozyme of the HSPB5 (Carra *et al.*, 2019). However, the sequence and size variability of NTD of sHSPs suggests it to be the main weapon of sHSPs for recognizing of substrates in different species (Jaya, Garcia and Vierling, 2009; Mogk, Ruger-Herreros and Bukau, 2019), and swapping NTDs among sHSPs have shown to transfer the protective ability specific to a sHSP (Specht *et al.*, 2011; Muranova *et al.*, 2021).

It has been seen that unfolded α -lactalbumin binds to XaHspA in a 1:2 substrate:chaperone ratio, suggesting that dimer of sHSP are active in binding to the substrate and further modify to keep them in soluble form (Hilario *et al.*, 2011). However, in Hsp20.2 of *D.radiodurans*, six substrate molecules bind to an 18-mer oligomer (Bepperling *et al.*, 2012). Thus it is safely said that the size and nature of

substrate-sHSP interaction complex varies with the nature and concentration of substrate, sHSP and the conditions of experiment.

1.4 sHSP mechanism of action

sHSPs are prompt in action and they tend to bind the misfolded protein in the early stage of its unfolding. This action helps in halting the further unfolding of the protein, thus keeping it in a near-native and minimally unfolded confirmation (Stengel *et al.*, 2012; Ungelenk *et al.*, 2016). This sequestration helps in preserving the confiscated substrate molecules separate from each other and also assists in easy release and refolding of the substrate post-release, where the release of the substrate is made possible in an energy-dependent manner involving high molecular weight, Hsp70 and Hsp100, chaperones (Mogk *et al.*, 2003). Thus, sHSPs are acting as a reservoir or a sponge for the absorption-like activity of the otherwise deleterious partially folded proteins. (Mogk, Ruger-Herreros and Bukau, 2019). However, there are exceptions when this substrate – sHSP association is irreversible (Haslbeck and Vierling, 2015).

1.5 Structural and functional role of NTD

Initially, either the whole N-terminal domain was truncated or a portion of it was deleted and later point mutations were performed to identify the role of different regions of NTD in forming/stabilizing a sHSP structure as well as in its chaperone function. Absolute absence of NTD in various sHSPs resulted not only in complete loss of the function of sHSP but also an alteration in the arrangement of sHSPs units in forming the quaternary structure, as seen in study with HSPB6 of human, Hsp26 of yeast and HSP16.9 of rice (Young *et al.*, 1999; Haslbeck *et al.*, 2004; Stromer *et al.*, 2004). On the other hand, partial truncation of NTD in *E.coli* IbpB, HspH and HspF of *B.japonicum* and HSPB6 resulted in decreased chaperone activity and small oligomeric structures of mutants (Studer *et al.*, 2002; Heirbaut *et al.*, 2014).

1.5.1 Studies in Eukaryotes

The initial 15 residues of sHSPs have shown to play an important role in determining and maintaining the oligomer size of the sHSPs and consequently its chaperone function. Deletion of the initial 15 aminoacids of sHSP in *C.elegans*, led to decrease in its oligomer size. However, further reducing the NTD of the sHSP, that is, deleting the initial 30 residues, did not affect the oligomer size but complete removal of NTD resulted in the formation of smaller oligomeric species, especially dimers, in the mutant. Interestingly, no subunit exchange was observed in all of the above truncated mutants and all of these truncated mutants failed to protect the substrate against aggregation, indicating the role of NTD in subunit exchange and therefore in substrate recognition and substrate binding (Leroux *et al.*, 1997).

Complete NTD removal led to diminish chaperone activity, also in human-HSPB6. However, in order to identify the imperative regions or residues in this NTD, that are responsible for substrate recognition and/or binding or involved in oligomerization, shorter NTD deletion mutants were prepared and compared to its WT using various substrate enzymes. All the NTD truncated mutants displayed similar oligomeric structure as the WT, proving little role of NTD in oligomerization of this sHSP. In addition to this, the chaperone activity of all the NTD truncated mutants were similar to WT, except one mutant, in which a highly conserved region in the middle (residues 31 – 40) of the NTD was truncated. This mutant showed increase protective effect than the WT. Further study involving point mutation helped to indicate that glutamine at position 31 regulates the chaperone activity of HSPB6 while phenylalanine at position 33 is involved in affecting its oligomerization (Heirbaut *et al.*, 2014).

NTD of a sHSP does not always participate in oligomerization. (HSPB7). The expression of HSPB7, with intact NTD, in human embryonic kidney cell lines (HEK293) and analysis via sucrose gradient assay indicates that it is present as monomers and dimers, in the cells (Vos *et al.*, 2010). Numerous experimental studies have shown that the intrinsically disordered NTD of the HSPB7 is responsible for

binding to proteins with the polyQ stretches and thus playing a role in prevention of their aggregation. This NTD when fused with HSPB1, inferred this property to HSPB1 itself, which inherently lacks binding to and is non-functional towards proteins possessing polyQ regions, even in active/phosphorylated state. Also, the HSPB1 possessing fused NTD of HSPB7 was found to present in only dimers, although HSPB1 innately exist as polydisperse molecules varying from dimers to oligomers with even 50-mers (Wu *et al.*, 2019). On the other hand, over expression of recombinant HSPB7 when analyzed by SEC, showed presence of two species, a dimer and a larger oligomer with 600 KDa MW. However, it is worthy of noticing that this NTD in recombinant HSPB7 when lost its initial 13 residues, or the serine rich region (17 – 29) resulted in formation of dimers in opposition to larger oligomers as found in recombinant HSPB7 WT. Interestingly, this NTD-truncated HSPB7, existing as dimer, displayed a loss of its activity. The above studies indicate that the proximal NTD as well as the serine rich region of the NTD of HSPB7 from human, participates to the interactions to form larger assemblies (Muranova *et al.*, 2021).

Although NTD of sHSPs, throughout the three domains, present a large variability in terms of both the sequence and length, some regions of it are highly conserved in closely related or even distantly related orthologues. Residues in conserved motifs are hypothesized to play an important role either in protein structure or its function or both. Point mutation studies in such regions are performed to understand and underscore the residues vital for a specific role. One of such regions is the conserved motif RLFdqF_xFG, in HSPBs. Deletion of this whole motif resulted in decrease in oligomer size and increase in chaperone activity and stability towards denaturation by urea (Pasta *et al.*, 2003). Replacing the arginine in this motif with alanine (R29A mutation), where the hydrophobicity of the region is increased, resulted in drastic decrease in heat stability of HSPB8 and a slight increase in heat stability of HSPB1 while heat stability of other HSPBs was not affected. Although replacing this arginine did not affect the chaperone activity of any of the HSPBs, it altered their oligomeric assembly, as compared to the respective WTs. In HSPB1, it caused a concentration dependent shift towards higher oligomeric species (similar to that

found in WT), while at lower concentrations the major portion was as smaller species. For HSPB5, the mutant shifts towards formation of slightly low MW oligomers as compare to WT (540 KDa → 460 KDa). While in HSPB8, this mutation led to the formation of tetramers, opposed to dimers in WT. Although this mutation did not affect the oligomeric nature of HSPB6, it is assumed that this highly conserved arginine execute an imperative role in the structural assembly of vertebrae HSPBs. Where HSPB1, HSPB4, HSPB5 and HSPB6 and their respective mutants were able to protect the substrate enzymes, ovotransferrin and myosin subfragment-1, HSPB8 WT and its mutant both were unable to protect these substrates. (Shatov *et al.*, 2018). However, using other substrates, insulin and yeast-ADH, it was found that HSPB8 WT is able to protect stress induced aggregation in former. However, initially the aggregation of the latter was protected but then increased more, as compared to the control. On the contrary, R29A mutant was unable to protect insulin against aggregation and only protects yeast-ADH aggregation in the early stages. This result indicates two striking points, one, not all sHSPs protect all substrates and second, highly conserved arginine in this motif plays a role in structure and function of HSPBs. Deletion of the initial part of this conserved motif (SRLLD) in HSPB8, also results in loss of chaperone activity, indicating charge and hydrophobicity richness in this region necessary for sHSP function (Shatov, Sluchanko and Gusev, 2021). Replacing of the conserved R29 or deletion of the initial penta-peptide of the conserved motif, also alters hetero-oligomerization in HSPBs, further emphasizing the role of NTD in oligomerization (Shatov, Strelkov and Gusev, 2020). This arginine when replaced with glutamine in HSPB4 results in increase in surface hydrophobicity and ability to bind membrane proteins which is assumed to be the reason of decrease in lens transparency and cataract formation (Phadte, Santhoshkumar and Sharma, 2018).

A motif (SNVFD) similar to RLFQDFFG conserved motif in human HSPBs is also seen in a distant eukaryotic neighbour, wheat sHSP, TaHsp16.9 (Kim, Kim and Kim, 1998). The x-ray structure of the TaHsp16.9 shows that V4 and F10 (of the conserved motif) in N-terminal arm are liable for intermolecular and intramolecular

hydrophobic interactions with hydrophobic residues (W48 and F110) of its ACD, in dormant state and are exposed when oligomers dissociate to smaller species. In TaHsp16.9 oligomer, each monomer is first bonded in a dimer, followed by interaction of three dimers to form a hexameric disk. Two of the hexameric disks then come close to form the dodecamer oligomer, which is characteristic of this particular sHSP. This oligomeric assembly is stabilized at different interfaces. First is the dimer interface, which involves stabilization by the interacting ACDs. Next, dimers form tetramers stabilized by C-terminal extensions interacting with dimerization loop. The oligomer is further stabilized by N-terminal arms interactions from the top and bottom hexameric disk forming a domain on the inside of the dodecamer. Thus the NTD not only helps in creating compact oligomer but also endows rigidity to it. Detailed zooming structural studies reveal that the F10 (in NTD) interacts with W48 and F110 of partnering monomer of the same dimer, while V4 is involved in interactions with W48 and F110 of a monomer of another dimer, where the hydrogen bonds between the nitrogen of V4 and oxygen of R109 and R111, further stabilize this interaction. When compared to an archaeal sHSP MjHsp16.5, the building block of oligomer, that is, the dimer is similar in both the sHSPs but the final oligomer differs completely, due to short dimerization loop, in TaHsp16.9. Moreover, the TaHsp16.9 lacks the β 1 strand in ACD which the archaeal sHSP possess. This β 1 strand packs alongside β 7 and contributes to the hydrophobicity of the core. Thus it is proposed that the NTD of monomers, are forming a core in the oligomer and the accessibility to these arms increases as a result of subunit exchange when sHSPs disassemble, essentially, at elevated temperatures (Van Montfort *et al.*, 2001). Thus, NTD is, generally, responsible for formation of oligomers through intermolecular interactions, between hydrophobic and charged amino acids of neighboring dimers. This belief has been further strengthened by a recent study about the crystal structure of a heterodimer of HSPBs (HSPB2 and HSPB3). The tetramer arrangement of these HSPBs gives rise to groove like structures as a result of ACD dimer interaction which are occupied by NTD in an unfolded confirmation. This may provide plasticity to the oligomeric structure of

sHSPs, which is resilient enough to maintain its integrity but loose enough to allow for subunit exchange (Clark *et al.*, 2018).

The proposed mechanism of action for sHSPs suggest the interaction of hydrophobicity rich NTD with the exposed hydrophobic region of the partially denatured client proteins. It is believed that increase in hydrophobicity will increase the chaperone activity of the sHSPs; however, in some cases, this may negatively effect the cells in-vivo system such that S10F and P20L mutations of HSPB6 cause cardiomyopathy even though those mutations increased protection effects against stress induced aggregation of model substrates (GAPDH and yeast-ADH) in in-vitro system (Shatov and Gusev, 2020).

In Chinese hamster sHSP, the truncated mutants of HSPB1, which lacked the highly conserved motif WD/EPF between eukaryotes, resulted in complete loss of chaperone activity of HSPB1. Also the quaternary structure of these mutants (by ultracentrifugation and SEC), was found to be smaller oligomeric species, tetramers and dimers, similar to pseudo phosphorylated mutants. Therefore, in this sHSP, residues between 15 and 22 were suggested to be important for oligomer stabilization and chaperone activity (Thériault *et al.*, 2004).

The main purpose of sHSP is not just to interact with a partially unfolded substrate. It includes a complex mechanism where the whole sHSP recognizes its substrate followed by interaction with it and to keep it in a stabilized/non-aggregated form, along with maintaining its own structural integrity, to limit the potential cytotoxic effects of these aggregates until mechanism of refolding or lysosomes get activated for removal of such aggregates/aggresomes. Recent experiments have shown that ACDs of HSPB1 and HSPB5 when together with their respective NTDs were proved effective in not only delaying the amyloid fibril formation but also in disaggregating the apoC-II preformed fibrils (Selig *et al.*, 2020). Similar results have been previously obtained by experiments revealing that HSPB1 mutants with NTD phosphorylation mimicking ability results in higher protection of tau protein against aggregation as compared to the respective NTD truncated mutant, while solo ACD,

of HSPB1, was incapable of exhibiting any remarkable protection, at all, to the intrinsically disordered, microtubule-associated protein (Freilich *et al.*, 2018).

Point mutation studies have highlighted residues, in sHSPs, important for substrate protection. In human α B-crystallin, HSPB5, M68 (a residue in the beginning of ACD), when substituted with more or less hydrophobic residues, resulted in loss of chaperone activity. Increase in hydrophobicity by valine or isoleucine substitution at M68 increased the heat protecting effect of the sHSP as compared to WT, while this effect was significantly decreased by threonine, a more hydrophilic residue replacement, when insulin was used as substrate protein at 37°C. Moreover, the heat stability of M68T variant was significantly reduced, while M68V and M68I showed even higher stability than WT at 62°C for 15 minutes (Shroff *et al.*, 2001).

In addition to methionine, presence of another hydrophobic residue, phenylalanine, in the NTD of sHSPs, is of pronounced importance. It has been observed that introduction of positive charge (arginine) in place of highly conserved F27 in HSPB5 of mouse, (Figure 1.5), led to complete loss of chaperone activity of the sHSP at elevated temperature. However, same mutation for F24 resulted in loss of chaperone activity even at optimum temperature of the sHSP, *i.e.*, 37°C. Moreover, reducing hydrophobicity at position 27 (F27A) also did not prove sufficient to protect chemically and thermally denatured protein substrates. All mutants had similar monomer/tertiary structure as compared to their WT (Plater, Goode and Crabbe, 1996; Horwitz *et al.*, 1998). Similarly, decrease in chaperone activity, as well as decreased oligomer size, was observed when the hydrophobicity was reduced (with alanine replacement) at F6, F7 and F9 in SpHsp16.0 from *Schizosaccharomyces pombe* (Hanazono *et al.*, 2013a). In addition, this study suggested that in an oligomer, NTDs are present, on the inside of the oligomeric core. The phenylalanine residues in the NTD interact through aromatic and cation- π interactions with phenylalanine and arginine residues of the ACD, resulting in Arg/Phe clusters in the core of the oligomer. All these results imply a role for phenylalanine in interaction with the substrate, and also in the thermal stability of sHSP (α B-crystallin). In addition to this, the phenylalanine residues in the beginning of the NTR might play a role in

stabilizing oligomers by specific interactions as suggested for F10 in TaHsp16.9 from wheat (Van Montfort *et al.*, 2001).

Although many studies have shown that phenylalanine in NTD is a noticeable substrate interaction point which is lost after any mutation in the residue. However, not all phenylalanine/hydrophobic amino acids in NTD are responsible for binding to the same substrate (different hydrophobic residues in NTD bind to different substrates) for protection against stress-related aggregation. Also not all substrates are protected against thermal aggregation in the same sHSP: substrate ratio. One such study involved replacement of all phenylalanine in NTD in PsHSP18.1 from pea (*Pisum sativum*) with Bpa (benzoyl-phenylalanine an analog of phenylalanine). The resultant variants F41 and F48 had decreased protection effect against Luciferase as compared to WT, where WT required 4:1 molar ratio to fully protect thermally aggregated luciferase, while the former mentioned mutants required 3 times higher sHSP than the WT, to show the same effect. On the other hand, MDH was observed to be fully protected at 1:1 molar ratio by all the variants (namely Bpa replaced at positions, F7, F8, F16, F19, F30, F32, F41 and F48) *i.e.*, similar to WT. However, all the mutants showed similar dodecamer form, like WT, except F16 and F19, both of which were very large oligomers, with a smaller amount of specie as dimers or dodecamer. Interestingly, unstable oligomers were successful in displaying chaperone activity for Luc and MDH, equal to WT. The F16 and F19 of PsHsp18.1 correspond to those residues in TaHsp16.9 from wheat (*Triticum aestivum*), which are shown to be responsible for maintaining oligomeric assembly (Van Montfort *et al.*, 2001). Further analysis of these variants about cross-linking with substrates showed that crosslinking occurs only at elevated temperature and, the maximum crosslinking between PsHsp18.1 and MDH/Luc was seen in the N-terminal arm, suggesting it to be responsible for binding to denatured proteins (Jaya, Garcia and Vierling, 2009).

Another study in which phenylalanine in the NTD of a sHSP was analyzed for importance in structure and function is that involving F29 and G30 in *Drosophila melanogaster* HSP27 which are highly conserved, when compared to human and

other mammals' sHSPs. The F29 of DmHsp27 corresponds to F27 in cryaA and F28 in cryaB in humans and F17 of Hsp20_METM6 from *Methanococcus maripaludis*, an archaean. The whole NTD truncated mutant of DmHsp27 resulted in abundance of dimer structure (Leroux *et al.*, 1997), and was unable to protect the thermally aggregated substrates. Thus emphasizing the fact that oligomerization as well as chaperone activity in eukaryotic sHSP require its NTD. Deletion of a highly conserved FGFG motif in NTD of DmHsp27 resulted in considerably large oligomers while exhibiting chaperone activity as efficient as the WT. The major quaternary specie in point mutants F29Y and F29A showed increase in oligomeric size (~800 KDa) then WT. However, when the hydrophobicity was decreased (F29A) small amount of oligomeric specie similar to WT, were also produced. On the other hand, the replacement of glycine at position 30 and 32 in the FGFG motif with a charged residue arginine resulted in dramatic increase in the size of oligomer, to almost double than that of the WT, *i.e.* 1100 KDa. All the variants of FGFG motif exhibited chaperone activity similar to WT when heat aggregated substrates were used. This result helps to conclude that FGFG motif in DmHsp27, has noteworthy role in oligomerization and subunit exchange and little or no momentous role in exhibiting the chaperone activity (Moutaoufik *et al.*, 2017).

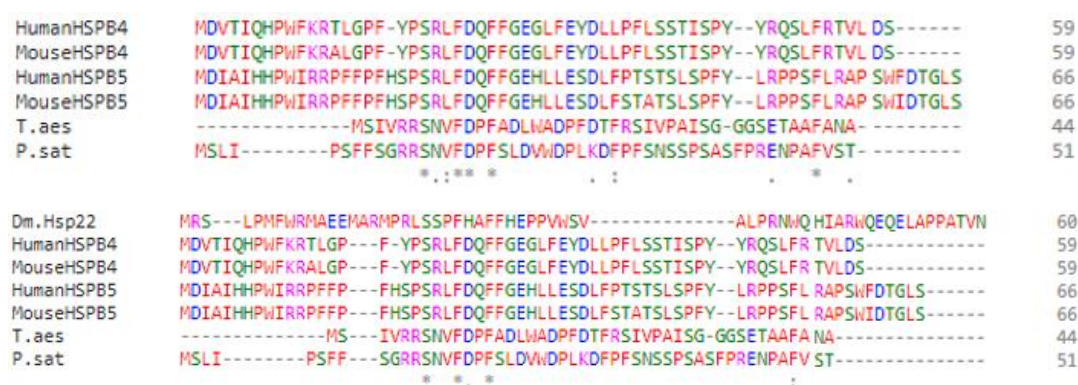


Figure 1.5. MSA of N-terminal domain of the sHSPs of different organism showing the conserved residues. HSPB4 from human with accession number P02489.2, HSPB5 from human with accession number P02511.2, HSPB4 from mouse with accession number AAH85172.1, HSPB5 from mouse with accession number NP_034094.1, TaHsp16.9 from wheat (PDB: 1GME), PsHsp18.1 from pea with accession number P19243.1 and Hsp22 from *Drosophila melanogaster* with accession number AAA28635.1.

In human α B-crystallin another point mutation in NTD involved proline, a hydrophobic residue, critical for its effects on structure of protein. A combination of mutations of P39S, P46R and F61I or P39S, P51R/P52R and S41A, along with other mutations in ACD and CTD did not affect the chaperone activity of the protein at all, however, the oligomeric structure of the protein changed, resulting in increased multimeric mass (Horwitz *et al.*, 1998).

1.5.2 Studies in Eubacteria

There are limited information regarding the functional roles of NTD of prokaryotic sHSPs. A study on a bacteria *Bradyrhizobium japonicum* sHSP, (HspH) reported that the removal of the initial 15 residues of its NTD resulted in loss of chaperone function against thermal aggregation of CS as well as increase in oligomeric size of the sHSP (Studer *et al.*, 2002). Not many studies are available regarding the point mutation in the NTD of bacterial sHSPs. Two sHSPs (IbpA and IbpB) are found in the bacteria, *E.coli*. *Deinococcus radiodurans* also is reported to possess two sHSPs, Hsp17.7 and Hsp20.2, where the former is present as dimers and tetramers (which could be due to short CTD and IXI motif lying closer to ACD), the latter shows the typical oligomer assembly mostly, the spherical 36-mer and 18-mer, which dissociates into smaller species in response to stress (Bepperling *et al.*, 2012). When the initial 11 NTD residues of the bacterial sHSP, AgsA, from *Salmonella typhimurium*, were removed, 18-mer oligomers were also observed. In addition to this truncation, if CTD was also deleted, the sHSP formed 24-mer oligomers. The NTD truncated variants had decreased protection effect towards chemically denatured lysozyme as compare to WT (Mani *et al.*, 2016).

1.5.3 Studies in Archaea

Most of the species in Archaea domain possess one or two sHSPs, with a few exception, for example, *Halobacterium* possessing five homologues of them

(Laksanalamai and Robb, 2004). The sHSP of an archaea *S.tokodaii*, StHsp14.0, and its various truncated NTD variants were engendered and evaluated for their structure as well as chaperone function. All the mutants, namely Δ N8, Δ N15 and Δ N21, had oligomers, sizes of which were similar to that of WT, (*i.e.*, 24-mer). The mutants' oligomers retained the spherical structure, similar to WT but displayed a hollow core, implying the fact that the major portion of NTD in a sHSP quaternary composition, occupies its core and is involved in subunit exchange. The efficiency of subunit exchange was decreased in Δ N15 and Δ N21 which, indirectly, led to decrease in their protective effect, thus correlating the molecular chaperone activity and subunit exchange (Usui *et al.*, 2004). Similarly, in *S.solfataricus* sHSP, Hsp20.1, elimination of NTD alone or in addition to CTD, resulted in mutant protein with a complete loss of protection of substrate enzyme, MDH, against thermal aggregation. This may signify the hypothesis that NTD is used by sHSPs for interaction with aggregating substrates (Liu, J. Chen, *et al.*, 2015). Hydrophobic surface of the NTD of StHsp14.0 is facing outward and expected to participate in inter-dimer contacts for oligomer formation. The WT of StHsp14.0 exist as 24-mer, however, its CTD truncated variants in crystal structure studies were found to be present as dimers, indication yet again, the role of CTD in formation of a stabilized oligomer (Takeda *et al.*, 2011; Hanazono *et al.*, 2012).

Bioinformatics tool analysis has revealed that Hsp27 of human has a unique 12-14 residue long peptide near the junction of NTD and ACD, which is absent in other sHSPs of human. When this short peptide is deleted stable large oligomers of Hsp27 and Hsp27D3 are produced with minimum polydispersity (Shi *et al.*, 2006). This peptide is rich in proline residues at the N-terminal end and alanine residues at the carboxyl end. When it is inserted in corresponding region of Hsp16.5 of the archaea *Methanococcus jannaschii* (the variant name Hsp16.5-P1), its monodispersed nature of 24-mer oligomer was altered to polydisperse nature. The largest oligomers of exactly double the size, *i.e.*, 48-mer, had higher substrate binding affinity than the WT (McHaourab, Lin and Spiller, 2012). The larger oligomers also had a highly ordered structure, very similar to the 24-mer of WT (Mishra *et al.*, 2018), with NTR

forming its core. In the oligomerization process, 30-mer and 36-mer were the intermediates which were disordered or molten-like in structure. By deletion mutagenesis in this peptide, a region PLPP was identified, as the region necessary for subunit exchange. When it was deleted in Hsp27, the resultant mutants formed stable large oligomers with negligible subunit exchange and chaperone activity.

1.6 Mini-Chaperones/Micro-Peptides as chaperones

The usage of short peptides as therapeutics is, an attractive and continuously developing strategy (Raju, Santhoshkumar and Sharma, 2016). It was found that smaller fragments of a protein, including the active site, possessed some percentage of the parent activity in the isolated form. For example, a shorter peptide (155 residues; 191 – 345) of the molecular chaperone GroEL, possessed the refolding and reconditioning abilities towards its substrate proteins (Zahn *et al.*, 1996; Altamirano *et al.*, 1997). After identification of the chaperone activity of α -crystallins (Horwitz, 1992), their substrate binding sites, within the ACD, were identified (Krishna Sharma, Kaur and Kester, 1997; Sharma *et al.*, 1998). These substrate binding sites were short peptide 4 – 18 residues in length and possessed the activity of the parent molecule (Sharma *et al.*, 2000). Since then, an extensive research has been carried out to identify the substrate binding sites within the ACD of α -Crystallins, synthesis of the peptides and assay them for chaperone activity with respect to the parent molecule (Bhattacharyya and Sharma, 2001; Santhoshkumar and Sharma, 2004; Bhattacharyya *et al.*, 2006; Raju, Santhoshkumar and Sharma, 2016). As mentioned above, ACD is highly conserved in terms of structure and function, therefore, identification of such sequence in the ACD of other sHSPs is a relatively easy task. However, due to scarcity of such conserved motifs in the NTD, defining such peptides in this region is not straight forward. Gliniewicz (2019) first reported that, the whole NTD of HSPB1 from human, possessed similar chaperone activity to its parent molecule for selected substrates (MDH and lysozyme) (Gliniewicz *et al.*, 2019). Recently, shorter peptides within the NTD of HSPB1, with length as short as

five aminoacids, were reported to have chaperone effect (protection against CS) and were referred as micro-chaperones (Kho *et al.*, 2021). These peptide included the highly conserved WDPF sequence, along with serine (SWDPF peptide), previously shown to be responsible for chaperone activity and oligomerization of this sHSP (Thériault *et al.*, 2004; McDonald *et al.*, 2012).

1.7 Why study sHSPs/Significance of sHSPs in life

Initially sHSP are thought as the saver of life in human body, by being the first line of response to stress and preventing formation of insoluble aggregates of proteins which might lead to cell death. But, recent studies have shown that any disturbance in the balance of production of sHSPs may have drastic effects. sHSPs play an important role in maintaining cell viability, however, they sometimes function in favor of certain cancer tissues by regulation of various properties of cancer stem cell which include giving strength of better survival, acquiring chemoresistance and being more invasive as compared to normal cells. For example, besides a few instances, where sHSPs are observed to decrease, most of the cancers show over-expression of HSPBs. Therefore, several cancer drugs act by limiting these effects, for example, quercetin, an anticancer agent, in addition to decreasing the expression of HSPB1, it blights its phosphorylation thus curtailing its activity (Xiong *et al.*, 2020). The carcinogenic effect and increase in expression of sHSPs, especially HSPB1 and HSPB2 in several cancers could be attributed to their anti-apoptotic and anti-oxidant functions.

As obvious, small Heat Shock Proteins are involved, in-vivo, in proteostasis, therefore, any mutation in the primary structure of these sHSPs leads to compromised chaperone activity and thus consequences in the form of various diseases. Of the many natural human sHSPs mutations known, those related to N-terminal arm are few. Of these, G34R, P39L and E41K in the NTD of HspB1 in humans, result in various diseases. For example, associated with distal hereditary motor neuropathy (dHMN) is attributed to G34R and E41K mutations, whereas

P39L mutation is associated with type II Charcot-Marie-Tooth disease (Houlden *et al.*, 2008). For the former two mutations, disease symptoms are detected rather late as opposed to childhood (less than 10 years) for the E41K mutation and severity of these symptoms also vary. Therefore, these mutants were studied in-vitro for their physico-chemical properties. All the mutants showed aggregation at temperatures lower than that for WT. The chaperone activity of all the mutants was lower than WT for MDH, lysozyme and significantly lower for insulin. The mutants showed oligomer size larger than the WT (the oligomer size of WT was 540 KDa while that of P39L was 700 KDa), in SEC profile, where the size of oligomer was dependent on loading concentration in G34R and E41K but independent in P39L. These mutations also decreased the phosphorylation induced dissociation of HspB1 oligomer (Muranova *et al.*, 2015). Thus, these results suggest that all the three positions in human HspB1 (34, 39, 41) and its residues (Glycine, Proline and Glutamic acid respectively) are important for maintaining oligomeric assembly, its subunit exchange and dissociation as well as the chaperone function.

Recent studies, using RT-PCR and immunohistochemistry techniques, have also shown that acute-inflammation induced by ethanol leading to increase in human-HSPB1 expression resulted in increased pro-inflammatory response and high levels of TNF α but decreased apoptosis in transgenic mice neural cells, suggesting a complex yet protective role of HSPB1 in neural cells following acute-inflammation. (Dukay *et al.*, 2021). HSP27 has a well-established role in epithelial to mesenchymal tissue transformation (EMT) in tissues which is crucial for carcinogenesis. The complexity of the matter increases to an extent, where over-expression of HSPB1 in lungs promotes EMT leading to fibrosis while in renal and myocardial tissue increase in HSPB1 expression downregulates fibrosis by exerting anti-inflammatory effects (Zhang *et al.*, 2021). Also levels of HSPB1 in serum following oxidative stress and/or hypoxia in kidney, serves as a promising alternative to present as a marker/predictor for analysis of contrast induced-acute kidney injury (Jaroszyński *et al.*, 2021). Another study has reported increase in levels of Hsp22 and α B-crystallin while decrease in production of Hsp27 in STZ-induced Diabetic mice (Reddy *et al.*,

2021). In addition to this, many diseases arise following mutations that occur in the HSPBs.

Therefore, there is growing evidence that Hsp27 can be a potential target when developing new and novel, site specific treatments for neuronal diseases especially in Parkinson's and Alzheimer's diseases (Navarro- zaragoza *et al.*, 2021).

1.8 Scope of this Study

In the recent past, there has been numerous studies to understand the mechanism of action of sHSPs, but, it is increasingly clear that no single model is sufficient to describe structure, function and dynamic behaviour of sHSP. It is generally agreed that substrate binding is facilitated by an increase in exposure of hydrophobic surfaces. Identification of specific residues that contribute to substrate interaction is an area of research that still needs to be further pursued. Complete or partial truncation of NTD of the sHSPs, provided evidences about the role of NTD in substrate recognition, binding and stabilizing sHSP oligomers in some sHSPs (Usui *et al.*, 2004; Heirbaut *et al.*, 2014; Webster *et al.*, 2020). However, single residue targeting mutagenesis studies, especially those involving NTD, are very limited. In the reports involving PsHsp18.1, *S.pombe* Hsp26 and HSPB5 of human and mouse, hydrophobic residues mainly phenylalanine has been shown to play role in chaperone activity (Plater, Goode and Crabbe, 1996; Horwitz *et al.*, 1998; Jaya, Garcia and Vierling, 2009). The 3-D crystal structure analysis of the sHSP from *S.solfataricus*, indicated the involvement of hydrophobic interactions between NTD of adjacent dimers to form tetramers (Hanazono *et al.*, 2013a). However, none of the studies provide detailed insight into roles of hydrophobic residues individually and in combination in different segments of the NTD. In addition, previously, the role of highly conserved positively charged residues in chaperone function are investigated in human HSPB1 (Shatov *et al.*, 2018; Shatov, Sluchanko and Gusev, 2021). Results pinpoint the role of NTD glutamate in formation of oligomers by salt bridges, thus affecting chaperone activity. Nevertheless, there is no report that provides an

exemplary detail about negative charges in the NTD segment of any archaeal sHSP. Furthermore, the hydrophobicity around the junction of NTD and ACD also remains to be explored in terms of its role in stabilizing the dimer and thus the chaperone function.

In this study we aimed to fill the gaps in literature concerning the role of hydrophobic residues in different segments of the NTD alone and in combination, in substrate interaction, oligomeric organization and chaperone function. In addition, negatively charged residues of the NTD are probed for functional analysis. The model organism, whose sHSP is being investigated is an facultative anaerobic heterotroph, *Thermoplasma volcanium*. Phylogenetically, it belongs to the genus *Thermoplasma*, which is classified under archaeobacteria. According to the 16S rRNA sequence studies, it is a member of the Thermoplasmatales archaea (Yelton *et al.*, 2011). It can grow at temperatures between 33°C and 67°C (Kawashima *et al.*, 2000) and pH 1 – 4 (Seegerer, Langworthy and Stetter, 1988). However, the optimum temperature for its growth is 60°C and optimum pH is 2.0 (Kawashima *et al.*, 2000). It was first isolated from solfatara fields of vulcano, in Italy. The cells lack a cell wall but possess flagella (Seegerer, Langworthy and Stetter, 1988). Previously, *Tpv* sHSP 14.3 gene has been cloned first time in our laboratory and this gene, with locus name TVN0775, for sHSP 14.3 lies between 790978 and 791352 nucleotide and it consists of 375 nucleotides. Recombinant sHSP protein which is His-tagged has been expressed, purified and characterized (Kocabiyik and Aygar, 2012). The 14.3 sHSP is spherical, with sphere size between 10 and 20 nm. This sHSP is 124 amino acids long and has a monomeric molecular weight of 14.33 KDa with isoelectric point of 9.2. In this study, it has been shown that heterolog expression of *Tpv* sHSP 14.3, increased the temperature tolerance of *E.coli*. In-vitro it has prevented heat inactivation of the enzymes citrate synthase and L-glutamate dehydrogenase from pig heart and bovine source, respectively (Kocabiyik and Aygar, 2012).

Here, we analyzed the proximal and distal part of NTD of *Tpv* HSP 14.3 that possess hydrophobic amino acid cluster. Besides single point mutations, double and triple mutations were also designed to understand the effect of simultaneous decrease of

the hydrophobicity in this cluster of the NTD for the first time. Also, the role of negative charge in the NTD was planned to be explored in detail by targeting both the glutamate residues (E11 and E22), separately. Thus all the phenylalanines in the NTD (F7, F8 and F26) and highly conserved I5, I27, V23 and V31 were targeted for mutagenesis. Single, double and triple mutants generated were extensively studied for their contribution in chaperone activity and structure integrity of the *Tpv* HSP 14.3. The I5 is located in the intrinsic disordered region of the NTD and the effect of its substitution alone or along with F7 and F8 was investigated to see the changes if there would be any in this region, which provided new information to the literature. The results obtained from this study will provide new insights of the functional and structural characteristics of the hydrophobic cluster and negative charge in the sHSP's NTD. This in turn will help in analogous interpretation and hypothesizing the single residue effects of NTD in sHSPs of other origin, including humans.

This study also involved designing of small peptides from the NTD sequence of the *Tpv* HSP 14.3, which possessed chaperone activity. There is only one report about mini peptides from the NTD of human HSPB1 and no such peptides have been produced from the NTD segment of any sHSP and investigated, so far.

We have used the bioinformatics tools to identify the critical residues which may contribute to the structural and/or functional integrity of the *Tpv* sHSP 14.3. Those residues were highlighted which were found to be highly conserved in paralogs and orthologues proteins and some of which have already been studied in orthologues.

This was followed by employing a set of techniques, like molecular cloning and site-specific mutagenesis, to generate a series of NTD mutants. The mutants were designed in such a way, to understand the effect of a single residue alone and/or in combination with other residues on the structure and function of the resulting mutant. Therefore, a collection of single, double and triple mutants were expressed in *E.coli* cells and purified for structural and functional analysis.

The mutants were tested for their protective effect against heat induced aggregation by light scattering analysis using pig heart CS as substrate. Chaperone activity of the

mutant sHSP variants were also studied by measuring their CS activity protection effect against heat inactivation as compared to the WT.

For structural analysis, BN-PAGE was performed for selected mutants and WT. The 3-D model structure for the all mutants was generated using online computational tools and software, Display studio Visualizer and Chimera 1.11.2. The 3-D model structures were analyzed for the structural and configurational changes. Also, alterations in inter and intramolecular bonds, that occur as a result of mutation, within the NTD of *Tpv* sHSP 14.3, were also determined.

Lastly, short peptides of *Tpv* sHSP 14.3 were designed with the aim of possessing either similar or higher chaperone activities than the whole sHSP protein, for their potential uses in various biotechnological and therapeutical applications.

CHAPTER 2

MATERIAL AND METHODS

2.1 Materials

The chemicals used in this study were of HPLC grade or Molecular Biology grade with maximum purity. The water used in this study is distilled and sterilized.

All primers used in mutagenesis were prepared from Alfagen, Ankara, Turkey.

For DNA experiments, O'GeneRuler DNA Ladder Mix, ready-to-use and GeneRuler DNA Ladder Mix, ready-to-use were used as molecular size markers, while for protein experiments, Pre-stained protein ladder PageRuler, ready to use, was used (Fermentas, U.S.A.).

2.1.1 Chemicals

Sodium hydroxide (NaOH), sodium chloride (NaCl), sodium dihydrogen phosphate, magnesium sulfate, potassium dihydrogen phosphate, potassium hydroxide, β -mercaptoethanol, HEPES, α -D-Glucose, sodium dodecyl sulfate (SDS), glycine, ethylenediaminetetraacetic acid (EDTA), methanol, isopropyl- β -D-thiogalactoside (IPTG) and coomassie blue-G250 were purchased from Merck, Darmstadt, Germany.

Glycerol, ammonium persulphate, ethidium bromide, Brilliant Blue R, TRIS, agarose, TEMED, ampicillin antibiotic, acrylamide, bis-acrylamide, oxaloacetate, DTNB, acetyl-CoA, 6-amino hexanoic acid, tricine, ethanol and glacial acetic acid were purchased from Sigma Chemical Company, Missouri, USA.

Imidazole, tryptone and yeast extract were purchased from Fluka, U.S.A.

Agarose Low EEO was purchased from AppliChem, Darmstadt, Germany.

2.1.2 Enzymes

NdeI, *BamHI* (used in sub-cloning procedure) were purchased from New England Biolabs U.S.A. *NdeI*, *BamHI*, *Sall*, *SacI*, *HindIII*, *BglII* (for characterization) were purchased from Fermentas, U.S.A.

Sall (For enzyme protection assay) was purchased from Roche, Switzerland.

Porcine heart citrate synthase (C3260) and alcohol dehydrogenase was purchased from Sigma, U.S.A.

2.1.3 Kits

QIAprep Spin Miniprep Kit and QIAquick Gel Extraction Kit were purchased from Qiagen, Germany. GeneArt Site-Directed Mutagenesis Kit and GeneJET Plasmid Miniprep Kit were purchased from ThermoFisher Scientific, U.S.A. QuikChange Site-Directed Mutagenesis Kit was purchased from Agilent Technologies, U.S.A. DNA Ligation Kit was purchased from Novagen, Merck, Germany.

2.1.4 Competent Cells

BL21(DE3) competent cells were purchased from New England Biolabs, U.S.A. XL1-Blue supercompetent cells were purchased from Agilent, California, USA. One Shot Max Efficiency DH5 α TMT1^R cells were purchased from ThermoFisher Scientific, Massachusetts, USA. NovaBlue competent cells were purchased from Novagen, Merck, Germany.

2.2 Methods

2.2.1 Subcloning of TVNO775 gene for expression

The *Tpv* sHSP 14.3 gene was subcloned into pET23a expression vector (Novagen, Merck, Germany) from recombinant cloning vector pDrive/2_14.3 to express the gene from its own AUG. The pDrive/2_14.3 plasmid, containing together with the TVN0775 gene its 232 nucleotides upstream and 74 nucleotides downstream sequences, was from our laboratory collection. The cloning strategy is as follows: The recombinant pDrive/2_14.3 plasmid was isolated using the QIAprep Spin Miniprep Kit (Qiagen, Germany). It was used as the gene source to amplify the TVN0775 gene by PCR using the primers, TVN0775FP: (5'-CATATGTATACACCCATAAAGTTCTTTACG-3'), and TVN0775RP: (5'-GTATGCTTGATGTGATTGGGTGGGATCC-3') containing the *NdeI* and *BamHI* restriction sites (underlined), respectively. The primers we designed were synthesized by New England Biolabs, U.S.A. Then, PCR products were purified from gel by using QIAquick Gel Extraction Kit (Qiagen, Germany) and digested with *NdeI* and *BamHI* restriction enzymes. Digested fragments and the pET23a vector cut with the same restriction enzymes, after purification using QIAquick spin column (QIAquick Gel Extraction Kit), were ligated using DNA Ligation Kit (Novagen, Merck, Germany). Ligation products were introduced into BL21(DE3) competent *E.coli* cells (New England Biolabs, USA), with transformation efficiency of 4×10^6 cfu/ μ g pUC19 DNA, and plated onto LB agar (pH 7.5; containing 100 μ g ampicillin/ml). Positive clones among the other transformants that carry *Tpv* sHSP 14.3 gene were identified by restriction enzyme digestion analysis of the plasmids isolated from randomly selected colonies. Putative recombinant plasmids were further confirmed by sequencing both DNA strands. The recombinant plasmid from one of the selected positive clones was named as pET23a Mix2/15-WT to be used in downstream experiments and was transformed into NovaBlue competent cells (Novagen, Merck, Germany) for efficient long term storage.

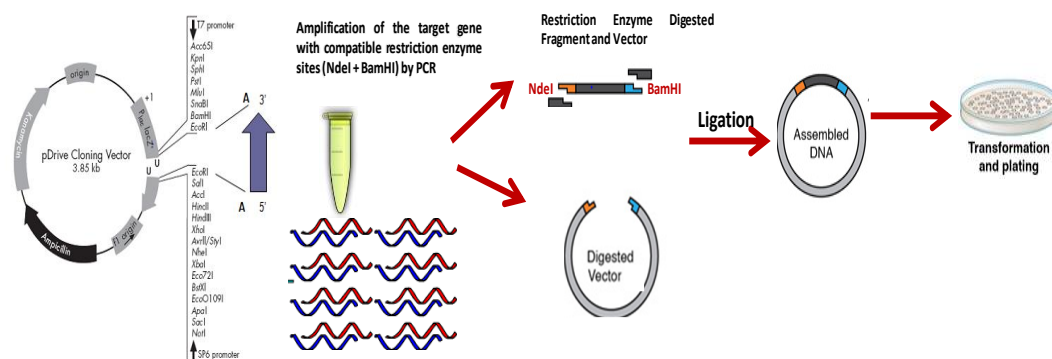


Figure 2.1. Flow chart for cloning

2.2.2 Site-Directed Mutagenesis

A series of single, double and triple mutations at selected positions were attempted in the NTD of the *Tpv* sHSP 14.3 to understand their impact on structure and/or function of the sHSP. QuikChange II Site-Directed mutagenesis kit (Agilent, California, USA) was used for the mutagenesis. The plasmid pET23a Mix2/15 was used as template in all the mutagenesis experiments. This plasmid DNA was isolated from the NovaBlue competent cells (Novagen, Merck, Germany) using QIA prep Spin Miniprep Kit (Qiagen, Germany). 26 ng of template plasmid was used in a reaction mixture of 50 μ l. For the purpose of constructing each respective mutation, two complementary primers, the forward primer and the reverse primer, were designed, and named mutagenic primers. The list of the FP and RP for all the mutants generated for this study are given in Table 2.1.

The both strands of the template plasmid were replicated in-vitro by PCR using High fidelity polymerase enzyme, *PfuUltra* HF DNA polymerase in the presence of mutagenic primers set for 16 amplification cycles. Then PCR products containing WT and mutant plasmids, were digested with *DpnI* enzyme to eliminate the non-mutant parental plasmid DNA and minimize the chance of its transformation into cells, in the next step.

Table 2.1. Forward and reverse complementary primer for each mutant generated. Exchanged codons are underlined

Mutation Name	Primer Design
I5T (FP)	5'-GTATACACCC <u>ACA</u> AAGTTCTTTACG-3'
I5T (RP)	5'-CGTAAAGA <u>ACTTTGT</u> GGGTGTATAC-3'
F8Y (FP)	5'-CACCCATAAAGTTCT <u>TAT</u> ACGAATGAGATG-3'
F8Y (RP)	5'-CATCTCATTTCGT <u>ATAGA</u> ACTTTATGGGTG-3'
I5TF8Y (FP)	5'-GTATACACCC <u>ACA</u> AAGTTCT <u>TAT</u> ACGAATGAG-3'
I5TF8Y (RP)	5'-CTCATTTCGT <u>ATAGA</u> ACTTT <u>TGT</u> GGGTGTATAC-3'
F7SF8Y (FP)	5'-CACCCATAAAG <u>TCCTAT</u> ACGAATGAGATG-3'
F7SF8Y (RP)	5'-CATCTCATTTCGT <u>ATAGGACT</u> TTTATGGGTG-3'
I5TF7SF8Y (FP)	5'-GTATACACCC <u>ACA</u> AAG <u>TCCTAT</u> ACGAATGAGATG-3'
I5TF7SF8Y (RP)	5'-CATCTCATTTCGT <u>ATAGGACT</u> TT <u>TGT</u> GGGTGTATAC-3'
E11V (FP)	5'-GTTCTTTACGAAT <u>GTG</u> ATGATAAAAAACG-3'
E11V (RP)	5'-CGTTTTTATCAT <u>CAC</u> ATTTCGTAAAGAAC-3'
M12T (FP)	5'-CTTTACGAATGAG <u>CG</u> GATAAAAAACGTATCG-3'
M12T (RP)	5'-CGATACGTTTTTTAT <u>CGT</u> CTCATTTCGTAAAG-3'
E22G (FP)	5'-CGAATACTGTGAAA <u>GGG</u> GTCTCATCC-3'
E22G (RP)	5'-GGATGAGAC <u>CCC</u> TTTCACAGTATTCG-3'
V23GF26YI27T (FP)	5'-CTGTGAAAGAG <u>GGG</u> CTCATCCT <u>TATACAT</u> ATCCACCAGTC-3'
V23GF26YI27T (RP)	5'-GACTGGTGGATAT <u>TGTATAGGATGAG</u> CCCTCTTTACACAG-3'
F26YI27T (FP)	5'-GAGGTCTCATCCT <u>TATACAT</u> ATCCACCAGTC-3'
F26YI27T (RP)	5'-GACTGGTGGATAT <u>TGTATAGGATGAG</u> ACCTC-3'
V31G (FP)	5'-CCTTTATATATCCACC <u>AGC</u> ACGTTATATCAAGATAGC-3'
V31G (RP)	5'-GCTATCTTGATATAACGT <u>GCCT</u> GGTGGATATATAAAGG-3'
V31IL33I (FP)	5'-CCTTTATATATCCACCA <u>ATCACGATAT</u> ATCAAGATAGCTCTG-3'
V31IL33I (RP)	5'-CAGAGCTATCTTGATAT <u>TATCGTGAT</u> TGGTGGATATATAAAGG-3'

The resultant nicked plasmid DNAs that contain our desired mutations were immediately transformed into XL1-Blue supercompetent cells (Agilent, California, USA) for nick repair, following the recommended protocol: briefly, 1 µl of the *DpnI* treated plasmid mixture was mixed with 50 µl of XL1-Blue supercompetent cells in BD falcon polypropylene tube. This mixture after incubation on ice for 30 minutes, was subjected to heat shock at 42°C for 45 seconds followed by cooling on ice for two minutes. Next, 0.5 ml of reconstituted NZY⁺ broth was added to these tubes and the cells were allowed to multiply at 37°C, with shaking at 250 rpm for one hour. The transformed cells were then plated on ampicillin (100 µg/ml) containing LB (pH 7.0) agar plates. Positive clones containing the desired mutations were identified by restriction enzyme digestion analysis and DNA sequencing. The mutant plasmids were then transformed into competent BL21(DE3) *E.coli* cells (New England Biolabs, USA), for expression. The transformation procedure for BL21(DE3) *E.coli* cells (New England Biolabs, USA) was similar to that mentioned for XL1-Blue

supercompetent cells, with minor changes mentioned here. Instead of NZY⁺ broth, SOC broth was used, the incubation on ice was five minutes and the heat shock was just for 30 seconds. While for BL21(DE3) cells prepared in our lab, the step of heat shock was omitted. The LB agar plates used for BL21(DE3) were pH 7.5.

For the F8YV31G double mutation, a combination of forward and reverse primers of single mutants (F8Y and V31G) were used by employing the GeneArt Site Directed Mutagenesis Plus kit (ThermoFisher Scientific, Massachusetts, USA). Kit protocol, to generate a double mutant product with mutations far apart, on the same protein was followed. The FP of one mutation and the RP of another mutation, was added to the two separate PCR tubes and by the help of AccuPrime *Pfx* DNA polymerase, amplification of the plasmids containing the desired mutations was achieved in 15 replication cycles. The products of both PCR tubes then were mixed and allowed for a recombination reaction which was stopped by addition of EDTA for 15 minutes. The recombined plasmids were transformed into One Shot Max Efficiency DH5 α TMT1^R cells (ThermoFisher Scientific, Massachusetts, USA). The transformation mixture was incubated in ice for 15 minutes, followed by heat shock at 42°C for 30 seconds and then incubated again on ice for 2 minutes. After addition of 250 μ l of SOC medium the mixture was immediately plated on Luria Bertani agar plates, pH 7.0.

2.2.3 Preparation of BL21 competent cells.

In order to express mutant proteins, the recombinant mutant plasmids of *Tpv* sHSP 14.3 were transferred into BL21(DE3) cells. This *E.coli* strain is suitable for expression of recombinant proteins, which possess IPTG inducible T7 RNA polymerase gene that promotes the transcription of the gene lying down stream of T7 promoter (Novy *et al.*, 2001, Studier & Moffatt, 1986). In the transformation experiments, besides BL21(DE3) competent cells purchased from (NEB, U.S.A), competent cells that we prepared were used. For preparation of the BL21(DE3) competent cells, 10 ml of the Luria Bertani broth was inoculated with a single colony

and incubated overnight at 37°C with shaking at 160 rpm (Heidolph, Germany). Then 0.2 ml of overnight culture was added to 20 ml of LB Broth in a conical flask and incubated at 37°C by shaking at 262 rpm until the OD₆₀₀ reached 0.5. The cells were harvested by centrifugation (Sigma) at 4000g and 4°C for 10 minutes. The BL21(DE3) cell pellet was dissolved in Transformation and Storage solution (Chung, Niemela and Miller, 1989). Aliquots of 100 µl were transferred into eppendorf tubes and stored at -80°C.

2.2.4 Small scale lysate preparation and heat-treatment

The recombinant BL21(DE3) *E.coli* cells were grown in 50 ml Luria Bertani broth (pH 7.5; containing 100 µg ampicillin/ml) at 37°C by monitoring the optical density, with continuous shaking at 300 rpm. The protein expression was induced by addition of IPTG to a final concentration of 0.4 mM, when OD₆₀₀ = 0.6 was reached (OD₆₀₀ = 0.5 for mutant I5TF8Y). The cells were harvested at 4°C for 20 minutes at 5000g (Sigma Centrifuge, St. Louis, USA), and the pellets were immediately suspended in 8 ml of Lysis Buffer. Then, cells were disrupted by sonication (Sonics sonicator, Merck, Germany), with tip diameter 1/2" (13 mm). Sonication was done with a 30 seconds interval at an amplitude of 40%. The cell debris was removed after centrifugation at 10,000g for 30 minutes. Then the supernatant was subjected to heat treatment at 65°C, 70°C and then 80°C for 15 minutes, to analyze the heat stability of the recombinant protein. For those recombinant proteins which were found stable, the heat treatment was carried out upto 30 minutes. While for those which showed decreased heat stability at 65°C and above, temperature as low as 60°C for 20 minutes was also used. This step was employed to remove the heat labile host proteins. Following heat treatment, the suspension was cooled on ice for 30 minutes and then centrifuged at 12,000g for 60 minutes at 4°C, to remove the denatured host proteins. The clear supernatant, *i.e.*, cell free extract of the WT and mutant cells were kept at -80°C for further use. The presence of protein was confirmed at each step by SDS-PAGE.

2.2.5 Recombinant Protein Expression and Purification (Large scale)

The recombinant BL21(DE3) *E.coli* cells were grown in Luria Bertani medium (1000 ml) (pH 7.5; containing 100 µg ampicillin/ml) at 37°C by monitoring the optical density, with continuous shaking (300 rpm) for expression of the recombinant WT and the engineered mutant proteins. The protein expression was induced by addition of IPTG to a final concentration of 0.4 mM, when OD₆₀₀ = 0.6 was reached. The cells were harvested at 4°C for 20 minutes at 5000g (Sigma Centrifuge, St. Louis, USA), and the pellets were kept at -80°C until purification. For preparation of the cell lysates, after thawing the pellets were suspended in Lysis Buffer (25 mM Tris-HCl, 1 mM EDTA, 30 mM NaCl, pH 7.5). Then, cells were disrupted by sonication (Sonic sicator, Merck, Germany), with tip diameter 1/2" (13 mm). Sonication was done with a 30 seconds interval at an amplitude of 40% for an average of 200 seconds. The cell debris was removed after centrifugation at 20,000g for 20 minutes. Then the supernatant was subjected to heat treatment at 70°C for 30 minutes, to remove the heat labile host proteins. Soluble fraction free of denatured proteins, containing *Tpv* sHSP 14.3 WT or mutant protein, was obtained after centrifugation of the suspension at 12,000g for 60 minutes at 4°C, after cooling the suspension on ice for 30 minutes. The supernatant obtained was first filtered through 0.45 µm membrane (Millipore) and then purified by HPLC (ÄKTA prime, Amersham Biosciences USA). The clear supernatant was applied to a HiTrap Q XL anion exchange column (GE Life Sciences, Chicago, USA) pre-equilibrated in anion exchange buffer (20 mM Tris at pH 8.02) and eluted, at a flow rate of 5 ml/min with a linear gradient of 0 – 1000 mM NaCl in the same buffer. The fractions containing the protein *Tpv* sHSPs as determined by SDS-PAGE analysis were pooled, and then desalted (Phosphate buffer: 50 mM Sodium Phosphate and 150 mM Sodium Chloride, pH 7.2) and concentrated by ultrafiltration (VivaSpin 20 ml 5MWCO). Purified proteins were stored at -20°C. The concentration of the proteins in fractions and in purified samples was determined by PicoDrop at 280 nm.

2.2.6 Citrate Synthase enzyme aggregation assay

The Citrate Synthase (Sigma C3260) enzyme is a dimeric protein having two identical subunits, each with a molecular weight of 48.969 KDa (Remington, Wiegand and Huber, 1982). At elevated temperature, the CS unfolds and form high molecular weight particles that scatter light. Molecular chaperones including sHSPs are known to suppress heat-induced aggregation of CS (Buchner *et al.*, 1991; Usui *et al.*, 2004).

Citrate synthase from pig heart is a mesophilic enzyme with an optimum temperature of 35°C and is inactivated at temperatures higher than 43°C (Kocabiyik and Aygar, 2012). The ability of *Tpv* sHSP 14.3 variants to suppress the CS aggregation at elevated temperatures, was studied by incubating CS alone and in the presence of sHSP. Porcine heart citrate synthase (Sigma EC. 4.1.3.7) is in the form of ammonium sulfate suspension. It was subjected to ultrafiltration (5 KDa molecular weight cut-off filter unit, Low Binding cellulose, Millipore Amicon) to exchange its buffer with HEPES buffer (50 mM HEPES-KOH; pH 7.5). Picodrop (Picopet 01, Picodrop Ltd. UK) was used to measure the concentration of the CS as well as the chaperone protein. Aggregation assay was performed using Corning 96 well flat bottom plate (Microplate Corning 3631, 96 well, Corning Incorporated Life Sciences, Acton, MA). The reactions were prepared in 150 µl of HEPES-KOH buffer (50 mM HEPES-KOH; pH 7.5) including 9.8 µM or 49 µM sHSP and 1.4 µM CS (1:7 or 1:35 substrate:sHSP molar ratio). Thermal aggregation of CS at 45°C was monitored as increased light scattering at 320 nm in Microplate Spectrophotometer (Multiskan GO ThermoScientific) for 120 minutes through continuous shaking (with one minute interval). Each data is the mean of three independent experiments and the mean ± SD of all three independent experiments is shown by error bars on the data. The results were interpreted as the percentage heat protection when CS aggregation without chaperone was taken as 0% protection.

2.2.7 Citrate Synthase enzyme activity protection assay

The model substrate used in the enzyme activity protection assay in our experiment was pig heart citrate synthase enzyme (Sigma EC. 4.1.3.7). The assay protocol used was as described by Kocabiyik and Aygar, (2012). Serial dilutions of the citrate synthase enzyme was performed to end up with a solution of 0.03124U/ μ l enzyme in the assay buffer (20 mM Tris, 1 mM EDTA, pH 8.0). The protection effect of chaperones on the enzyme activity of citrate synthase was assayed at two different molar ratios 1:1700 and 1:850 substrate:sHSP, while, the final enzyme concentration in the reaction mixture was approximately 0.1562 U/ml, The concentration of chaperone was 214 μ g/ml and 107 μ g/ml. The negative and positive controls were devoid of chaperones. The negative control was subjected to heat treatment at accurately maintained temperature of 47°C for 10 minutes, while the positive control was not. The protection effect of the chaperones was analyzed by incubation of the test reactions, at 47°C for 10 minutes. After heating, the mixtures were immediately transferred to ice and rapidly cooled. These mixtures were then assayed for the remaining citrate synthase activity at 35°C according to DTNB method of Srere (Srere *et al*, 1963). Briefly, DTNB was added to this mixture of citrate enzyme and chaperone protein, along with the assay buffer and incubated at 35°C for 3 minutes. This was followed by addition of acetyl CoA and oxaloacetate. The absorption was monitored at a wavelength of 412 nm continuously, at controlled temperature of 35°C for 10 minutes using thermostatically controlled spectrophotometer (Schimadzu UV-1601A Spectrophotometer, Kyoto, Japan).

2.2.8 Native - Polyacrylamide Gel Electrophoresis

For the analysis of the oligomeric states and determination of the molecular masses of the oligomeric complexes of the *Tpv* sHSP 14.3 WT and its selected mutants, initially Native PAGE was employed. In this procedure, the migration of the protein, through the native gel, is dependent upon the pore size of the gel and the size and the

pI of the globular protein. The staining and de-staining solution for Native-PAGE gel was similar to the one used for SDS-PAGE gels. An acrylamide gradient gel of 4% to 20% was used in Native-PAGE. The purified protein samples of *Tpv* sHSP 14.3 WT and I5T mutant without heat treatment were loaded onto the gel. The following standard proteins were used to estimate the MW of the oligomers: Ferritin (440 KDa, 24-mer), Catalase (232 KDa, tetramer), Aldolase (158 KDa, tetramer), Ovalbumin (44 KDa, monomer) and Ribonuclease (13.7 KDa, monomer).

2.2.9 Blue Native - Polyacrylamide Gel Electrophoresis

For the analysis of the oligomeric states and determination of the molecular masses of the *Tpv* sHSP 14.3 WT and its selected mutants, Blue-Native PAGE was employed. In this procedure, under the effect of electric field, the proteins pass through the gel, in their native state. The migration of the protein, through the gel, is dependent upon the pore size of the gel and the size of the globular protein. The charge shift is achieved by including coomassie dye (0.02%) into cathode buffer to the hydrophobic domains on the surface of the proteins (Wittig, Braun and Schägger, 2006). After the destaining of the gel, these proteins are visible as blue bands on the gel.

An acrylamide gradient gel of 4% to 13% was used in BN-PAGE. Equal amounts of the purified protein samples of *Tpv* sHSP 14.3 WT and selected mutants without heat treatment (at room temperature) were loaded onto the gel. To verify the effect of heat on the oligomerization of sHSPs, samples of all proteins also were pre-heated at 70°C for 10 minutes and immediately loaded. The following standard proteins were used to estimate the MW of oligomers: Ferritin (440 KDa, 24-mer), Catalase (232 KDa, tetramer), Aldolase (158 KDa, tetramer) and Ribonuclease (13.7 KDa, monomer). Buffers and conditions of BN-PAGE were same as described by Wittig *et al.*, (2006).

2.2.10 3-Dimensional molecular modeling of *Tpv* sHSP 14.3 and its bioinformatics

Three dimensional monomer structure of *Tpv* sHSP 14.3 was generated from the primary sequence which was first determined in our lab (Kocabiyik and Aygar, 2012), while National Centre for Biotechnology Information (NCBI) annotation is from genome sequencing, using the program Easy Modeller 4.0 (Kuntal, Aparoy and Reddanna, 2010). Template structures for molecular modelling employed here were the crystal structures of HSP14.1 from *S.tokodaii* (PDB entry code: 3VQK) and HSP20 from *S.solfataricus* (PDB entry code: 4YL9). Five models for monomer were generated, while the one with highest GA341 score was chosen for the generation of dimer. The dimer structure was built by UCSF Chimera 1.11.2 (Pettersen *et al.*, 2004), using *S.solfataricus* HSP20 sHSP as the template. The model 3-D structures of *Tpv* sHSP 14.3 monomers and dimers were visualized, analyzed and compared by using Discovery Studio Visualizer (DSV) program (BIOVIA). The multiple sequence alignments were performed mostly by Clustal W Program (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) available in the EMBL-EBI data base. sHSPs with the most similar regions to NTD of *Tpv* sHSP 14.3 WT were identified by the Basic Local Alignment Search Tool (BLAST) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). The Rosetta Energy for the WT and mutant variants was calculated using online Tool ZipperDB. The hydropathy bar charts of N-terminal domain of *Tpv* sHSP 14.3 WT and its mutants were generated using a web server/online tool, pepcalc. These charts are based on Hopp-Woods scale where the aminoacids are ranked according to their solubility in water or hydrophilicity (Hopp and Woods, 1981). Two online tools/servers MUpro and I-Mutant2.0 were used to analyze the protein thermodynamic stability after introduction of mutation. The IDR of *Tpv* sHSP 14.3 protein was studied using two online tools PONDR and PrDOS. While, analysis of surface hydrophobicity of *Tpv* sHSP 14.3 WT and mutants was performed by UCSF Chimera 1.11.2.

2.2.11 Procedure for Mini peptides

Short peptides were designed, comprising of the NTD of *Tpv* sHSP 14.3 or its shorter region. Citrate synthase and alcohol dehydrogenase aggregation assays were carried out to demonstrate the functional role of our mini peptide in chaperone activity. The assays were carried out at 45°C (for citrate synthase) and 43°C (for alcohol dehydrogenase) in a 96-well microplate. The thermal aggregation of the substrate enzymes was monitored at 320 nm (for citrate synthase) and 340 nm (for alcohol dehydrogenase) in Microplate Spectrophotometer (Multiskan GO ThermoScientific) for 120 minutes through pulsed shaking. Each data is the mean of three independent experiments and the mean \pm SD of all three independent experiments is shown by error bars on the data. The results were interpreted as the percentage heat protection when substrate aggregation without chaperone was taken as 0% protection. The reactions were prepared in 150 μ l assay buffer (50 mM HEPES-KOH, pH 7.6 for CS; 50 mM Sodium phosphate + 100 mM NaCl pH 7.04 for ADH) including fixed weight of substrate (65 μ g/ml CS; 233 μ g/ml ADH) while varying the chaperone quantity.

CHAPTER 3

RESULTS

3.1 Subcloning

The sHSP gene of *T.volcanium* was previously cloned in the cloning vector pDrive, in our lab. In order to transfer the gene from recombinant pDrive/sHSP plasmid to the expression vector, pET23a(+) (Appendix D, Figure 1), the cloning vector containing the gene was first isolated from the recombinant *E.coli* cells. Following isolation, the plasmid was digested with restriction enzyme EcoRI, to make sure of the presence of gene TVN0775. Agarose gel electrophoresis (AGE) of the samples from 7 different plasmid isolations and digestions revealed the existence of the cloned gene (Figure 3.1).

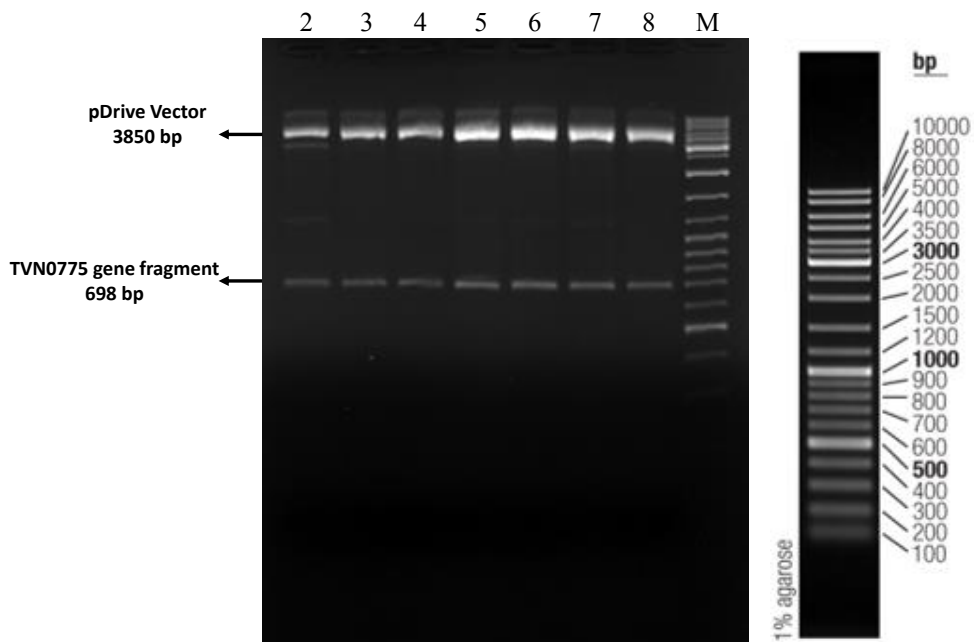


Figure 3.1 Restriction enzyme digestion of several replicates of cloning vector. (Lane 2 – Lane 8) Recombinant pDrive vector, containing TVN0775 gene (by *EcoRI* Restriction Enzyme digestion). The fragment released, containing TVN0775 gene together with some upstream and downstream sequences, is 698bp. 1.5 % agarose gel used to view the samples. M = marker (DNA Ladder Mix from ThermoFisher).

3.1.1 Amplification and purification of the *Tpv* sHSP 14.3 gene

The gene of interest, TVN0775, was amplified by PCR using the pDrive sHSP vector DNA as template and primers with *Nde*I and *Bam*HI restriction sites at the 5' ends. The amplified gene fragment possessed sticky ends which will allow its ligation into the expression vector, pET23a(+), at the intended site and direction. Four replicates of the gene amplification were run and visualized on AGE using 1% agarose prior to purification (Figure 3.2a). The PCR products were then cleaned/purified by gel extraction using the protocol mentioned in QIAGEN QIAquick Gel Extraction Kit Manual and again visualized by AGE on 1% agarose to check the purity (Figure 3.2b).

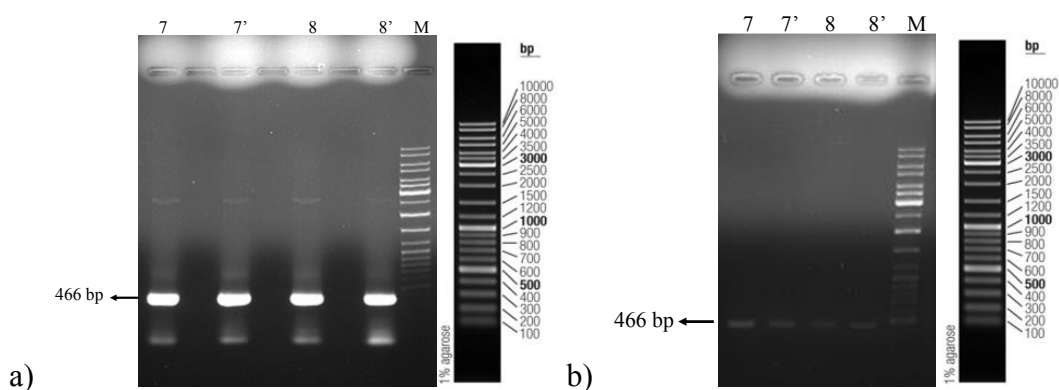


Figure 3.2 Amplified (non-purified) gene (a) and purified gene (b), TVN0775. Amplification done by using primers that possess restriction site for restriction enzymes, *Nde*I and *Bam*HI. Both are viewed on 1% agarose gel. The product in **Lane 7, 7', 8 and 8'** is 466 bp long. M = marker (DNA Ladder Mix from ThermoFisher).

3.1.2 Digestion of the amplified TVN0775 gene and the vector pET23a(+) and purification of the digested products

Along with the expression vector pET23a(+), the purified PCR fragments were double digested with *Nde*I and *Bam*HI and then analyzed by gel electrophoresis, (Figure 3.3). The Figure 3.3 shows that the double digestion as well as the purification procedure was achieved successfully. The expected bands

corresponding to double digested TVN0775 fragment (451 bp) and pET23a(+) vector DNA (3626 bp) were clearly observed on the gel pictures.

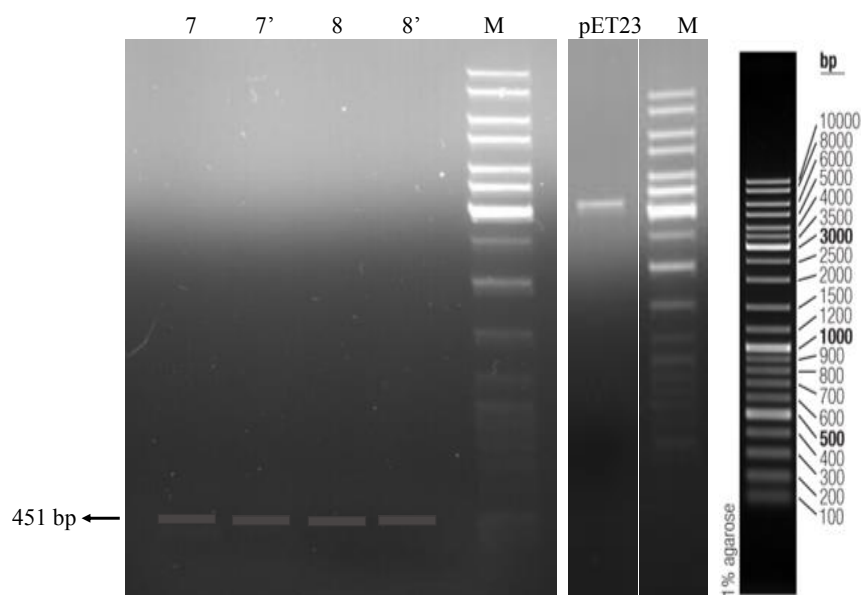


Figure 3.3 Double digested and purified insert, TVN0775, and vector, pET23a(+) on 1% agarose gel. M = marker O'GeneRuler DNA Ladder mix. Lane 7, 7', 8 and 8', all correspond to several replicas of the insert (451 bp) subjected to double digestion and then purification. Lane pET23 corresponds to a band of 3626 bp.

3.1.3 Ligation and transformation

The ligation of double digested TVN0775 insert into the double digested expression vector pET23a(+) was achieved by virtue of sticky end ligation. To this end, 3 ligation mixes were prepared using varied concentrations of the pET23a(+) vector and insert DNA at 1:4, 1:5 and 1:8 vector:insert ratios. Next, *E.coli* competent cells, BL21(DE3), were transformed with a portion of the ligation mixes, separately. The protocol of transformation was as described in Chapter 2. Transformation efficiency, calculated for each ligation mix is given in the Table 3.1. The highest yield of transformants were obtained at vector/insert ratio of 1:8.

Table 3.1 Transformation Efficiency of various ligation mixes in *E.coli* BL21(DE3) cells plated directly and with serial dilutions.

Ligation Mix and dilution name	Transformation Efficiency cfu/μg DNA
Mix1 Plate 1	6×10^4
Mix1 Plate 2	5×10^4
Mix1 Plate 3	3×10^4
Mix1 Plate 4	4×10^4
Mix1 Plate 5	5×10^4
Mix2 Plate 1	3×10^2
Mix2 Plate 2	5×10^2
Mix2 Plate 3	-
Mix2 Plate 4	-
Mix2 Plate 5	-
Mix3 Plate 1	1×10^4
Mix3 Plate 2	2×10^4
Mix3 Plate 3	3×10^4
Mix3 Plate 4	4×10^4
Mix3 Plate 5	-

Randomly selected 24 putative recombinant colonies were screened by restriction enzyme (*NdeI* and *BamHI* combination) digestion of their plasmid DNA followed by agarose gel electrophoresis. Out of 24 colonies, seven were found to be recombinant. After double digestion, a DNA segment of 451 bp corresponding to cloned *Tpv* sHSP gene was excised from the recombinant plasmids (Figure 3.4). The picture of non-recombinants plasmid colonies is shown in Appendix D, Figure 2. Most of the recombinant colonies were obtained with transformation using Mix1.

3.1.4 Verification of the cloning in pET23a(+) expression vector

The recombinant colonies were further characterized by cutting with combinations of different restriction enzymes. The plasmid samples 6 and 7 (of Figure 3.4) were digested with *HindIII/XbaI*, *EcoRI/XbaI*, *NdeI/BamHI*, *HindIII/BglII* restriction enzymes. Double digestion of the plasmid 6 with *NdeI/BamHI* and *HindIII/XbaI* released insert fragment containing the cloned gene sequence at correct positions of 451 bp and 514 bp, respectively (Figure 3.5). However, plasmid sample 7 proved to

be an unsuccessful recombinant, based on unexpected bands after (see Appendix D, Figure 2), double digestion with *EcoRI/XbaI* and *HindIII/BglII* restriction enzyme combinations.

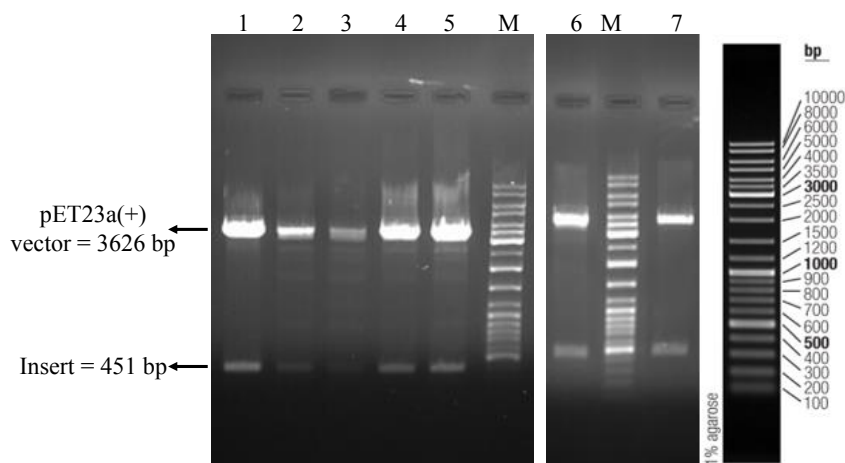


Figure 3.4 Isolated plasmids of successful clones after double digestion by restriction enzymes *NdeI* and *BamHI*. Lane, 1 – 7, the upper band represents the double digested vector pET23a(+) of 3626 bp, while the corresponding lower band show the cloned gene of 451 bp. M = marker O’GeneRuler DNA Ladder Mix.

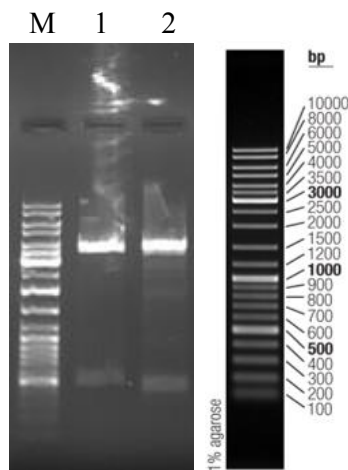


Figure 3.5 Characterization of plasmid sample 6 by double digestion with various enzymes. Lane 1, double digestion by *HindIII/XbaI*, resulting in two bands, upper one is excised vector (3563 bp) and lower band is TVN0775 gene containing insert (514 bp). Lane 2, double digestion by *NdeI/BamHI*, where upper band is excised vector (3626 bp) and lower band is insert of 451 bp. M = marker O’GeneRuler DNA Ladder Mix.

The recombinant plasmid samples 1 to 5 were characterized by double digesting with *HindIII/XbaI*, *EcoRI/XbaI* and single digesting with *SacI* restriction enzymes. All

the five samples were found to be successful recombinants, whereby, their plasmids released the gene fragment, after double digestion, at the expected position (514 bp for *HindIII/XbaI* and 497 bp for *EcoRI/XbaI* restriction enzyme combinations), as will be seen in Figure 3.6. Cutting at unique *SacI* site of the plasmid 5 resulted in linearization of this recombinant plasmid DNA (Figure 3.6). The position for the cut sites of restriction enzymes is mentioned in Table 3.2, while the circular diagram in Figure 3.9 also indicates the multiple cloning site and restriction enzyme digestion site in the pET23a(+) + TVN0775 recombinant plasmid.

Table 3.2 Positions of cut sites of various restriction enzymes in the restriction map of recombinant plasmid pET23a+TVN0775 (4077 bp)

Restriction Enzyme	Restriction site (bp number)
<i>HindIII</i>	173 bp (upstream of gene)
<i>Sall</i>	179 bp (upstream of gene)
<i>SacI</i>	190 bp (upstream of gene)
<i>EcoRI</i>	192 bp (upstream of gene)
<i>BamHI</i>	198 bp (upstream of gene)
<i>NdeI</i>	651 bp (downstream of gene)
<i>XbaI</i>	689 bp (downstream of gene)
<i>BglII</i>	747 bp (downstream of gene)

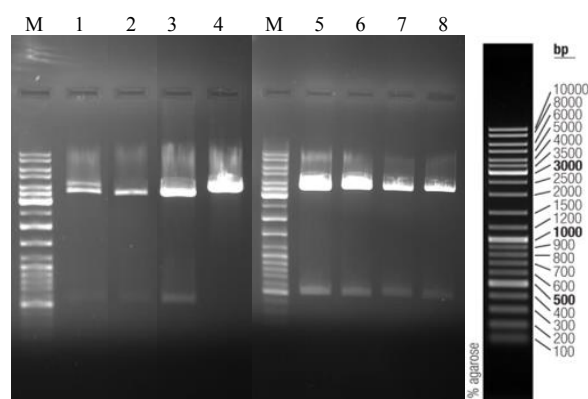


Figure 3.6 Characterization of plasmid samples 1 - 5 by single and double digestion with various enzymes. Sample 2 is represented in **Lane 1** (double digestion by *HindIII/XbaI*); Sample 3 is represented in **Lane 2** (double digestion by *HindIII/XbaI*); Sample 5 is represented in **Lane 3** (double digestion by *HindIII/XbaI*) and **Lane 4** (single digested by *SacI*); Sample 1 is represented in **Lane 5** (double digestion by *HindIII/XbaI*) and **Lane 6** (double digestion by *EcoRI/XbaI*) and Sample 4 is represented in **Lane 7** (double digestion by *HindIII/XbaI*) and **Lane 8** (double digestion by *EcoRI/XbaI*). M = marker O'GeneRuler DNA Ladder Mix.

Out of the six samples successfully identified, two (plasmid sample 1 and 5) were chosen for further verification by sequencing. Recombinant plasmids (sample 1 and 5) with high concentration and purity, single digested with *Hind*III, and after analysis by agarose gel electrophoresis, they were sent for sequencing (Figure 3.7).

The sequencing data (with primers complementary to + and – strand) were compared to the original sequence of pET23a + TVN0775 gene, using online tools, and the result for pET23a/tvsHSP15 recombinant plasmid is shown in Figure 3.8.

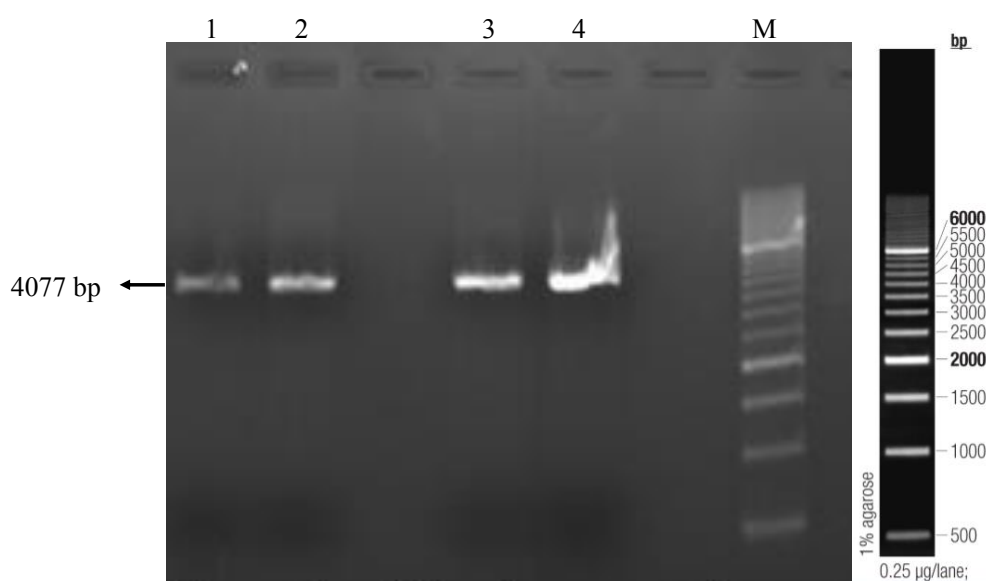


Figure 3.7 Several replicas of recombinant plasmids of pET23a(+) containing the cloned TVN0775 gene. All plasmid samples single digested with *Hind*III restriction enzyme. M = marker O'GeneRuler DNA Ladder 500 bp.

The complete MSA result of whole pET23a(+) with the inserted gene versus the sequence of TVN0775 gene acquired from NCBI is shown in appendix A, Figure 3. The sequencing results for the recombinant plasmids confirmed the cloning of our gene of interest (TVN0775) into pET23a successfully in the correct sequence. A recombinant vector map generated using SnapGene software, displaying the position and the direction of the insert is shown in Figure 3.9.

```

pET23T7Ter      TTTAAGAAGGAGATATACATATGTATACACCCATAAAGTTCTTTACGAATGAGATGATAA
TVN0775         -----ATGTATACACCCATAAAGTTCTTTACGAATGAGATGATAA
pET23T7         TTTAAGAAGGAGATATACATATGTATACACCCATAAAGTTCTTTACGAATGAGATGATAA
                  *****

pET23T7Ter      AAAACGTATCGAATACTGTGAAAGAGGTCTCATCCTTTATATATCCACCAGTCACGTTAT
TVN0775         AAAACGTATCGAATACTGTGAAAGAGGTCTCATCCTTTATATATCCACCAGTCACGTTAT
pET23T7         AAAACGTATCGAATACTGTGAAAGAGGTCTCATCCTTTATATATCCACCAGTCACGTTAT
                  *****

pET23T7Ter      ATCAAGATAGCTCTGATCTGGTATTGGAAGCAGAAATGGCCGGGTTTGACAAGAAAAACA
TVN0775         ATCAAGATAGCTCTGATCTGGTATTGGAAGCAGAAATGGCCGGGTTTGACAAGAAAAACA
pET23T7         ATCAAGATAGCTCTGATCTGGTATTGGAAGCAGAAATGGCCGGGTTTGACAAGAAAAACA
                  *****

pET23T7Ter      TAAAGGTCTCGGTAATAAGAATGTACTACTATAAGTGCGGAGAGAAAGAGAGAATACT
TVN0775         TAAAGGTCTCGGTAATAAGAATGTACTACTATAAGTGCGGAGAGAAAGAGAGAATACT
pET23T7         TAAAGGTCTCGGTAATAAGAATGTACTACTATAAGTGCGGAGAGAAAGAGAGAATACT
                  *****

pET23T7Ter      CTACCGTATATATCGATCAGCGCGTTGACAAAGTGTATAAAGTAGTTAAGCTGCCCGTAG
TVN0775         CTACCGTATATATCGATCAGCGCGTTGACAAAGTGTATAAAGTAGTTAAGCTGCCCGTAG
pET23T7         CTACCGTATATATCGATCAGCGCGTTGACAAAGTGTATAAAGTAGTTAAGCTGCCCGTAG
                  *****

pET23T7Ter      AGATTGAGCAGCAGGACATATCTGCTAAGTATAGTGAAGGCATACTTACAGTTAGAATGA
TVN0775         AGATTGAGCAGCAGGACATATCTGCTAAGTATAGTGAAGGCATACTTACAGTTAGAATGA
pET23T7         AGATTGAGCAGCAGGACATATCTGCTAAGTATAGTGAAGGCATACTTACAGTTAGAATGA
                  *****

pET23T7Ter      AAACCAAGAACATAAAGAACGTAGAAATAGAATAAAAATCATTTTTTTAATAATAATATATA
TVN0775         AAACCAAGAACATAAAGAACGTAGAAATAGAATAA-----
pET23T7         AAACCAAGAACATAAAGAACGTAGAAATAGAATAAAAATCATTTTTTTAATAATAATATATA
                  *****

```

Figure 3.8 MSA of the gene sequence of TVN0775 obtained from National Centre for Biotechnology Information (NCBI) against the sequencing result obtained for pET23a/tvsHSP15.

3.2 Transformation of the cloned gene in NovaBlue cells and making stocks

The recombinant plasmids, which were verified by sequencing (pET23a/TvsHSP15), were isolated from the host *E.coli* BL21(DE3) cells and transformed into NovaBlue cells. NovaBlue cells are convenient hosts for pET vectors due to high transformation efficiency and since they contain no source of T7 RNA polymerase, they are ideal for the establishment of recombinant plasmids under nonexpression conditions. Four colonies were selected randomly for identification of recombinant *E.coli* NovaBlue cells. This was accomplished by double digestion of isolated plasmids with restriction enzymes *Xba*I and *Hind*III to establish the presence of the *tpv* sHSP 14.3 gene, Figure 3.10.

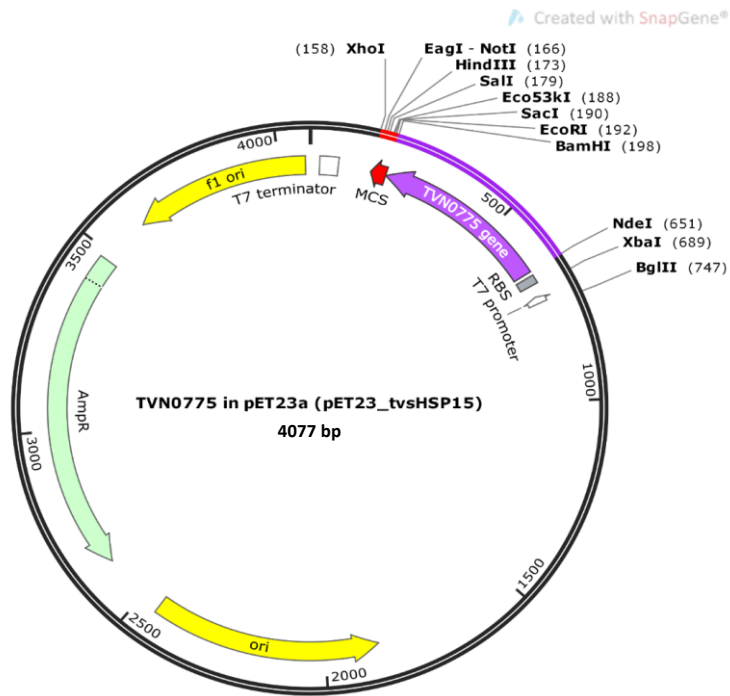


Figure 3.9 Circular view of Recombinant plasmid map of pET23a + TVN0775. Multiple Cloning Site (MCS) lies between *XhoI* (158 bp) and *NdeI* (651 bp). TVN0775 gene is inserted between *NdeI* (651 bp) and *BamHI* (198 bp).

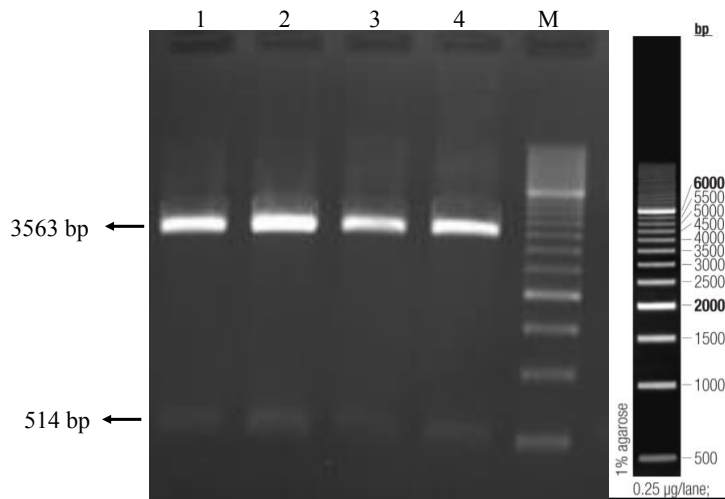


Figure 3.10 Recombinant plasmids isolated from *E.coli* (NovaBlue) cells and double digested with *XbaI* and *HindIII* Restriction enzymes releasing the insert corresponding to 514 bp. M = marker O'GeneRuler DNA Ladder 500 bp.

3.3 Expression of *Tpv* sHSP 14.3 gene

When the cloning of *Tpv* sHSP 14.3 gene into pET23a vector was confirmed, through sequencing results, next procedure was as follows. Expression of the cloned *Tpv* sHSP 14.3 gene using pET23a(+) system was optimized for *E. coli* BL21(DE3) cells. The untransformed BL21(DE3) cells were taken as control and their lysate was also prepared, (Appendix D, Figure 4). The recombinant plasmid (pET23a containing the insert) was already transformed in BL21(DE3), during sub cloning, which is expression cell for pET vectors, therefore, direct expression was done. During optimizing the process of protein expression, two replicates of the pET23_Mix2/15 *Tpv* sHSP cultures were used. This recombinant *E. coli* clone was chosen for downstream experiments. The cultures of the recombinant cells for expression were prepared using the inocula of overnight (O/N) culture and log phase culture. The expression of protein was induced in both cultures, when the OD was 0.6, with IPTG addition for four hours. Fractions of the cell lysates that were subjected to SDS-PAGE analysis (Figure 3.11) indicated that, *Tpv* sHSP 14.3 protein is present in the soluble fraction of the cell lysate.

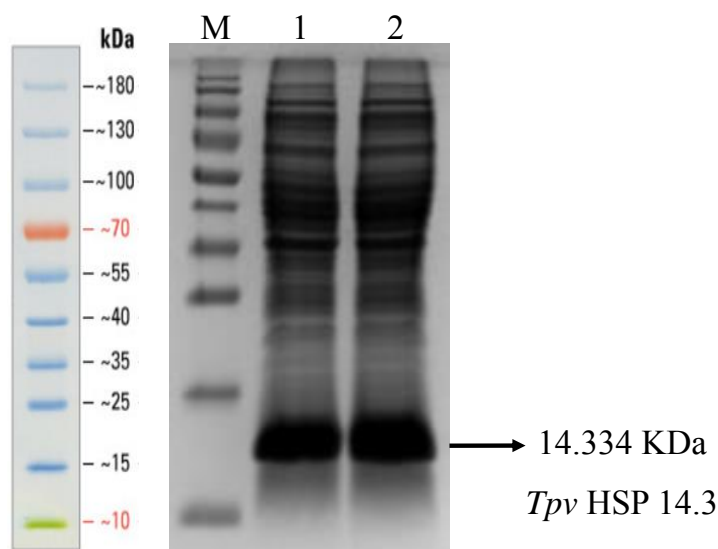


Figure 3.11 *Tpv* sHSP 14.3 (protein of TVN0775) viewed on SDS-PAGE. Lane 1: Inoculum log phase culture, Lane 2: Inoculum o/n culture. M = marker SM Page Ruler Prestained Protein Ladder.

High level expression of the sHSP protein was observed in both the samples as shown by arrow in Figure 3.11. However, lysozyme was used for assistance in cell lysis and lysozyme is also present in the cytoplasmic fraction. In addition, the molecular weight of lysozyme (MW = 14.307 KDa) is almost same as the molecular weight of our protein *Tpv* sHSP 14.3 (MW = 14.333 KDa). In order to get rid of heat sensitive *E.coli* proteins, together with lysozyme, the protein samples were heat treated to 65°C for 15 minutes. The SDS-PAGE showed presence of bands of same thickness at the expected position, (Figure 3.12).

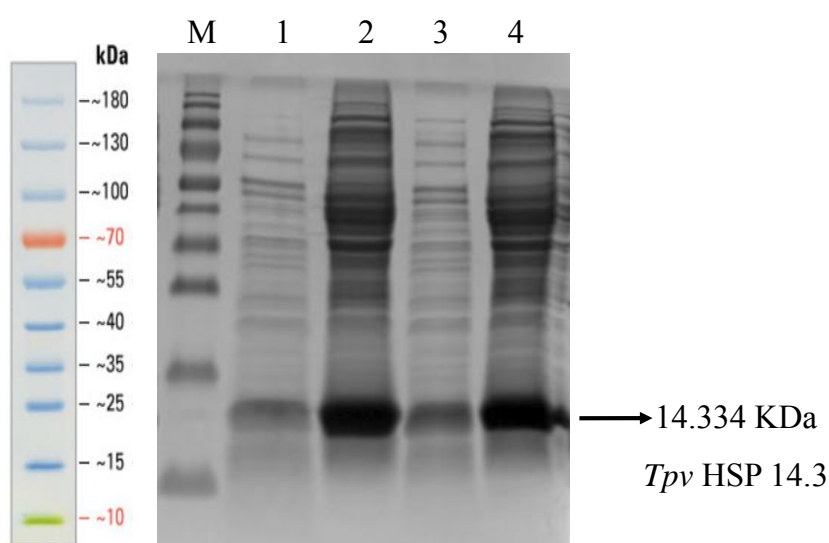


Figure 3.12 *Tpv* sHSP 14.3 (protein of TVN0775) viewed on SDS-PAGE after heat treatment. Lane 1 and 2, After heat and Before heat treatment of the samples prepared from cultures inoculated with log phase inocula, Lane 3 and 4: After heat and Before heat treatment of the samples prepared from cultures inoculated with o/n inocula. M = marker SM Page Ruler Prestained Protein Ladder.

Next, protein expression was performed by skipping the lysozyme, and obtaining the cell lysis solely with sonication. Cell extracts of two recombinant colonies, pET23_Mix2/11 and pET23_Mix2/15, on SDS-PAGE showed high level protein expression, as indicated by arrow in Figure 3.13. The protein samples were subjected to same heat treatment as mentioned above (heating the protein at 65°C for 15 minutes). *Tpv* sHSP 14.3 protein was found to be stable enough according to its band thickness, as seen in Figure 3.13. The *Tpv* sHSP 14.3 WT from pET23_Mix2/15 was also heat treated at 70°C and 80°C for 30 minutes and was found to be heat stable.

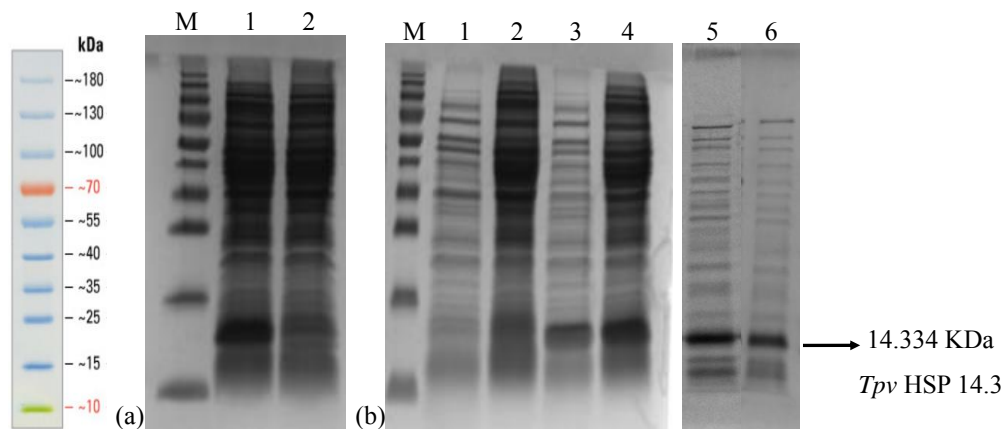


Figure 3.13 *Tpv* sHSP 14.3 (protein of TVN0775), viewed on SDS-PAGE. The cell lysis was achieved only with sonication. The protein is expressed in both the samples and the band indicated by arrow shows the protein is stable after heat treatment at 65°C. (a) **Lane 1** is pET23_Mix2/15 and **Lane 2** is pET23_Mix2/11. (b) **Lane 1** and **2** is after heat (65°C) and before heat of pET23_Mix2/11; **Lane 3** and **4** is after heat (65°C) and before heat of pET23_Mix2/15. **Lane 5** is after heat (70°C) and Lane 6 is after heat (80°C) of pET23_Mix2/15. M = marker SM Page Ruler Prestained Protein Ladder.

The plasmid for protein *Tpv* sHSP 14.3 WT was also transformed into the T7-Express competent cells and their cell lysates were also prepared, Appendix D, Figure 4. However, the yield of the protein was very low as compared to the lysate of BL21(DE3) cells. Therefore, it was not used for downstream experiments.

3.4 Selection of residues for mutation

For this, first BLAST was performed to shortlist the species sharing most similarities in their sHSP protein with *Tpv* sHSP 14.3. BLAST was done separately for the three kingdoms, archaea, eubacteria and eukaryotes. The top 8-17 organisms were chosen and their sHSPs' N-terminal domain (NTD) alone as well as full sHSPs were compared to that of *Tpv* sHSP 14.3. The MSA results based on alignment of NTD of the species' sHSPs having highest similarities with *Tpv* sHSP 14.3 are shown in Figures 3.14, 3.15 and 3.16.

```

WP_010917033.1  ---MYTPIKFFFTNEMIKNVSNTVKEVSSFIYPPVT 32
WP_010901274.1  ---MYTPVRFFTNEMLKNVSNTVKEMSSFIYPPIT 32
EQB69930.1      MTTIYGPLKFMADDFMKNVNDRAKEVLTLYLPPVR 35
EQB71208.1      MINMYNPLKYFTGEVAKNLSRAKEIMTFLYPPVT 35
EQB67620.1      MASLYGPVKYFTDEMKNANERAKEILTFLYPAIT 35
EQB71751.1      ---MYRPLKYYSDEFIKNINRAKEIMTFMYPVPT 32
KJR78384.1      -----MTSAIKNELIKRTEELSRGFYEATYPPV- 28
WP_054845465.1  -----MMEPIKKELSKRIEEASREFYENVLPPM- 28
WP_048197442.1  ---MGIWVYMAKEVAKEIGNKSREFYEFVLPPV- 30
WP_012718226.1  -----MMNVIMREIGKLDLSELYESVFPPI- 28
WP_007551556.1  ---MGLVKSMAKEMIKEIGNKSREFYEFVLPPV- 30
WP_010189564.1  ---MGLVKSMAKEMIKEIGNKSREFYEFVLPPV- 30
WP_054838418.1  -----MMETVKKKESKRLLEELSKEFYENVIPPL- 28
AFU59559.1      ---MIGGGYVARSIAKELDNRSREFYEFVMPAI- 30
WP_009989320.1  -----MMNVIMREIGKLDLSELYESVFPPI- 28
WP_048118634.1  ---MGLVKSMAKEMIKEIGNKSREFYEFVLPPV- 30
WP_042684269.1  ---MGIWRYMAGEMMREIGNKSREFYELVLPV- 30
BAB66734.1      -----MYYLKGKELQKRSSEELSRGFYELVYPPV- 27
.. : . : : . * :

```

Figure 3.14 MSA of (NTD) of sHSP of *T.volcanium* with sHSPs of top (17) Archaea from BLAST. WP_010917033.1 *Thermoplasma volcanium*, WP_010901274.1 *Thermoplasma acidophilum*, EQB69930.1 *Thermoplasmatales archaeon Gpl*, EQB67620.1 *Thermoplasmatales archaeon E-plasma*, EQB71208.1 *Thermoplasmatales archaeon A-plasma*, EQB71751.1 *Ferroplasma sp. Type II*, AFU59559.1 *Candidatus Nitrososphaera gargensis Ga9.2*, WP_042684269.1 *Candidatus Nitrosotenuis chungbukensis*, WP_048197442.1 *Thaumarchaeota archaeon N4*, WP_048118634.1 *Candidatus Nitrosopumilus adriaticus*, WP_007551556.1 *Candidatus Nitrosoarchaeum koreensis*, WP_010189564.1 *Candidatus Nitrosoarchaeum limnia*, WP_012718226.1 *Sulfolobus islandicus*, WP_009989320.1 *Sulfolobus solfataricus*, WP_054845465.1 *Sulfolobus sp. JCM 16833*, WP_054838418.1 *Sulfolobus metallicus*, KJR78384.1 *Sulfolobales archaeon AZ1*, BAB66734.1 *Sulfolobus tokodaii str. 7*. Color code (Red = small + hydrophobic, Blue = acidic, Magenta = basic and Green = hydroxyl + sulfhydroxyl + amine). Symbol code (Asterisk * = fully conserved residue, Colon : = conservation between groups of strongly similar properties, Period . = conservation between groups of weakly similar properties).

```

WP_010917033.1  ---MYTPIKFFFTNE-----MIKNVSN-----TVKEV-----SSFIYPPVT 32
ABC82633.1      ---MAKPMKHHDHGKPKERGRLAALRESMPALPPWIERID-ELVGER---WSAFWPVLR L GEELAFKVPV- 64
KUK22688.1      -----MLLGRRRED--IFRPFRELQREIDRLFDFFRTEVR---PV---KEFFAPDM----- 43
WP_011943168.1  -----MLLGRRRED--IFRPFRELQREIDRLFDFFRTEVR---PA---KEFFAPDM----- 43
WP_008193529.1  -----MLLGRRRED--IFRPFRELQREIDRLFDFFRTEVR---PA---KEFFAPEM----- 43
WP_031002291.1  -----MTGHMVER--RHSLFPDFNDFNREFPGL-----PGWRPATA AHSIPV----- 41
AKJ63312.1      -----MFWLDYRD--PFDSMERFRDDMNRLLSGY----- 27
WP_061079853.1  -----MAPKKFHN--PFHGVDMITEMNRISDSMSSMETGNAGERERGHADAWSPPT- 50
WP_017566186.1  -----MAPKKFHN--PFHGVDMITEMNRISDSMSSMETGNAGERERGHADAWSPPT- 50
KPK45640.1      -----MEISLVP--RRGELSTFRREMDRLDFRFFEGWPFKPSQ---EGPWAPS- 46
CUS97547.1      -----MLMRY- --PFKEIEMLEREINRFNDFRFGFES---GY---EY---PLI- 38
WP_014271265.1  -----MKYYVTY-----NHQNPVSNFESLFDIWSDWGV-----SSSKIPPV- 37
WP_017604807.1  -----MAPRKFN--PFHGVDMITEMNRMSDLSLSSFETSQAGERERGFSDAWSPPT- 50
WP_024977322.1  -----MSDLFFGTD--LLGEFDRLQRQMATLFAGAPASLRA---TR---IGTFPPV- 43
WP_027369409.1  MLRKYLPELRKRSQ---EVQRPNMFDNMFESFFNEPFTRAF-----GEFSYPOV- 47

```

Figure 3.15 MSA of (NTD) of sHSP of *T.volcanium* with sHSPs of top (14) eubacteria species from BLAST. WP_010917033.1 *Thermoplasma volcanium*, ABC82633.1 *Anaeromyxobacter dehalogenans*, KUK22688.1 *Thermotoga naphthophila*, WP_011943168.1 *Thermotoga petrophila*, WP_008193529.1 MULTISPECIES *Thermotoga*, WP_031002291.1 *Streptomyces sp. NRRL F-5727*, AKJ63312.1 *Kiritimatiella glycovorans*, WP_061079853.1 *Nocardiopsis dassonvillei*, WP_017566186.1 *Nocardiopsis synnemataformans*, KPK45640.1 *Nitrospira bacterium*, CUS97547.1 *Candidatus Chrysopegis kryptomonas*, WP_014271265.1 *Sphaerochaeta pleomorpha*, WP_017604807.1 *Nocardiopsis alkaliphila*, WP_024977322.1 *Ralstonia pickettii*, WP_027369409.1 *Desulfovermiculus halophilus*.

```

WP_010917033.1 -----MYTPIKFF----TNEMIKNVSN-TVK-----E---VSSFIYPP VT 32
KNZ81133.1      -MSSLF---YYEPFYHF EKLLDDAFSTRVAD-NDK---QIQCCGQDAG---VTGFLRPR M- 49
KXN92576.1      -MSRVF---FYDPFYEF DLF EHALRPWAQF-SEA---TAQLERSNAV---VSRTFKPR M- 49
KNZ81132.1      -MSSLL---YYEPFYQF EKLLDDAFSTRMAE-NDK---QLQRRGQDAG---VTGFLKPR M- 49
XP_014527854.1 -MSYL---SFYSPLLR E-----LGQ-GSVHLARTMPVWGSDVR---VDHGFGDA -- 41
XP_007340101.1 --MSLFWTATPSPFADF ERAFNGVFNGNLGR-----IGDSSNDT---TFQGFRRPR M- 46
XP_007773645.1 MSLTLG--VYDPFSELERFFDDAFLTRFSG-GNA---NANRE---VA---ARQPFRPK M- 47
KYR02364.1      -MSCKVFYNRPNPFYFN---HIYNELVKQ-QQQQL-----AN---QDRFYTPN S- 41
XP_642604.1     -MATIF-----DILNLTNNNNNNKNYFESCKRQRTNKNKTKIIDI LPP M- 44

```

Figure 3.16 MSA of (NTD) of sHSP of *T.volcanium* with sHSPs of top (8) eukaryotes from BLAST. **WP_010917033.1** *Thermoplasma volcanium*, **KNZ81133.1** *Termitomyces* sp. J132, **KXN92576.1** *Leucoagaricus* sp. SymC.cos, **KNZ81132.1** *Termitomyces* sp. J132, **XP_014527854.1** *Blastocystis* sp. subtype 4, **XP_007340101.1** *Auricularia subglabra*, **XP_007773645.1** *Coniophora puteana*, **KYR02364.1** *Dictyostelium lacteum*, **XP_642604.1** *Dictyostelium discoideum*.

BLAST with archaea showed conservation of many residues/amino acids, (Figure 3.14 and Table 3.3). However, BLAST with bacteria species showed relatively less number of conserved residues, (Figure 3.15 and Table 3.2). When BLAST was performed for eukaryotes, the top 50 organisms mostly belonged to the Fungi domain. MSA of *Tpv* sHSP 14.3 with eukaryotic sHSPs, showed more conservation in residues as compared to MSA of *Tpv* sHSP 14.3 with sHSPs of eubacteria, (Figure 3.16 and Table 3.3). This may indicate a close evolutionary link between eukaryotic and archaeal sHSPs, than eubacterial sHSPs. The common feature of these residues regardless of the source of sHSPs is hydrophobicity.

There are several studies conducted on N-terminal domain of sHSPs from eukaryotes, in order to identify amino acids responsible in substrate binding and the role it plays in prevention of substrate aggregation during thermal stress (Horwitz *et al.*, 1998; Fu *et al.*, 2005). In some of these studies, complete NTD truncation resulted in drastic loss of substrate protection of sHSPs against thermal aggregation (Leroux *et al.*, 1997). In a number of studies, by point mutation/deletion amino acids in the NTD of the respective sHSPs, that might be responsible for substrate binding were determined (Heirbaut *et al.*, 2017; Shatov *et al.*, 2018). These amino acids and their anticipated functions are mentioned in Table 3.4. Many of these residues were also suggested to be critical for substrate interaction in three dimensional structure analysis.

Table 3.3 List of conserved residues in the NTD of *Thermoplasma volcanium* sHSP 14.3, after running MSA with most similar sHSPs from BLAST.

Organism	Comparison Kingdom	Conserved residues when NTDs is compared	Conserved residues when full sHSP is compared
<i>Thermoplasma volcanium</i>	Archaea	P4, I5, F7, F8, E11, M12, I13, V16, V20, V23, F26, I27, P29, P30, V31	I5, F8, E11, M12, I13, K14, V16, V20, E22, V23, F26, I27, P29, P30, V31
	Eubacteria	P4, I5, F7, I13, V16, F26, P29, P30, V31	I5, F8, M12, I13, V16, F26, I27, P29, P30, V31
	Eukaryotes	P4, I5, F8, M12, V16, I27, V31	P4, I5, F8, M12, I13, I27, P29, V31

Among the highly conserved NTD residues in archaeal sHSPs based on MSA results and those residues that are reported to have possible roles in substrate recognition and/or binding in the literature, we have selected eleven of them for site specific mutagenesis; I5, F7, F8, E11, M12, E22, V23, F26, I27, V31 and L33. The L33 does not belong to the N-terminal domain, rather, it is the first residue of alpha crystallin domain (ACD) of *Tpv* sHSP 14.3. This hydrophobic residue was seen highly conserved among all the three domains, being only substituted by similar (hydrophobic) residues (I or V). The chosen residues, all have equivalents in plants and mammals, (see Table 3.5). The position of these points in the primary structure of the *Tpv* sHSP 14.3 is shown in Figure 3.17.

For mutagenesis, two step procedure was planned. In the first step, the hydrophobic residues were replaced with hydrophilic residues, and in second step, the charged residues were replaced by uncharged hydrophobic residues.

Table 3.4 Important residues in different sHSPs and their functions

Organism/Protein Name	Important amino acids	Function
Pea sHSP18.1 (Jaya, Garcia and Vierling, 2009)	F7	Substrate binding (with malate dehydrogenase)
	F8	
	F16	
	F19	
	F30	
	F32	
	K72	Substrate binding (with Luciferase)
	K77	
	E79	
	D82	
	D83	
	Q87	
	I146	
	Wheat (Van Montfort <i>et al.</i> , 2001)	
F10		
W48		
F110		
Mouse/ α B-crystallin (Horwitz <i>et al.</i> , 1998)	F27	
Human/ α B-crystallin (Horwitz <i>et al.</i> , 1998; Shroff <i>et al.</i> , 2001)	F24	
	F27	
	F28	
	P46	
Human/HSPB6 (Heirbaut <i>et al.</i> , 2014)	M68	
	Q31	
	F33	

Table 3.5 Equivalent residues for the amino acids chosen for site specific mutagenesis in the NTD of *Tpv* sHSP 14.3.

Residues chosen for site specific mutagenesis	Equivalent residues in other archaea species	Equivalent residues in other Eukaryotic species	Equivalent residues in other Eubacteria species
I5	V, L, G, M, T	F, L	M, S, E
F7	F, Y, S, V, P, T, A	F, H, E, Q, R, D	H, L, G, W, P, I, M, Y
F8	F, M, Y, V, L, I	F, E, L, N	D, G, H, L, K, S, Y, R
E11	E, S	A, V, I	K, E, R, H, V, Y, N, T, Q
M12	M, F, V, I, L	F, L, Y	R, P, L, S, G, E, N, R
E22	E, G	G, V, R, T, A, N, K	V, T, P, E, G, T, R
V23	M, V, I, F	V, T, A, Q, T	R, G, F, S, V, A
F26	F, Y, S, N, A, L	F, T, G, P, D	F, V, S
I27	I, L, M, V	F, L, Y, I	K, S
V31	I, V, L	M, S	V


```

          I5   F7   F8           E11 M12
atgtatacaccataaaagttctttacgaatgagatgataaaaaacgtatcgaatactgtg
M   Y   T   P   I   K   F   F   T   N   E   M   I   K   N   V   S   N   T   V

          E22 V23           F26 I27           V31   L33
aaagagggtctcatcctttatatatccaccagtcacgttatatcaagatagctctgatctg
K   E   V   S   S   F   I   Y   P   P   V   T   L   Y   Q   D   S   S   D   L
gtattggaagcagaaatggccgggtttgacaagaaaaacataaagggtctcggtaaataag
V   L   E   A   E   M   A   G   F   D   K   K   N   I   K   V   S   V   N   K
aatgtactcactataagtgcggagagaaagagagaataactctaccgtatataatcgatcag
N   V   L   T   I   S   A   E   R   K   R   E   Y   S   T   V   Y   I   D   Q
cgcggttgacaaagtgtataaagtagttaagctgcccgtagagattgagcagcaggacata
R   V   D   K   V   Y   K   V   V   K   L   P   V   E   I   E   Q   Q   D   I
tctgctaagtatagtgaaagcatacttacagttagaatgaaaaccaagaacataaagaac
S   A   K   Y   S   E   G   I   L   T   V   R   M   K   T   K   N   I   K   N
gtagaaatagaataa
V   E   I   E   -

```

Figure 3.17 The blue labelled part is the N-terminal region of the sHSP. Positions labelled in red are those which are highly conserved and are chosen for single, double and triple mutations.

3.5 Mutagenesis

Recombinant pET23_tvshSP15 plasmid DNA was employed as the source of template sHSP gene in the mutagenesis experiments, (Figure 3.18).

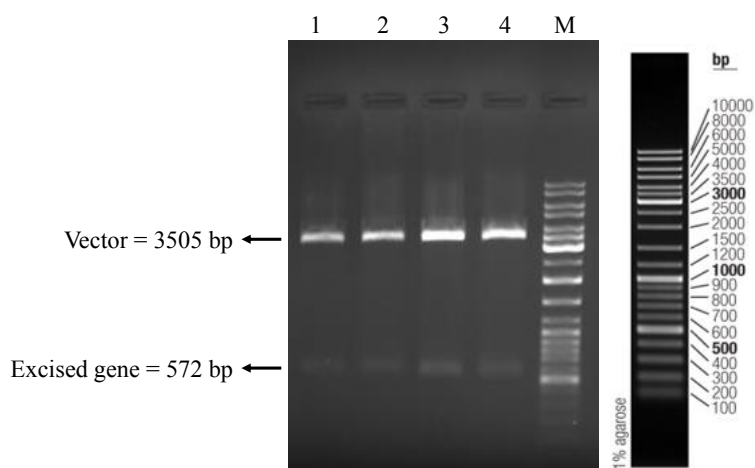


Figure 3.18 Recombinant plasmid pET23_tvshSP15 isolated from *E.coli* (NovaBlue) cells and double digested with *Xba*I and *Hind*III Restriction enzymes.

Single-site, double-site and triple-site mutants were generated using Agilent QuikChange method as described in the Material and Methods. The transformation efficiency calculated for each one of the mutations is shown in Table 3.6.

Table 3.6 Transformation Efficiency of mutagenesis with various samples

Mutation Name	Transformation Efficiency
I5T	1×10^6
F8Y	5×10^5
E11V	3×10^6
M12T	2×10^6
E22G	4×10^6
V31G	2×10^4
I5TF8Y	1×10^5
F7SF8Y	2×10^5
F26YI27T	2×10^5
V31IL33I	1×10^4
F8YV31G	5×10^6
I5TF7SF8Y	4×10^5
V23GF26YI27T	2×10^4

Putative mutants were randomly picked up among transformants, to isolate the plasmid and send for sequencing to verify the intended mutations. The isolated plasmids were double digested with restriction enzymes, *Bgl*II and *Hind*III to check for the release of the gene at a position corresponding to 572 bp (Figures 3.19 – 3.21). The isolated plasmids with the highest concentrations and purity were sent for sequencing.

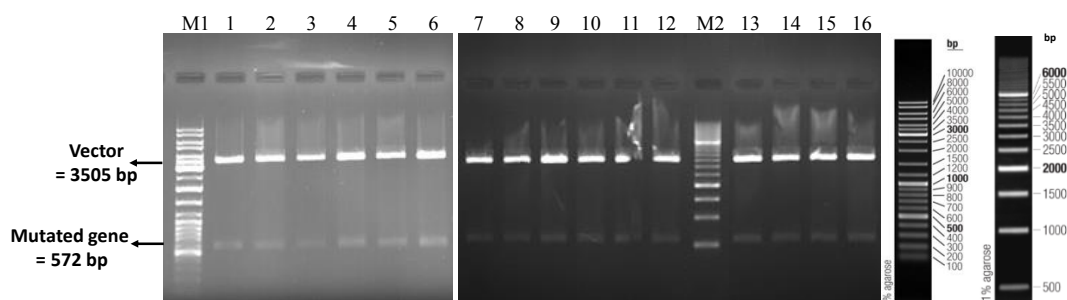


Figure 3.19 Double digestion of mutant plasmid with *Hind*III/*Bgl*II restriction enzymes. Lane 1 – 3: possible F8Y, Lane 4 – 6: possible I5TF8Y, Lane 7 – 9: possible I5T, Lane 10 – 12: possible M12T, Lane 13 – 15: possible I5TTF7SF8Y and Lane 16: possible V23GF26YI27T mutant gene (572 bp) and vector DNA (3505 bp). M1 = marker O’GeneRuler DNA Ladder Mix and M2 = marker O’GeneRuler DNA Ladder 500 bp.

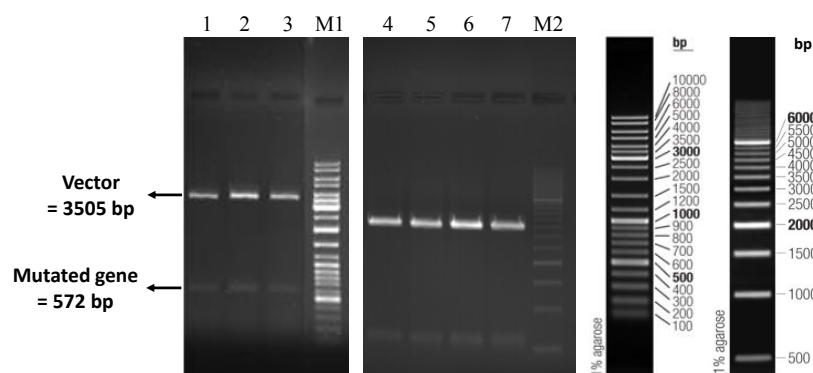


Figure 3.20 Double digestion of mutant plasmid with *HindIII/BglIII* restriction enzymes. Lane 1 – 3: possible F8YV31G, Lane 4 – 6: possible V31G and Lane 7: possible V31IL33I mutant gene (572 bp) and vector DNA (3505 bp). M1 = marker O’GeneRuler DNA Ladder Mix and M2 = marker O’GeneRuler DNA Ladder 500 bp.

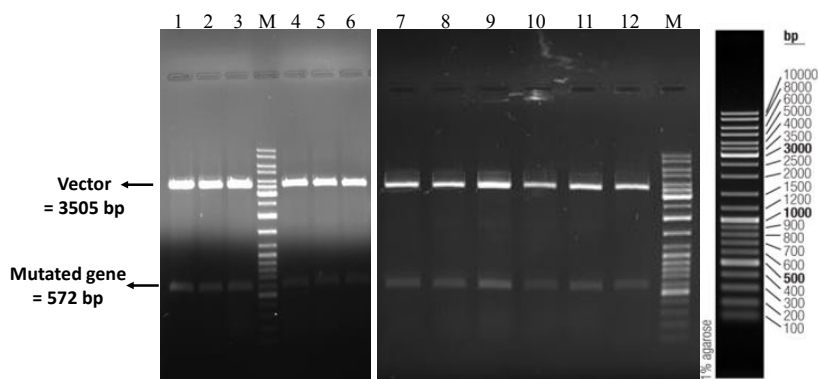


Figure 3.21 Double digestion of mutant plasmid with *HindIII/BglIII* restriction enzymes. Lane 1 – 3: possible E11V, Lane 4 – 6: possible E22G, Lane 7 – 9: possible F7SF8Y, Lane 10 – 12: possible F26YI27T mutant gene (572 bp) and vector DNA (3505 bp). M = marker O’GeneRuler DNA Ladder Mix.

3.5.1 Analysis of Sequencing Results for Mutagenesis

For each mutation, a total of three samples/colonies were sent for sequencing, and each sample was sequenced by forward as well as reverse primer. The results obtained were analyzed thoroughly by comparing each result with the nucleotide sequence of *Tpv* sHSP 14.3 protein, *i.e.*, TVN0775, that we have determined previously in our lab and is also available online at National Centre of Biotechnology Information (NCBI), by pairwise alignment and MSA using Clustal Omega online tool, keeping the intended mutation into consideration.

In addition to this, the chromatogram of sequencing was also comprehensively studied for overlapping peaks to eliminate any chances of computer error. The chromatograms of all the mutants, as well as their comparison with the WT is shown in Figure 3.22 – 3.29. For those samples, who's either the forward or reverse primer had some errors, sequencing was repeated. A summary of the analysis of sequencing results of 3 samples for each one of the mutations is given in Table 3.7.

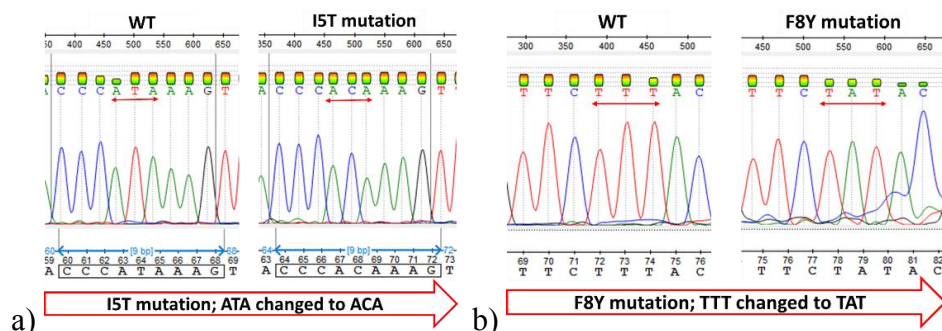


Figure 3.22 Chromatogram images of the mutants (a) I5T and (b) F8Y. The mutated codon is highlighted with a red line.

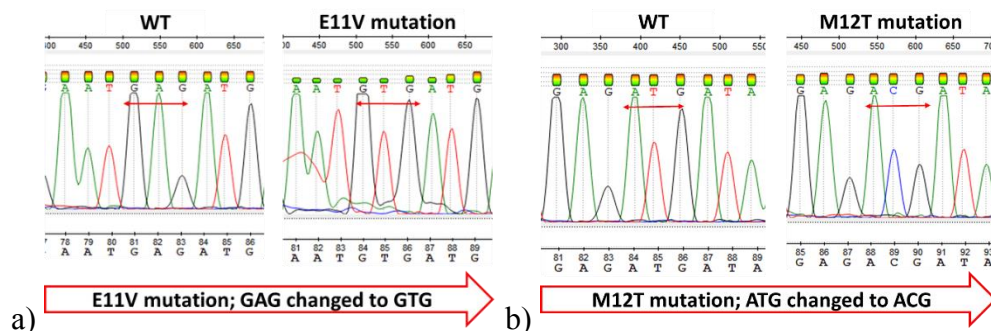


Figure 3.23 Chromatogram images of the mutants (a) E11V and (b) M12T. The mutated codon is highlighted with a red line.

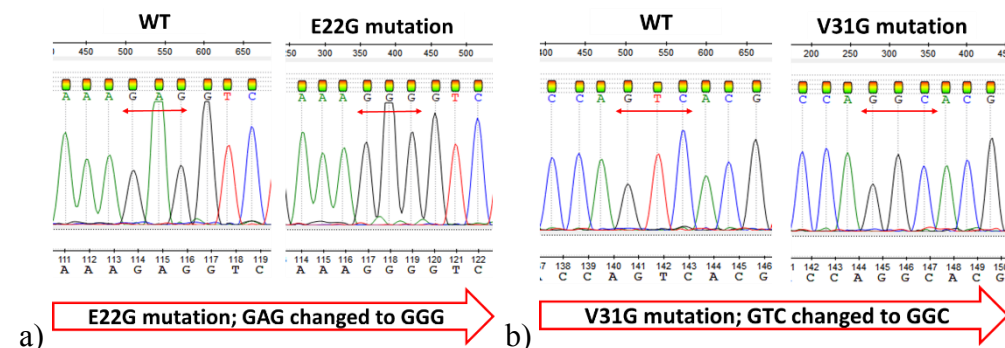


Figure 3.24 Chromatogram images of the mutants (a) E22G and (b) V31G. The mutated codon is highlighted with a red line.

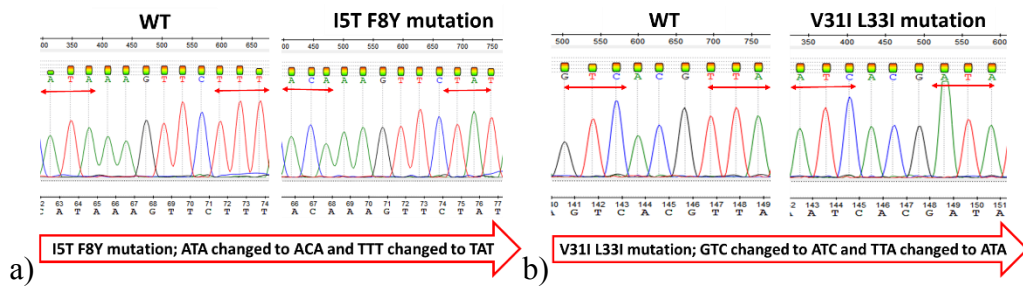


Figure 3.25 Chromatogram images of the mutants (a) I5TF8Y and (b) V31IL33I. The mutated codon is highlighted with a red line.

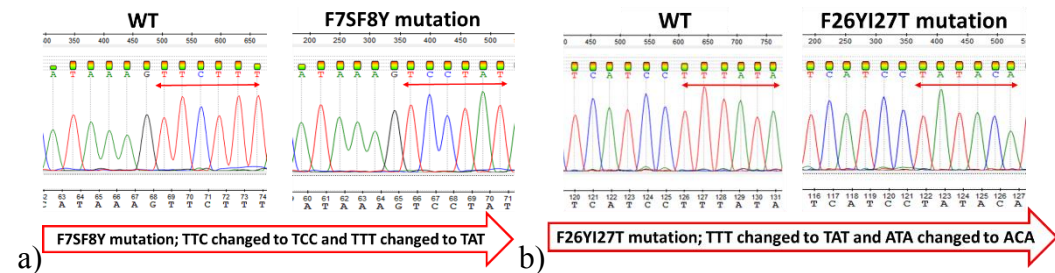


Figure 3.26 Chromatogram images of the mutants (a) F7SF8Y and (b) F26YI27T. The mutated codon is highlighted with a red line.

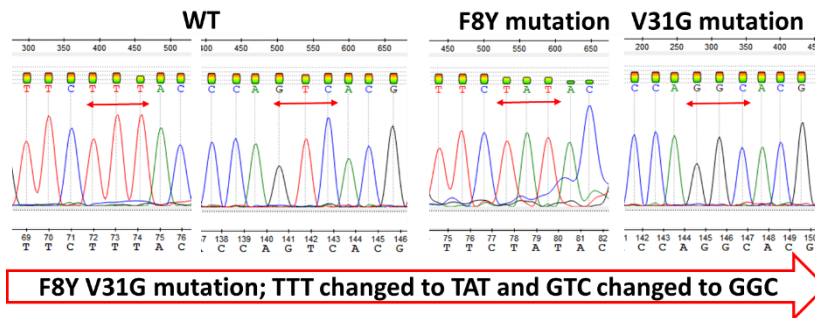


Figure 3.27 Chromatogram images of the mutant F8YV31G. The mutated codon is highlighted with a red line.

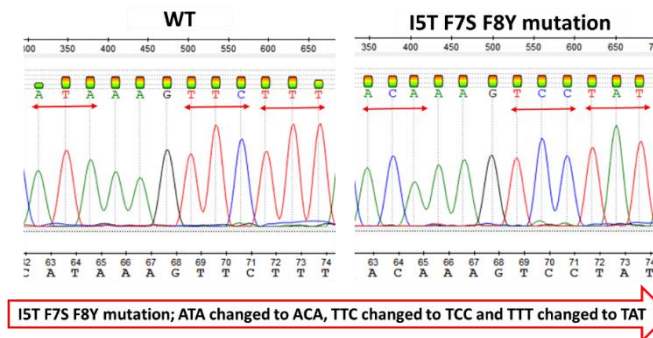


Figure 3.28 Chromatogram image of the mutant I5TF7SF8Y. The mutated codon is highlighted with a red line.

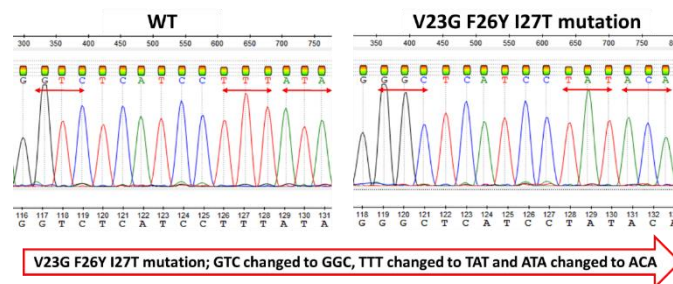


Figure 3.29 Chromatogram image of the mutant V23GF26YI27T. The mutated codon is highlighted with a red line.

Table 3.7 Summary of alignment results of mutation samples (present in XL1-Blue and DH5 α competent cells) after re-sequencing.

Mutation	Sample name	Forward primer result	Reverse Primer Result
I5T	1-1-2/1*	100 % as expected	100 % as expected
	1-1-2/2	100 % as expected	100 % as expected
	1-1-2/3	100 % as expected	100 % as expected
F8Y	1-3-2/1	100 % as expected	100 % as expected
	1-3-2/2*	100 % as expected	100 % as expected
	1-3-2/3	100 % as expected	100 % as expected
E11V	5-2/1	100 % as expected	100 % as expected
	5-2/2*	100 % as expected	100 % as expected
	5-2/3	100 % as expected	100 % as expected
M12T	2-1-4/1	100 % as expected	100 % as expected
	2-1-4/2	100 % as expected	100 % as expected
	2-1-4/3*	100 % as expected	100 % as expected
E22G	2-3/1	100 % as expected	100 % as expected
	2-3/2	100 % as expected	100 % as expected
	2-3/3*	100 % as expected	100 % as expected
V31G	4-5-2/1	100 % as expected	100 % as expected
	4-5-2/3	100 % as expected	100 % as expected
	4-5-2/4*	100 % as expected	100 % as expected
I5TF8Y	4-4-3/3	100 % as expected	100 % as expected
	4-4-3/4	No mutation	No mutation
	4-4-3/5*	100 % as expected	100 % as expected
F7SF8Y	4-4/1*	100 % as expected	100 % as expected
	4-4/1	100 % as expected	100 % as expected
	4-4/1	No mutation	No mutation
F26YI27T	2-13/1	100 % as expected	100 % as expected
	2-13/2*	100 % as expected	100 % as expected
	2-13/3	100 % as expected	100 % as expected
F8YV31G	1-2/1*	100 % as expected	100 % as expected
	1-2/2	V31G present only	V31G present only
	1-2/3	F8Y present only	F8Y present only
V31IIL33I	7-5-1/6*	100 % as expected	100 % as expected
I5TF7SF8Y	3-2-1/3	100 % as expected	100 % as expected
	3-2-1/5	100 % as expected	100 % as expected
	3-2-1/7*	100 % as expected	100 % as expected
V23GF26YI27T	4-2-1/8*	100 % as expected	100 % as expected

One sample from each mutation (indicated by asterisk in the Table 3.7) was selected for downstream experiment, *i.e.*, for transformation into expression cells and successive protein production and its structural and functional analysis.

3.6 Transformation of *E.coli* BL21(DE3) and *E.coli* T7-Express cells for expression of WT and mutant *Tp_v* sHSP 14.3 protein

Selected mutated plasmids were double digested with *Bgl*III and *Hind*III and/or *Xba*I and *Hind*III restriction enzyme combinations to release the insert. The mutated gene released from all selected plasmids travelled to a position corresponding to 572 bp after digestion by *Bgl*III/*Hind*III combination and to 514 bp after digestion by *Xba*I/*Hind*III combination of enzymes, as shown in Figure 3.30.

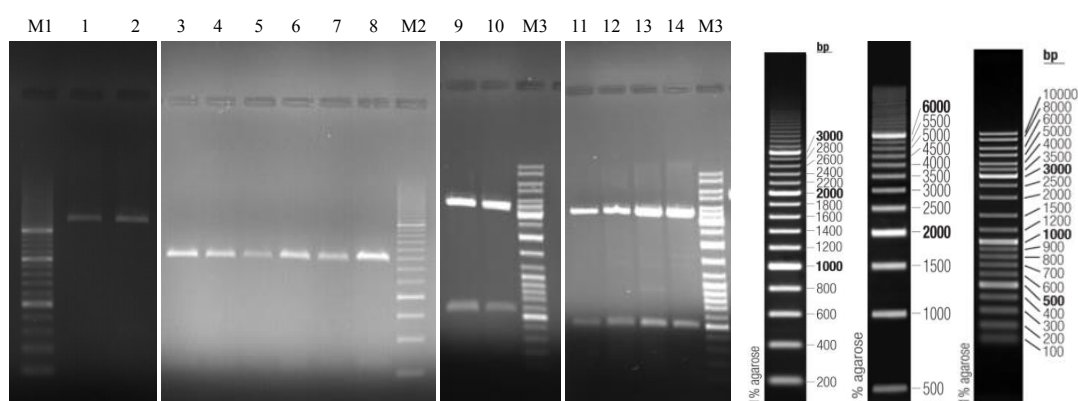


Figure 3.30 Mutated plasmids after double digestion with *Bgl*III and *Hind*III restriction enzymes (Lane 1 – 10) and *Xba*I and *Hind*III restriction enzymes (Lane 11 – 14), as seen on 1% agarose gel. **Lane 1** = I5T; **Lane 2** = F8Y; **Lane 3** = M12T; **Lane 4** = V31G; **Lane 5** = I5T F8Y; **Lane 6** = V31I L33I; **Lane 7** = I5T F7S F8Y; **Lane 8** = V23G F26Y I27T; **Lane 9** = F7SF8Y; **Lane 10** = F26YI27T; **Lane 11** = E11V; **Lane 12** = E22G; **Lane 13** = F8Y V31G, **Lane 14** = V31G. **M1** = marker O’RangeRuler 200bp DNA Ladder, **M2** = O’RangeRuler 500 bp DNA Ladder; **M3** = marker O’GeneRuler DNA Ladder Mix. In each lane, the upper band corresponds to vector (Lane 1 – 10 = 3505bp and Lane 11 – 14 = 3563bp), while the lower band corresponds to released gene (Lane 1 – 10 = 572bp and Lane 11 – 14 = 514bp).

In the pET expression system, *E.coli* BL21(DE3) cells are necessary for expression. Since they contain a chromosomal copy of the T7-RNA polymerase gene and *lac*I

gene (Novy *et al.*, 2001, Studier *et al.*, 1986). Transcription of the T7 RNA polymerase gene is directed by the *lacUV5* promoter, which is inducible by IPTG.

All the mutant plasmids were transformed into competent *E.coli* BL21(DE3) cells either prepared in the laboratory or those purchased from New England Biolabs, (U.S.A.). The transformation efficiency of each mutant is given in Table 3.8.

Table 3.8 Transformation efficiency of mutants in BL21(DE3) competent cells.

Mutant	Transformation efficiency
I5T	6×10^4
F8Y	4×10^3
E11V	5×10^4
M12T	3×10^4
E22G	3×10^5
V31G	1×10^4
I5TF8Y	4×10^3
F7SF8Y	2×10^5
F26YI27T	2×10^5
V31IL33I	4×10^4
F8YV31G	5×10^4
I5TF7SF8Y	3×10^3
V23GF26YI27T	2×10^5

For each mutation, two randomly selected colonies (mentioned in Table 3.9) were chosen for downstream experiments. Plasmids of the selected recombinant colonies were double digested with *Bgl*III and *Hind*III restriction enzymes (producing an insert size of 572 bp) to endorse that the selected colonies carry the gene of interest (Figure 3.31 and 3.32).

Agarose gel electrophoresis results for the plasmids of mutant variants (isolation repeated for some of them) are shown in the Figure 3.31. Insert for the variant V31G was only slightly visible, despite several plasmid isolations from different set of colonies for this specific mutation. Therefore, as a solution, another expression system, T7 Express competent *E.coli* cells were transformed with the plasmid containing this mutation, in the next step.

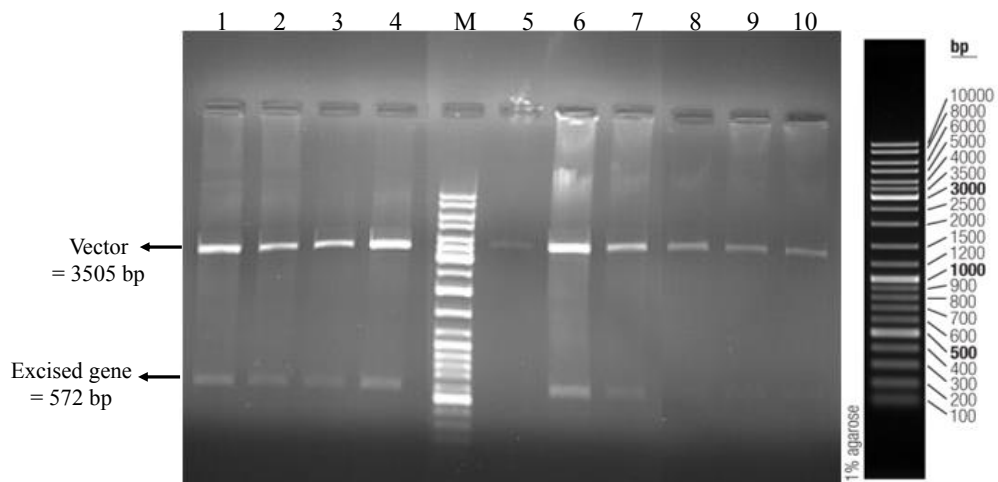


Figure 3.31 Mutated plasmids isolated from *E.coli* BL21(DE3) cells after double digestion with *Bgl*II and *Hind*III restriction enzymes. Lane 1 and 2: I5T variant; Lane 3 and 4: F8Y variant; Lane 5: V31G variant; Lane 6 and 7: I5TF8Y variant; Lane 8: V31IL33I variant; and Lane 9 and 10: I5TF7SF8Y variant. M = marker O'GeneRuler DNA Ladder Mix.

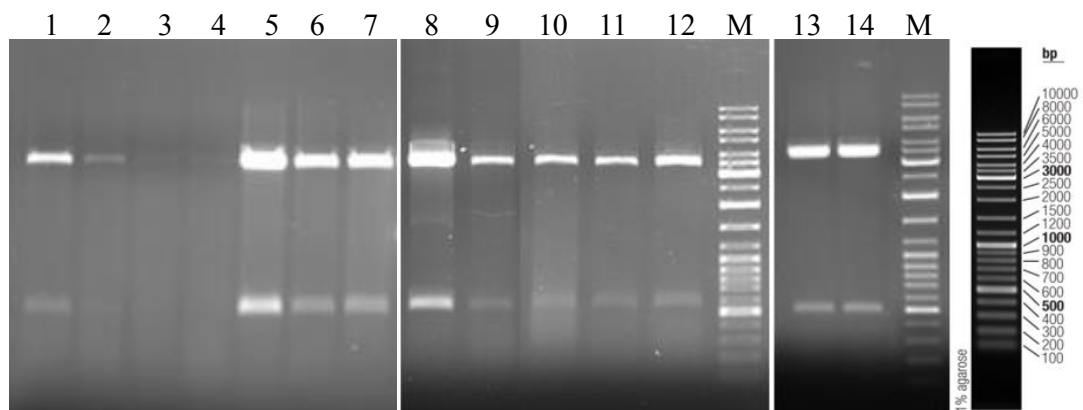


Figure 3.32 Mutated plasmids isolated from *E.coli* BL21(DE3) cells after double digestion with *Bgl*II and *Hind*III (Lane 1 – 12) and *Xba*I and *Hind*III (Lane 13 and 14) restriction enzymes. Lane 1 and 2: M12T variant; Lane 3 and 4: V31G variant; Lane 5 and 6: V31IL33I variant; Lane 7: V23GF26YI27T variant; Lane 8: F7SF8Y variant; Lane 9: F26YI27T variant; Lane 10: E11V variant; Lane 11 and 12: E22G variant and Lane 13 and 14: F8YV31G variant. M = marker O'GeneRuler DNA Ladder Mix. In each lane, the upper band corresponds to vector (Lane 1 – 12 = 3505bp and Lane 13 and 14 = 3563bp), while the lower band corresponds to released gene (Lane 1 – 12 = 572bp and Lane 13 and 14 = 514bp).

Table 3.9 Name given to selected colonies for each mutant

Mutant Name	Selected colony name
I5T	1-1-2/1 2 and 1-1-2/1 4
F8Y	1-3-2/2 6 and 1-3-2/2 3
I5TF8Y	4-4-3/5 4 and 4-4-3/5 7
F7SF8Y	4-4/1 5 and 4-4/1 9
I5TF7SF8Y	3-2-1/7 1 (only one colony obtained)
E11V	5-2/2 6 and 5-2/2 8
M12T	2-1-4/3 4 and 2-1-4/3 8
E22G	2-3/3 4 and 2-3/3 8
V23GF26YI27T	4-2-1/8 2 and 4-2-1/8 4
F26YI27T	2-13/2 5 and 2-13/2 8
V31G	4-5-2/4 1 and 4-5-2/4 3
V31IL33I	7-5-1/6 9 and 7-5-1/6 10
F8YV31G	1-2/1 2 and 1-2/1 6

Besides V31G mutant variant, the plasmids from four more variants generated, (E11V, M12T, E22G and F8YV31G), were also transferred into the highly efficient recombinant protein expressing cell system, *i.e.*, *E.coli* T7-Express competent cells. These cells have the T7 RNA polymerase gene inserted into the lac operon on their chromosome which is expressed under the control of the lac promoter. This configuration provides controlled induction of the polymerase and consequently, inducible control of transcription of genes downstream of the T7 promoter. This system provides potential advantages over BL21(DE3), which carries the T7 RNA polymerase on a lysogenic prophage. Although λ DE3 is normally dormant in the host chromosome, the induction of the SOS cascade can occur as the result of expressing proteins that damage the *E. coli* chromosome, either directly or indirectly. This may lead to cell lysis.

After transformation of the T7-express cells by the selected mutated plasmids, Transformation efficiencies, were calculated and shown in Table 3.10. Transformation efficiencies up to 10^3 folds higher than BL21(DE3) cells, could be achieved with T7-Express cells. Two colonies were randomly selected for each mutant for downstream experiments. The plasmids of these colonies were isolated,

cut by *Xba*I + *Hind*III restriction enzymes to make sure that it carries our gene of interest *i.e.*, releasing an insert of size 514 bp (Figure 3.33).

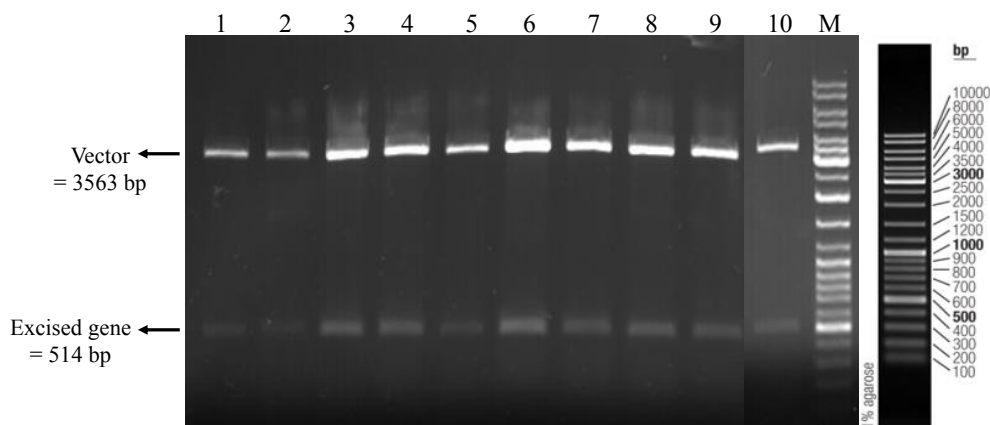


Figure 3.33 Mutated plasmids from transformed successfully in *E.coli* T7-Express competent cells after double digestion with *Xba*I and *Hind*III RE, Lane 1 and 2: M12T variant; Lane 3 and 4: V31G variant; Lane 5 and 6: F8YV31G variant; Lane 7 and 10: E11V variant; Lane 8 and 9; E22G variant. For each of the lanes, the upper band signifies the digested vector (3563 bp) while the lower band denotes the released insert (514 bp).

Table 3.10 Transformation efficiency of mutants in T7 Express competent cells

Mutant	Transformation efficiency
M12T	4×10^6
V31G	5×10^6
E11V	2×10^8
E22G	3×10^8
F8YV31G	4×10^7

Overall, 13 variants that express the different mutant *Tpv* sHSP 14.3 proteins were generated in this study. These were planned and categorized according to the segments of NTD of *Tpv* sHSP 14.3 protein (Figure 3.34).

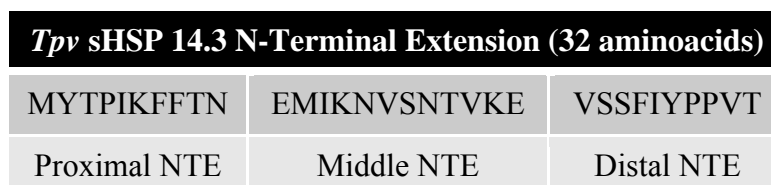


Figure 3.34 Schematic diagram for virtual division (into proximal, middle and distal sections) of N-terminal domain of *Tpv* sHSP 14.3.

All the variants' name and the category they lie in, is mentioned in the Table 3.11.

Table 3.11 List of *Tpv* sHSP 14.3 NTD variants generated in this study.

NTD (32 amino acids)	Proximal NTD (1-10 Residues)	Middle Part (11 – 22 residues)	Distal Part (23 – 32 residues)	Proximal/ Distal
Name of Mutant	I5T	E11V	F26YI27T	F8YV31G
	F8Y	M12T	V23GF26YI27T	
	I5TF8Y	E22G	V31G	
	F7SF8Y		V31IL33I	
	I5TF7SF8Y			

3.7 Expression of *Tpv* sHSP 14.3 N-Terminal Domain Mutants and heat treatment

Following transformation of the *E.coli* expression strains, BL21(DE3) and T7-Express *E.coli* cells, by the recombinant plasmids containing correctly mutated gene, all the variants were grown to log phase. The IPTG induced expression of the mutant protein was achieved under native conditions, which were optimized for the *Tpv* sHSP 14.3 WT protein as mentioned in Chapter 2. Cell free lysates for each mutant was prepared and analyzed at least two or three times.

The cell lysate of the each sample then was subjected to heat treatment at 60°C, 65°C, 70°C and/or 80°C. This was attempted not only for analyzing the stability of the variants at higher temperatures, but was also used as the first step of purification of the highly stable *Tpv* sHSP 14.3 NTD mutant proteins (Usui *et al*, 2004; Saji *et al*, 2008). The result of expression and heat treatment of each mutant variant is given below.

3.7.1 NTD Proximal Part Mutations

3.7.1.1 I5T Mutant

SDS-PAGE image of the cell lysate indicated high level expression of I5T mutant protein, (Figure 3.35). For analyzing the stability of this mutant at high temperatures,

heat treatment was accomplished at 65°C (for 15 and 30 minutes), at 68°C, 70°C and 80°C (for 30 minutes). The SDS PAGE results show that this mutant protein displayed high stability against heat induced protein denaturation at all elevated temperatures (Figure 3.35).

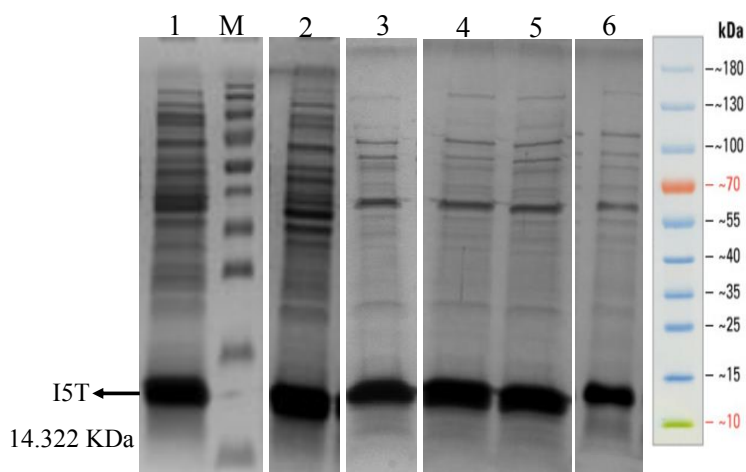


Figure 3.35 Cell lysate of BL21(DE3) *E.coli* cells possessing I5T mutation plasmid viewed on SDS-PAGE. Before Heat (**Lane 1**); after heating at 65°C for 15 minutes (**Lane 2**); after heating at 65°C for 30 minutes (**Lane 3**); after heating at 68°C and 70°C for 30 minutes (**Lane 4** and **Lane 5** respectively) and after heating at 80°C for 30 minutes (**Lane 6**). M = marker SM Page Ruler Prestained Protein Ladder.

3.7.1.2 F8Y, I5TF8Y, F7SF8Y and I5TF7SF8Y Mutants

Next, the clear cell lysates prepared for mutants, F8Y, I5TF8Y, F7SF8Y and I5TF7SF8Y (Figure 3.36) were initially heat treated at 65°C or 60°C for 15 minutes, thereby their stability was confirmed. When heat treated at 70°C and 80°C (for 30 minutes), also, all the four mutants showed good stability, and appreciable removal of the host proteins were achieved, especially at 80°C. The amino acid substitutions, either single, or double or triple, all reduced hydrophobicity and heat stability remained unchanged as compared to WT protein.

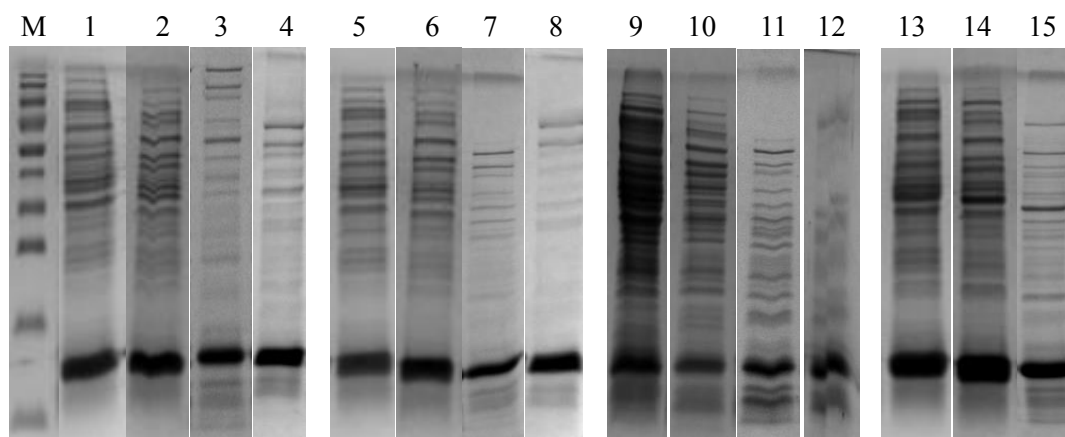


Figure 3.36 Cell free lysate of BL21(DE3) *E.coli* cells possessing mutant plasmid viewed on SDS-PAGE. F8Y variant (14.350 KDa): (**Lane 1, 2, 3 and 4**) before heat, after heating at 65°C (15 minutes), after heating at 70°C (30 minutes), after heating at 80°C (30 minutes). I5TF8Y variant (14.338 KDa): (**Lane 5, 6, 7 and 8**) before heat, after heating at 65°C (15 minutes) after heating at 70°C (30 minutes), after heating at 80°C (30 minutes). F7SF8Y variant (14.290 KDa): (**Lane 9, 10, 11 and 12**) before heat, after heating at 60°C (15 minutes), after heating at 70°C (30 minutes) after heating at 80°C (30 minutes). I5TF7SF8Y variant (14.277 KDa): (**Lane 13, 14 and 15**) before heat, after heating at 65°C (15 minutes), after heating at 70°C (30 minutes). M = marker SM Page Ruler Prestained Protein Ladder.

3.7.2 Combined NTD Proximal and Distal Point Mutation

3.7.2.1 F8YV31G Mutant

Of the three colonies selected for sequencing, only one had the double mutant plasmid. This double mutant plasmid, F8YV31G, was transformed first with high efficiency (1×10^7 cfu/ μ g of DNA) into recombinant *E.coli* T7-Express cells. The two randomly selected colonies for expression analysis although showed high yields of plasmid, the amount of protein in their cell lysate was quite low (Figure 3.37). Considering a possible effect of host cell on expression, the same mutant plasmid was then transferred into BL21(DE3) competent cells. The two sample colonies selected for expression analysis showed relatively high amount of protein in the cell lysate as deduced from the bands on SDS-PAGE. This mutant protein was stable at 65°C and its thermotolerance slightly decreased at 70°C and 80°C (Figure 3.37).

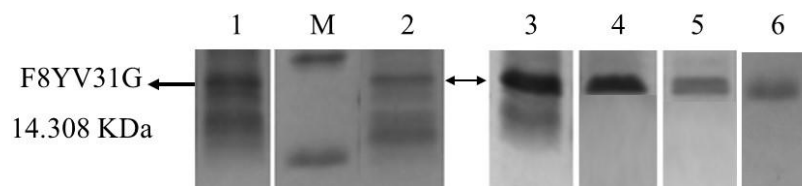


Figure 3.37 Cell lysate of T7-Express *E.coli* cells possessing F8YV31G (14.308 KDa) mutation plasmid viewed on SDS-PAGE. Mutant protein band is indicated by arrow. Before Heat (**Lane 1**); after heating at 60°C for 15 minutes (**Lane 2**). Cell lysate of BL21(DE3) *E.coli* cells possessing F8YV31G mutation plasmid viewed on SDS-PAGE. Before Heat (**Lane 3**); after heating at 65°C for 15 minutes (**Lane 4**) after heating at 70°C for 30 minutes (**Lane 5**); after heating at 80°C for 15 minutes (**Lane 6**). M = marker SM Page Ruler Prestained Protein Ladder.

3.7.3 NTD Middle Part Mutations

3.7.3.1 M12T Mutant

By Met12 to Thr substitution, a highly conserved hydrophobic methionine, at position 12 in the *Tpv* sHSP 14.3, is replaced with a similar size, but hydrophilic residue, Threonine. Two colonies were chosen for downstream experiments. However, the mutant protein concentration of cell lysates of both the colonies appeared quite low when analyzed by SDS-PAGE, (Figure 3.38). In the attempt to clear the cell lysate from the host *E.coli* proteins by heating the cell lysate at 60°C and 65°C, the band of the mutant protein was almost completely lost.

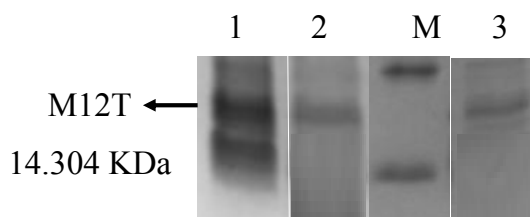


Figure 3.38 Cell lysate of BL21(DE3) *E.coli* cells possessing M12T (13.304 KDa) mutation plasmid viewed on SDS-PAGE. Before Heat (**Lane 1**); after heating at 65°C for 15 minutes (**Lane 2**); after heating at 60°C for 15 minutes (**Lane 3**). M = marker SM Page Ruler Prestained Protein Ladder.

3.7.3.2 E11V Mutant

Highly conserved glutamic acid at position 11, in the *Tpv* sHSP 14.3, is replaced, with hydrophobic residue valine. When E11V mutant sHSP was expressed in T7-Express competent cells, transformation efficiency for this mutant was high in this expression system. But the quantity of the protein, in the soluble part of cell lysate was found to be very low. Moreover, significant amount of the protein was lost on treating the cell lysate to 65°C, as evident from the protein band of 14.3 KDa (Figure 3.39).

Cell lysate was also prepared for E11V mutant in *E.coli* BL21(DE3) cells, where, the protein concentration in the cell lysate was observed to be higher than the T7-Express system (Figure 3.39). Decrease in the thermotolerance of the mutant protein was observed for this mutation, as is obvious from the SDS-PAGE images of the mutant (Figure 3.39). The protein band again was lost after heating at temperatures of 60°C, 65°C, 70°C and 80°C (Figure 3.39).

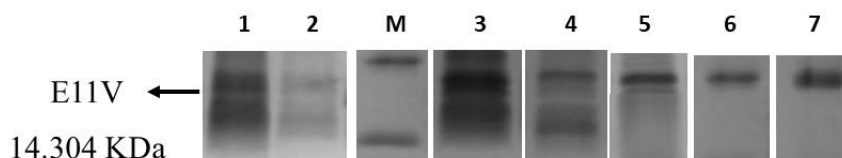


Figure 3.39 Cell lysate of T7-Express *E.coli* cells possessing E11V (14.304 KDa) mutation plasmid viewed on SDS-PAGE. Before Heat (Lane 1); after heating at 65°C for 15 minutes (Lane 2). Cell lysate of BL21(DE3) *E.coli* cells possessing E11V mutation plasmid viewed on SDS-PAGE. Before Heat (Lane 3); after heating at 60°C for 15 minutes (Lane 4) after heating at 65°C for 20 minutes (Lane 5); after heating at 70°C for 30 minutes (Lane 6) and after heating at 80°C for 15 minutes (Lane 7). M = marker SM Page Ruler Prestained Protein Ladder.

3.7.3.3 E22G Mutant

The second glutamic acid residue (E22), which is highly conserved in the NTD of *Tpv* sHSP 14.3 was exchanged with hydrophobic glycine. In the *E.coli* T7-Express cells the mutant protein concentration in the cell lysate appeared to be very low as deduced from the band thickness (Figure 3.40). When this mutant variant was

expressed in *E.coli* BL21 cells, there has not been a detectable change in the protein level. Decrease in the thermostability of the protein, at temperatures between 60°C and 80°C, was observed as compared to WT *Tpv* sHSP 14.3 (Figure 3.40).

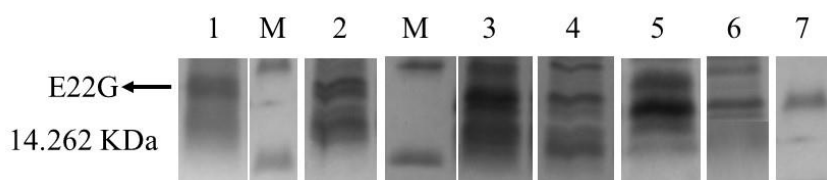


Figure 3.40 Cell lysate of T7-Express *E.coli* cells possessing E22G (14.262 KDa) mutation plasmid viewed on SDS-PAGE. Before Heat (**Lane 1**); after heating at 60°C for 15 minutes (**Lane 2**). Cell lysate of BL21(DE3) *E.coli* cells possessing E22G mutation plasmid viewed on SDS-PAGE. Before Heat (**Lane 3**); after heating at 60°C for 15 minutes (**Lane 4**) after heating at 65°C for 20 minutes (**Lane 5**); after heating at 70°C for 30 minutes (**Lane 6**) and after heating at 80°C for 15 minutes (**Lane 7**). M = marker SM Page Ruler Prestained Protein Ladder.

3.7.4 NTD Distal Part Mutations

3.7.4.1 F26YI27T and V23GF26YI27T Mutants

In this mutagenesis, hydrophobicity in the distal part of the NTD was reduced by amino acid exchange at multiple positions. Expression of the double mutant, F26YI27T, and triple mutant, V23GF26YI27T, variants in the *E.coli* BL21(DE3) cells were studied. For both the mutants, the concentration of expressed protein was not very high in the cell lysate before heat treatment, however, most of the protein of interest was lost along with the host proteins, when heated at 60°C, 70°C and 80°C (Figure 3.41 and 3.42).

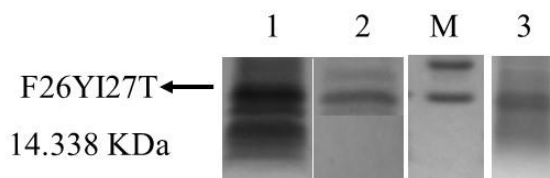


Figure 3.41 Cell lysate of BL21(DE3) *E.coli* cells possessing F26YI27T (14.338 KDa) mutation plasmid viewed on SDS-PAGE. **Lane 1, 2 and 3** is Before heat, after heat treatment at 70° and 80°C respectively. M = marker Unstained Protein Ladder SM0431.

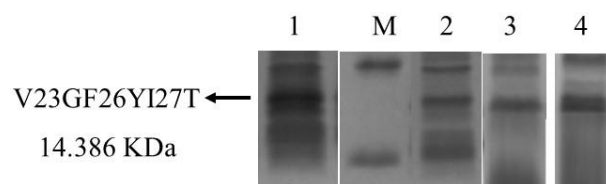


Figure 3.42 Cell lysate of BL21(DE3) *E.coli* cells possessing V23GF26YI27T (14.386 KDa) mutation plasmid viewed on SDS-PAGE. Before Heat (**Lane 1**); after heating at 60°C for 15 minutes (**Lane 2**); after heating at 70°C for 30 minutes (**Lane 3**); after heating at 80°C for 15 minutes (**Lane 4**); M = marker SM Page Ruler Prestained Protein Ladder.

3.7.4.2 V31G and V31IL33I Mutants

This group includes mutations present at the junction of NTD with the alpha crystallin domain in *Tpv* sHSP 14.3. The attempts to generate the single mutant V31G resulted in very limited number of mutant colonies (only 10 colonies were obtained), indicating that mutation at this position may not be virtuous for cell viability. The second mutation aiming at increasing the hydrophobicity at the position where NTD joins ACD, *i.e.*, V31IL33I, we could achieve only one colony after mutagenesis. The sequencing results were 100% as expected for both the mutants. Expression of the both mutant variants were studied in *E.coli* BL21(DE3) cells. The mutant plasmid V31G was also transformed into T7-Express competent cells. Although high transformation efficiency was attained the T7-Express recombinant cells did not show high concentration of protein in the cell lysate. However, the cell lysate of *E.coli* BL21(DE3) mutant cells containing this mutant plasmid, displayed satisfactory concentration of the mutant protein, although less than WT. However considerable loss of the sHSP mutant protein was observed when the cell lysate was subjected to high temperatures, between 60°C and 80°C, indicating low thermotolerance owing to this mutation (Figure 3.43).

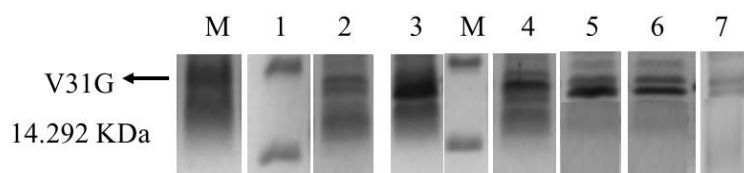


Figure 3.43 Cell lysate of T7-Express *E.coli* cells possessing V31G (14.292 KDa) mutation plasmid viewed on SDS-PAGE. Before Heat (**Lane 1**); after heating at 60°C for 15 minutes (**Lane 2**); Cell lysate of BL21(DE3) *E.coli* cells possessing V31G mutation plasmid viewed on SDS-PAGE. Before Heat (**Lane 3**); after heating at 60°C for 15 minutes (**Lane 4**); after heating at 65°C for 20 minutes (**Lane 5**); after heating at 70°C for 30 minutes (**Lane 6**); after heating at 80°C for 15 minutes (**Lane 7**); M = marker SM Page Ruler Prestained Protein Ladder.

Despite lower than the WT protein, satisfactory concentration of V31IL33I mutant protein was found in the cell lysate of the respective mutant BL21(DE3) cells. But when the cell lysate was heat treated at 60°C, 65°C, 70°C and 80°C, considerable loss of the mutant protein was observed (figure 3.44).

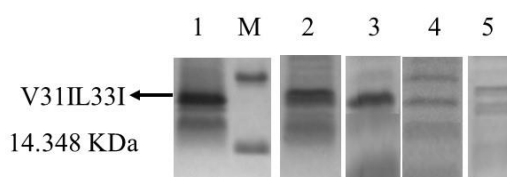


Figure 3.44 Cell lysate of BL21(DE3) *E.coli* cells possessing V31IL33I (14.348 KDa) mutation plasmid viewed on SDS-PAGE. Before Heat (**Lane 1**); after heating at 60°C for 15 minutes (**Lane 2**); after heating at 65°C for 20 minutes (**Lane 3**); after heating at 70°C for 30 minutes (**Lane 4**); after heating at 80°C for 15 minutes (**Lane 5**); M = marker SM Page Ruler Prestained Protein Ladder.

When expression results of the NTD mutant proteins are compared, it can be clearly seen that the highest amount and stability of the protein was observed in the mutant variants with the mutations in the NTD proximal region. On the other hand, the mutations in the middle and distal section of NTD resulted in less stable variant protein with low concentration. This indicates that this section might be important for the stability of the sHSP itself. Previously it has been shown in human HSPB6 that the distal NTD is involved in monomer – monomer interactions to form the dimer interface, while the proximal portion of NTD is thought to be involved in maintenance of oligomer stability (Weeks *et al.*, 2014; Bakthisaran, Tangirala and

Rao, 2015; Shatov *et al.*, 2018). Therefore, there is a possibility that mutations in the distal region of NTD, leading to decreased hydrophobicity, might have played a role in loosening of the dimer, thus dispersing the oligomer assembly, characteristic of the sHSP, leading to decrease stability of the variants.

3.8 Preparation of Cell Lysate on Large Scale and Purification by HPLC

3.8.1 Sample Preparation

Cell lysates of NTD proximal part mutant protein variants which exhibited high stability, together with WT *Tpv* sHSP 14.3 were produced on large scale, in order to purify protein by Ion Exchange Chromatography.

The cell lysate was first subjected to heat treatment at 70°C for 30 minutes, followed by centrifugation, to get rid of the host heat labile proteins (Figure 3.45).

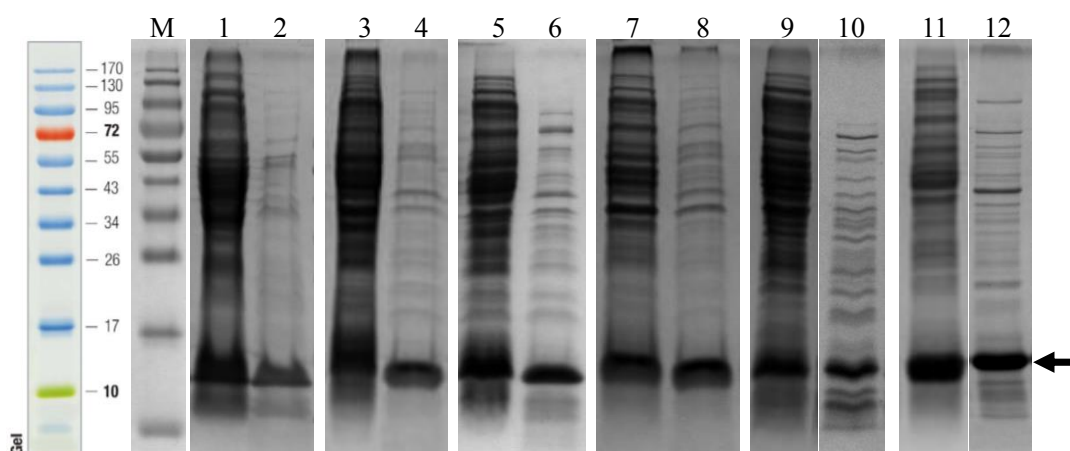


Figure 3.45 Cell lysate of BL21(DE3) *E.coli* cells possessing *Tpv* sHSP 14.3 WT (Before Heat: **Lane 1** and after heating at 70°C: **Lane 2**), I5T mutation plasmid (Before Heat: **Lane 3** and after heating at 70°C: **Lane 4**), F8Y mutation plasmid (Before Heat: **Lane 5** and after heating at 70°C: **Lane 6**), I5TF8Y mutation plasmid (Before Heat: **Lane 7** and after heating at 70°C: **Lane 8**), F7SF8Y mutation plasmid (Before Heat: **Lane 9** and after heating at 70°C: **Lane 10**) and I5TF7SF8Y mutation plasmid (Before Heat: **Lane 11** and after heating at 70°C: **Lane 12**), viewed on SDS-PAGE. M = marker SM Page Ruler Prestained Protein Ladder.

Next, the buffer was exchanged, from lysis buffer in to start buffer suitable for Ion-Exchange procedure and also concentrated by ultrafiltration. SDS-PAGE analysis was performed at each step to confirm the results (Figure 3.46). The concentrated sample obtained was then filtered using 0.45 μ M Millipore Filter and its concentration was measured by PicoDrop, to know the amount of protein being loaded on the column and to avoid over-stepping the maximum binding capacity of the column.

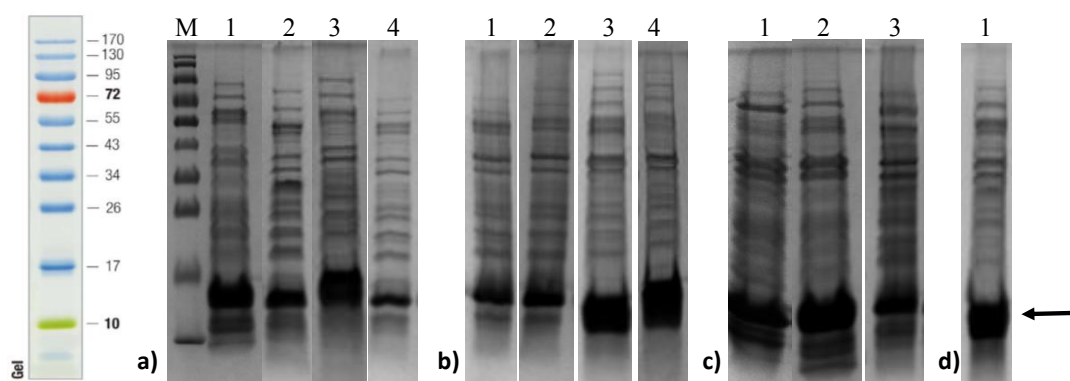


Figure 3.46 Production of multiple volumes of concentrated protein of (a) *Tpv* sHSP 14.3 WT, (b) I5T variant, (c) F8Y variant and (d) I5TF8Y variant (the band indicated by black arrow) in ion exchange start buffer, after ultrafiltration. M = marker SM Page Ruler Prestained Protein Ladder.

3.8.2 Anion Exchange Chromatography

In this study, the concentrated protein samples of 20 – 35 mg/ml were loaded to HiTrap® Q XL column for purification by Anion Exchange chromatography. A representative chromatogram for each *Tpv* sHSP 14.3 WT and the mutants I5T and F8Y are shown in Figure 3.47. *Tpv* sHSP 14.3 WT/mutant proteins were eluted in linear gradient of 0 to 1M NaCl. Fractions analyzed by SDS-PAGE indicated the presence of the protein, in the fractions from 6 to 13 in the gradient between 0.35 and 0.65 M NaCl. The highest concentration of *Tpv* sHSP 14.3 protein was found in the fractions 7 and 8, (Figure 3.47 and 3.48).

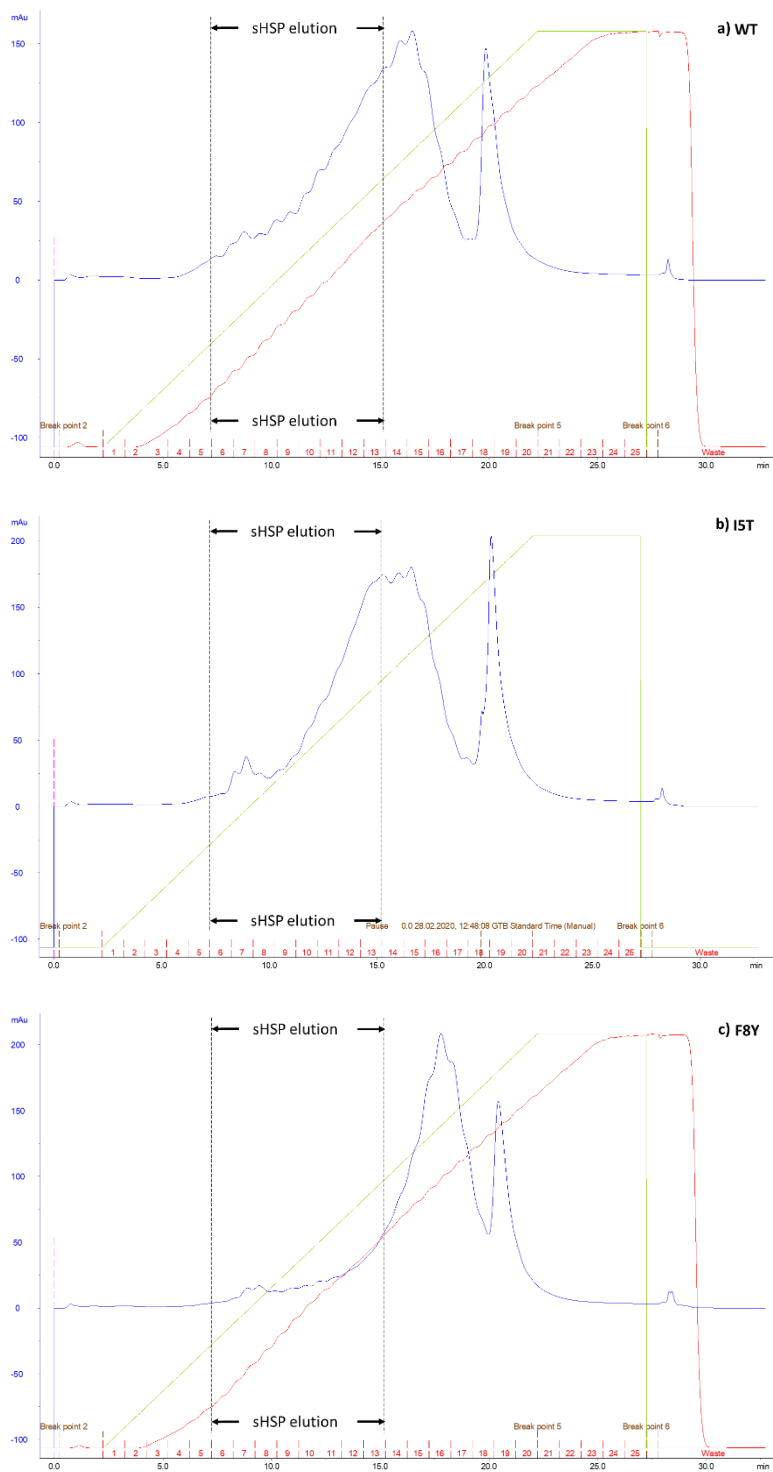


Figure 3.47 Chromatogram of (a) *Tpv* sHSP 14.3 WT, (b) I5T variant and (c) F8Y variant protein eluting from 0 to 1M NaCl linear gradient, indicating time and absorbance of each peak in Anion Exchange elution; The UV absorbance, conductivity and percentage of elution buffer are indicated on the chromatogram by the blue, red and green lines, respectively. The x-axis indicates the fractions.

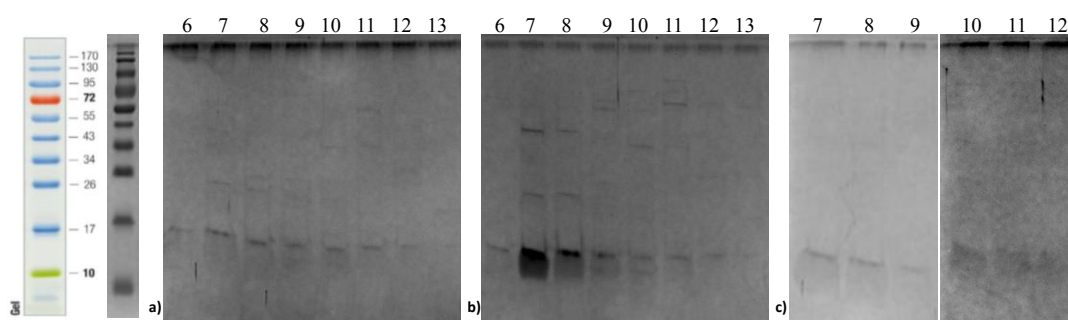


Figure 3.48 SDS-PAGE of fractions containing (a) *Tpv* sHSP 14.3 WT (fractions 6 – 13), (b) *Tpv* sHSP 14.3 I5T (fractions 6 – 13), (c) *Tpv* sHSP 14.3 F8Y (fractions 7 – 12). M = marker SM Page Ruler Prestained Protein Ladder.

3.8.3 Concentration of the eluted protein by Ultrafiltration Method

The eluted protein sample which was diluted when eluting were pooled and concentrated as described in Material and Methods. In addition to this, the buffer of the sample protein was also exchanged to the buffer with low concentration of NaCl. The representative figure of the concentrated and buffer exchanged samples of each, *Tpv* sHSP 14.3 WT and its mutant variants I5T and F8Y are shown in Figure 3.49. The black arrow indicates the presence of protein of interest *Tpv* sHSP 14.3 WT/mutant.

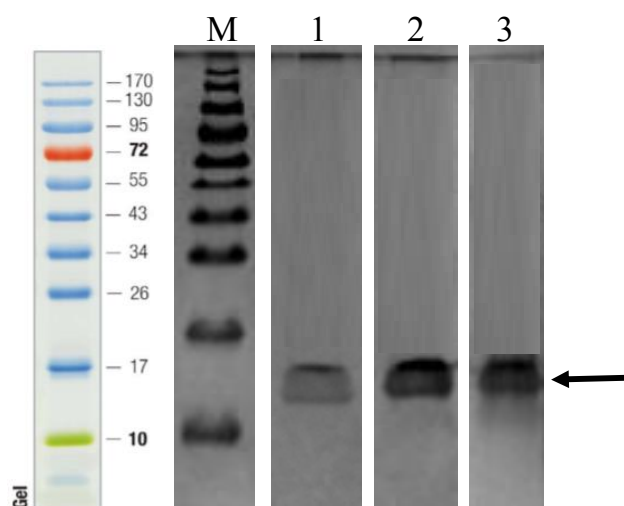


Figure 3.49 SDS-PAGE of (Lane 1) *Tpv* sHSP 14.3 WT, (Lane 2) F8Y variant and (Lane 3) I5T variant, after concentration of the eluted fractions into SEC buffer. M = marker SM Page Ruler Prestained Protein Ladder.

3.9 Citrate Synthase Enzyme Activity Assay

In this experiment, the chaperone activity, at two different molar ratios (CS:sHSP) 1:1700 (*i.e.*, 486 μg chaperone/ μg substrate) and 1:850 (*i.e.*, 243 μg chaperone/ μg substrate), was analyzed by the ability of the sHSP to prevent aggregation of citrate synthase enzyme and protect its activity at elevated temperature. The optimum temperature for the activity of citrate synthase from pig heart is 35°C (Shin *et al.*, 2010). But its activity is lost when heated at 47°C, (Figure 3.50). However, *Tpv* sHSP 14.3 WT protects this activity nearly 80% (slope = 0.58 Abs/min) at 1:1700 CS:chaperone ratio, while decreasing the WT concentration to half (1:850 CS:chaperone molar ratio) decreased the chaperone activity by 30% (slope = 0.46 Abs/min), as compared to positive control.

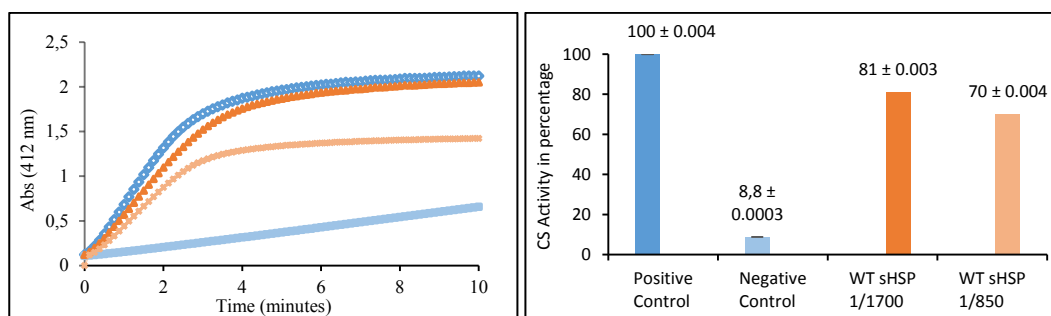


Figure 3.50 Enzyme activity plot and bar chart (representing the slopes) of the percentage protection of citrate synthase enzyme activity, at 47°C, by *Tpv* sHSP 14.3 WT at 1:1700 and 1:850 molar ratios. The CS activity is indicated (in terms of percentage when positive control is taken as 100%, or in terms of absorbance per minute) by the y-axis. Each data point represents the mean of three independent trials and is shown as the mean \pm standard deviation (SD). (—◆—) positive control. (—□—) negative control. (—▲—) WT 1/1700 and (—■—) WT 1/850.

Also our study pointed out that, mutations in the initial 8 residues of the NTD of *Tpv* sHSP 14.3, reduced the chaperone activity. Our results show that, decreasing the hydrophobic nature at residue number 5 by Thr (with polar sidechain) replacement (mutant I5T), decreased the protection effect of sHSP (slope = 0.27 Abs/min) by nearly half of the WT (slope = 0.58 Abs/min) at 1:1700 CS:chaperone molar ratio (Figure 3.51a and 3.51b). Decrease in the chaperone activity was the same when the

mutant sHSP concentration was reduced by half, *i.e.*, at 1:850 CS:chaperone molar ratio the slope obtained was 0.25 Abs/min.

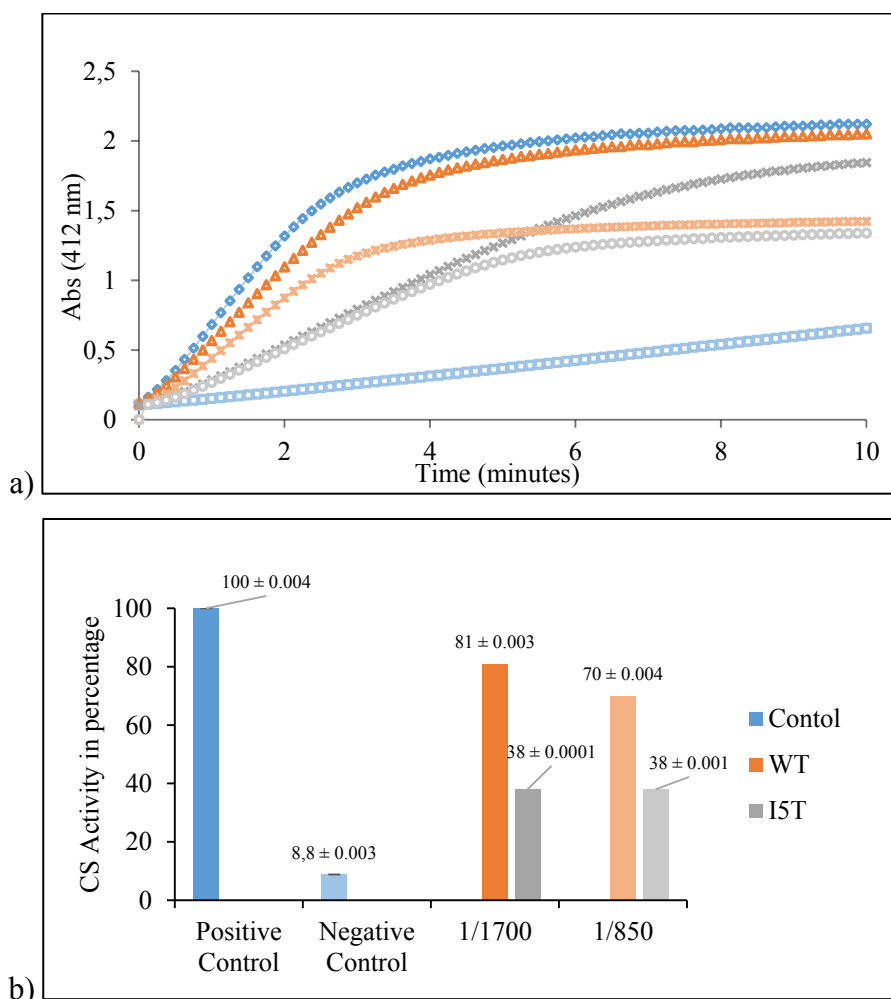


Figure 3.51 Enzyme activity plot (—◆—) positive control. (—□—) negative control. (—▲—) WT 1/1700, (—×—) WT 1/850, (—×—) IST 1/1700 and (—○—) IST 1/850 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase enzyme activity, at 47°C, by *Tpv* sHSP 14.3 WT (represented by orange color) and IST mutant (represented by grey color) at 1:1700 and 1:850 molar ratios. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

A single mutation at position 8, substituting the hydrophobic phenylalanine with tyrosine having hydrophobic sidechain (polar uncharged), reduced chaperone activity by 34% (as compared to positive control) at 1:1700 CS:chaperone molar ratio (slope = 0.47 Abs/min). But, the protective effect of F8Y mutant decreased almost by half (to 38%, as compared to WT at this ratio) when sHSP concentration was halved, *i.e.*, 1:850 substrate:chaperone ratio (slope = 0.24 Abs/min).

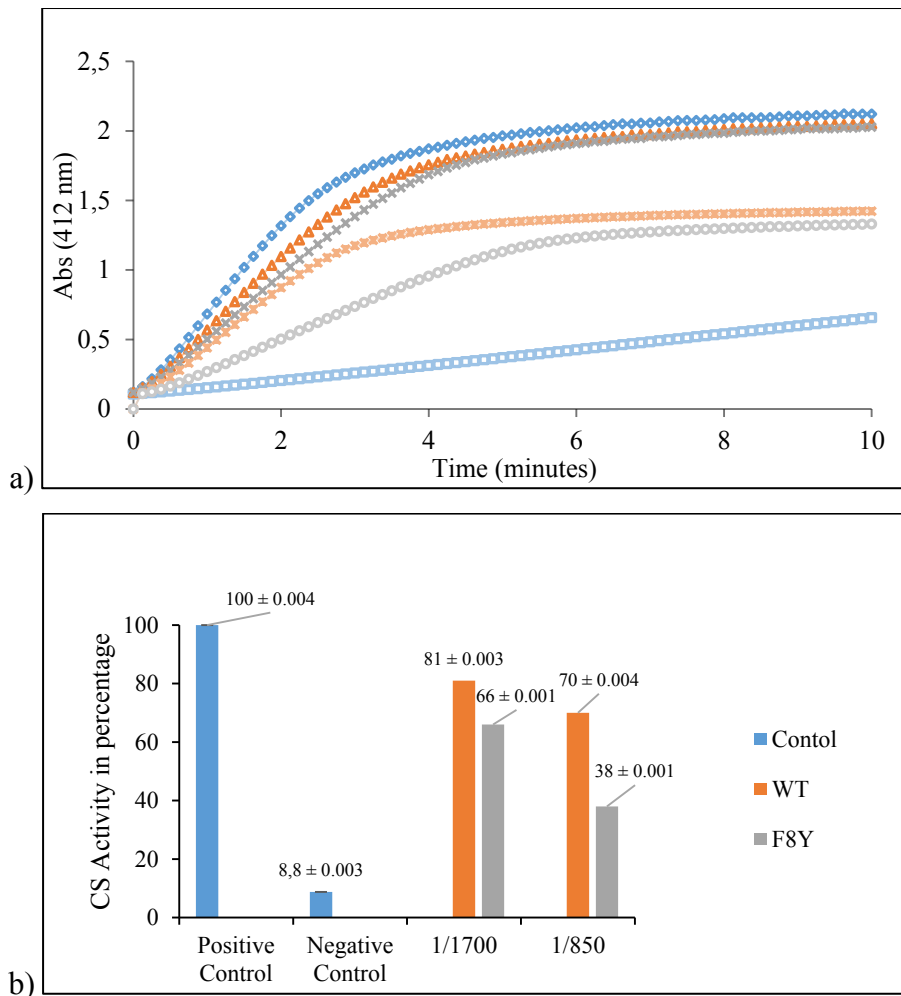


Figure 3.52 Enzyme activity plot (—◆—) positive control. (—□—) negative control. (—▲—) WT 1/1700, (—×—) WT 1/850, (—×—) F8Y 1/1700 and (—○—) F8Y 1/850 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase enzyme activity, at 47°C, by *Tpv* sHSP 14.3 WT (represented by orange color) and F8Y mutant (represented by grey color) at 1:1700 and 1:850 molar ratios. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

Interestingly, a combination of these two mutations (*i.e.*, I5 and F8, replaced by threonine and tyrosine respectively), resulted in protective effect which was more than the former and slightly less than the latter mutant alone, at 1:1700 CS:chaperone molar ratio (slope = 0.36 Abs/min). However, at 1:850 CS:chaperone ratio, double mutation reduced chaperone activity less than the individual point mutations (slope = 0.22 Abs/min) (Figure 3.53a and 3.53b).

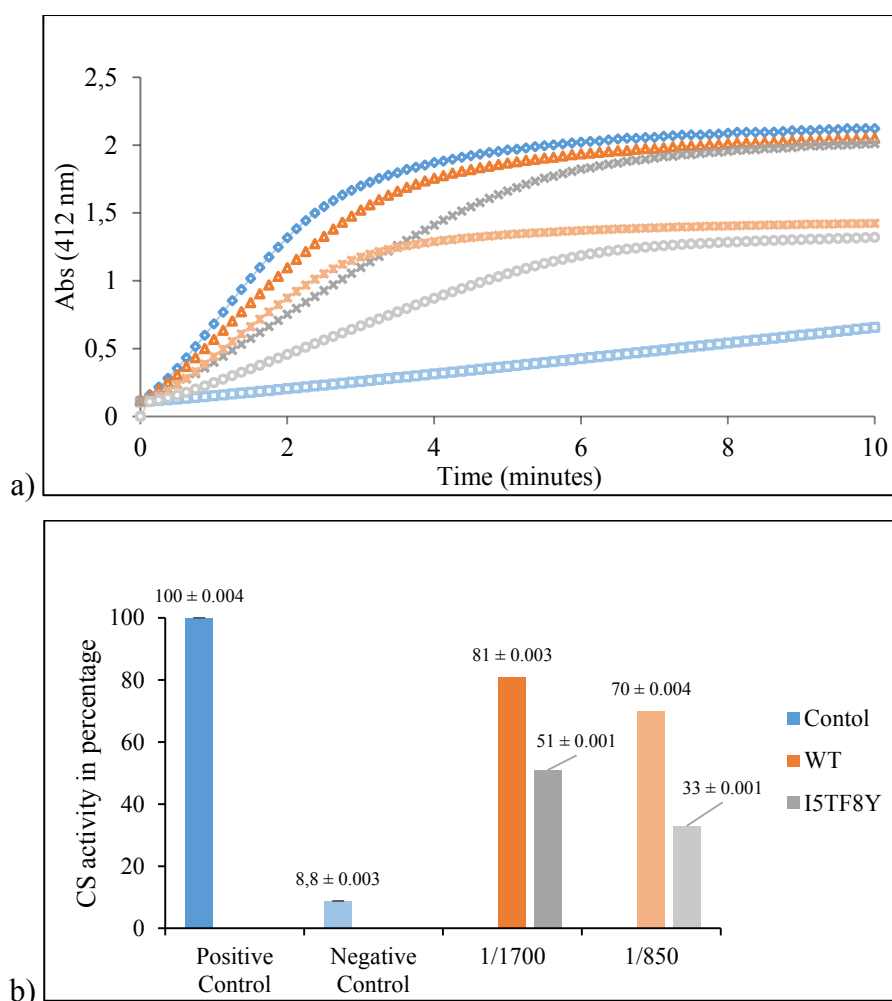


Figure 3.53 Enzyme activity plot (\diamond) positive control. (\square) negative control. (\triangle) WT 1/1700, (\times) WT 1/850, (\triangle) I5TF8Y 1/1700 and (\circ) I5TF8Y 1/850 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase enzyme activity, at 47°C, by *Tpv* sHSP 14.3 WT (represented by orange color) and I5TF8Y mutant (represented by grey color) at 1:1700 and 1:850 molar ratios. Each data point represents the mean of three independent trials and is shown as the mean \pm standard deviation (SD).

In order to study the combined effect of hydrophobic residues, we have substituted the two adjacent phenylalanine amino acids at position 7 and 8 with polar but uncharged serine and tyrosine respectively. The F7SF8Y double mutation resulted in drastic loss of chaperone activity (>70%) which was independent of the concentration of chaperone, (Figure 3.54). The slopes obtained at 1:1700 and 1:850 CS:chaperone molar ratio, for F7SF8Y mutant variant, were 0.18 Abs/min and 0.17 Abs/min, respectively. This result indicates the importance of the two adjacent

phenylalanine amino acids in NTD, in the chaperone activity, specifically for providing heat protection of the substrate protein.

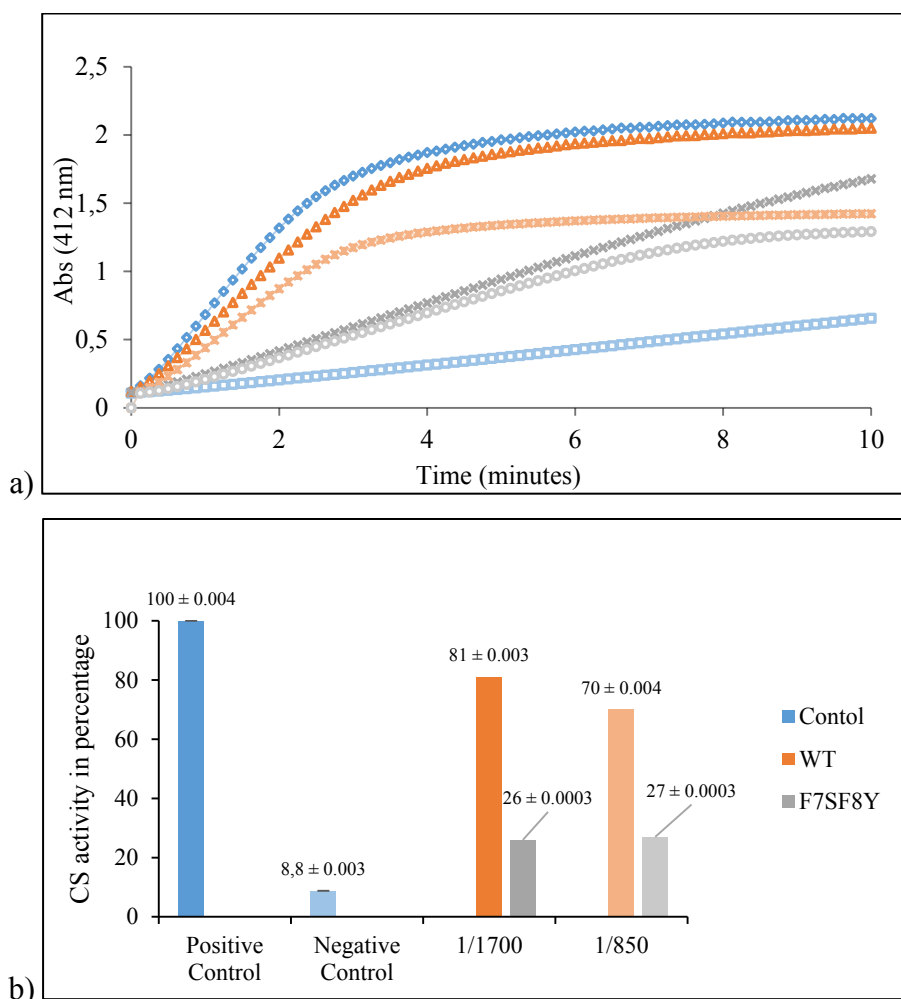


Figure 3.54 Enzyme activity plot (—◆—) positive control. (—□—) negative control. (—▲—) WT 1/1700, (—×—) WT 1/850, (—×—) F7SF8Y 1/1700 and (—○—) F7SF8Y 1/850 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase enzyme activity, at 47°C, by *Tpv* sHSP 14.3 WT (represented by orange color) and F7SF8Y mutant (represented by grey color) at 1:1700 and 1:850 molar ratios. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

However, in addition to replacing the two phenylalanine amino acids at 7 and 8 position, if hydrophobicity was even decreased by replacement of Ile with Thr at position 5 (I5TF7SF8Y), the chaperone activity was lower than that of the two single (I5T and F8Y) and double (I5TF8T) mutants and similar to that of F7SF8Y mutant at 1:850 molar ratios (slope = 0.17 Abs/min) (Figure 3.55). However, addition of

F7S mutation to I5TF8Y double mutation, slightly decreased the chaperone activity in the triple mutant than the double mutant (I5TF8Y) but increased it than that of F7SF8Y mutant variants, at 1:1700 molar ratio (slope = 0.28 Abs/min). This result may suggest that at high molar ratio, the adverse effect of I5T and F7S was partially compensated by F8Y mutation in the triple mutant variant.

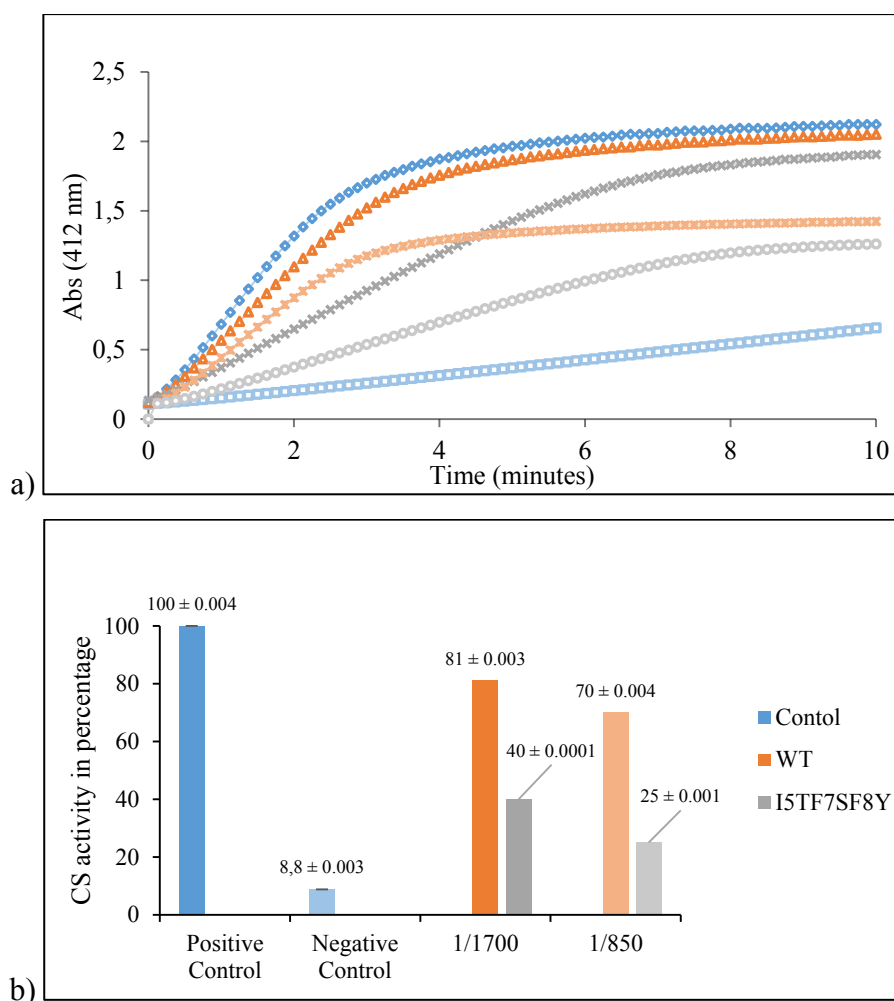


Figure 3.55 Enzyme activity plot (—◆—) positive control. (—□—) negative control. (—▲—) WT 1/1700, (—×—) WT 1/850, (—*—) I5TF7SF8Y 1/1700 and (—○—) I5TF7SF8Y 1/850 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase enzyme activity, at 47°C, by *Tpv* sHSP 14.3 WT (represented by orange color) and I5TF7SF8Y mutant (represented by grey color) at 1:1700 and 1:850 molar ratios. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

3.10 Citrate Synthase Aggregation Assay

Aggregation was studied by measuring of light absorbed after light scattering by aggregated particles, using 96-well plate, as described in Chapter 2, Material and Methods.

3.10.1 NTD Proximal Mutants

Two different substrate:sHSP molar ratios, 1:7 (which accounts for 140 μg of sHSP and 70 μg of citrate synthase/mL in the reaction mixture) and 1:35 (which accounts for 700 μg of sHSP and 70 μg of citrate synthase /mL of the reaction mixture), were used for studying the protection of sHSPs against aggregation of the substrate protein at higher temperature. At 1:7 substrate:sHSP molar ratio, WT sHSP suppressed heat aggregation of the CS to 69% as compared to the control (the aggregation of the control, without sHSP, was taken as 100% aggregation and, therefore, 0% protection against heat induced aggregation). When chaperone concentration was increased by 5-times, protection from aggregation increased to 80% (Figure 3.56).

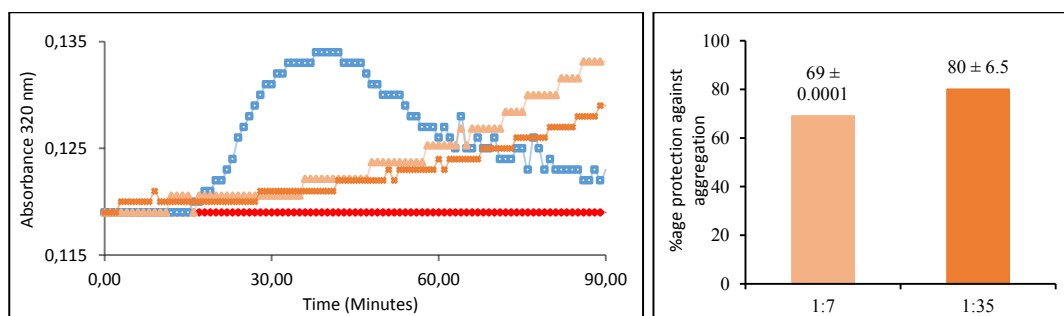


Figure 3.56 Light scattering plot and bar chart (representing the slope) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Tpv* sHSP 14.3 WT, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean \pm standard deviation (SD). (— \diamond —) Blank, (— \square —) CS control, (— \triangle —) WT 1:7 and (— \times —) WT 1:35.

When hydrophobicity at position 5 was reduced by Ile to Thr exchange, protection of CS from thermal aggregation was reduced to 40% at 1:7 substrate:chaperone ratio.

At 1:35 substrate:chaperone ratio heat protection effect was about 60%, but it was still less than WT at the same ratio (Figure 3.57).

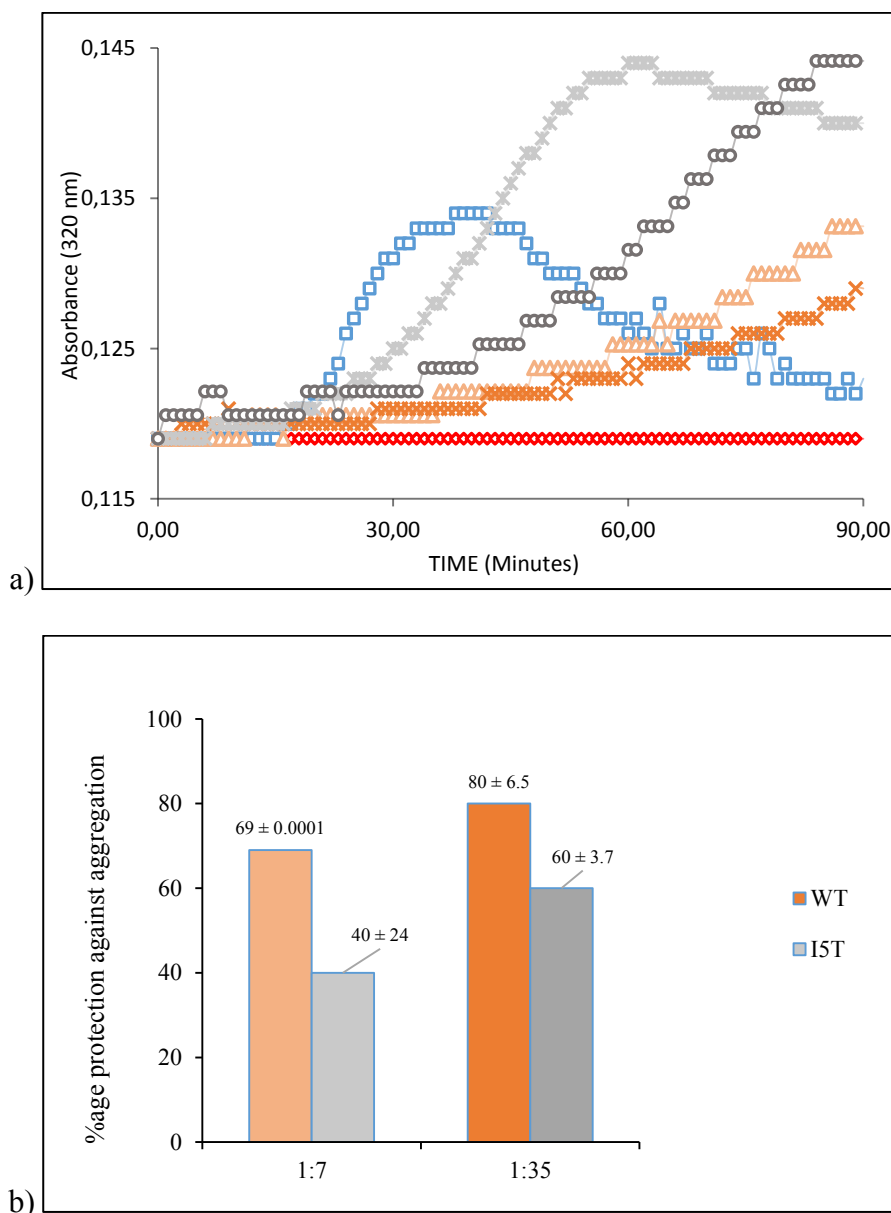


Figure 3.57 Light scattering plot (—◇—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—×—) I5T 1:7 and (—○—) I5T 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Tpv* sHSP 14.3 WT and its mutant variant *Tpv* sHSP 14.3 I5T, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

Decreasing hydrophobicity by Phe to Tyr substitution at position 8, the protection effect was same as the WT sHSP (*i.e.*, 69%) at 1:7 substrate:chaperone ratio. However, increased chaperone concentration (at 1:35 ratio) did not increase the protection any further, for this mutant, and thus at 1:35 substrate:chaperone, also resulted in 31% aggregation or 69% protection, which was less than that of the WT sHSP at the same chaperone concentration (Figure 3.58).

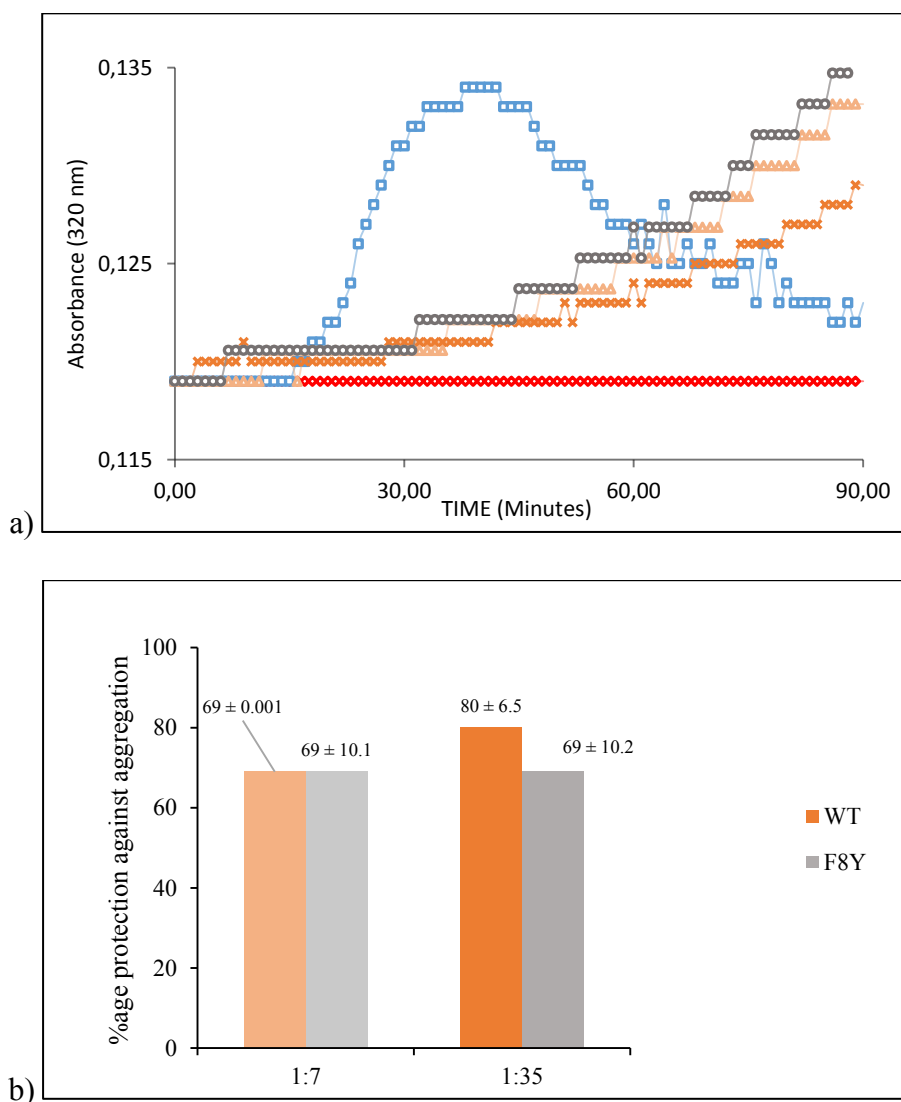


Figure 3.58 Light scattering plot (—◇—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—+—) F8Y 1:7 and (—○—) F8Y 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Tpv* sHSP 14.3 WT and its mutant variant *Tpv* sHSP 14.3 F8Y, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

Decreased hydrophobicity at positions 5 and 8 simultaneously in the I5TF8Y by double mutation provided about 60% protection from heat aggregation at both, low and high (1:7 and 1:35) substrate:chaperone molar ratio (Figure 3.59). Protective effect of this mutant was less than the WT sHSP at low and high chaperone concentrations.

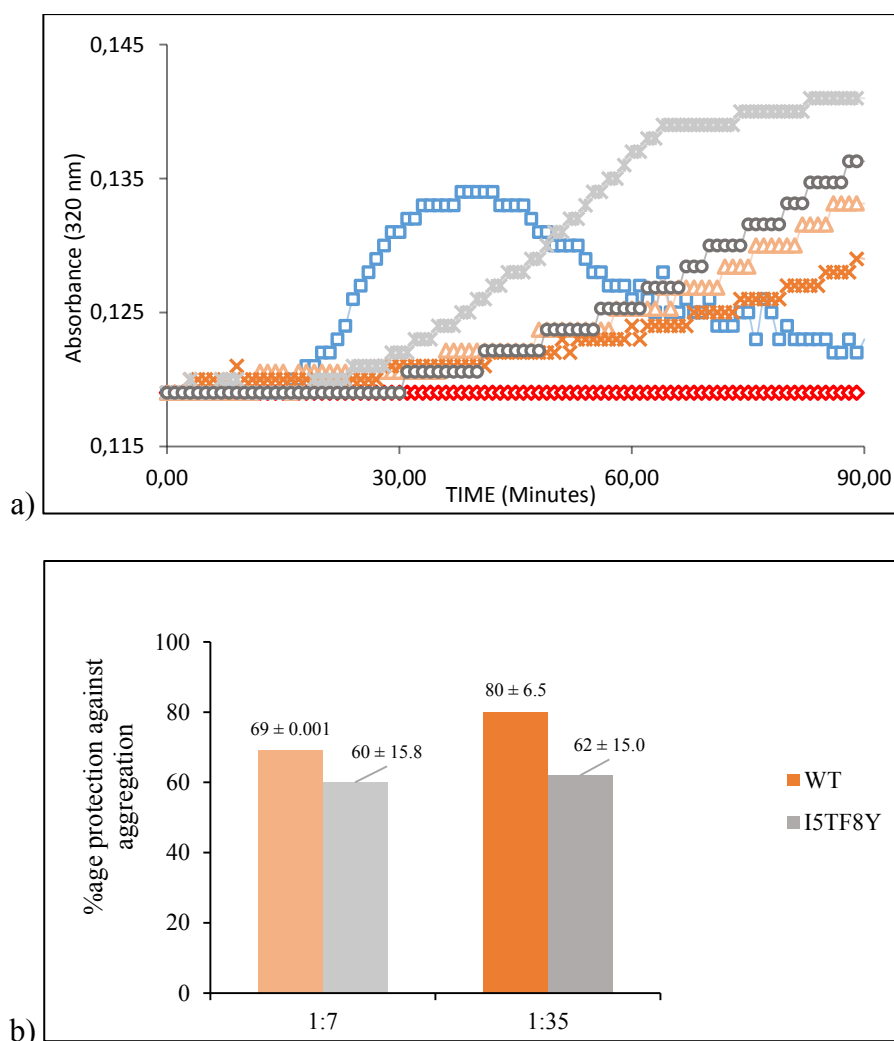


Figure 3.59 Light scattering plot (—♦—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—■—) I5TF8Y 1:7 and (—○—) I5TF8Y 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Tpv* sHSP 14.3 WT and its mutant variant *Tpv* sHSP 14.3 I5TF8Y, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

Another double mutation of decreased hydrophobicity at positions 7 and 8 (F7SF8Y) resulted in significant decrease in the protection of the CS from heat aggregation to 40%, which was almost half of the WT sHSP at both substrate:chaperone concentrations (Figure 3.60).

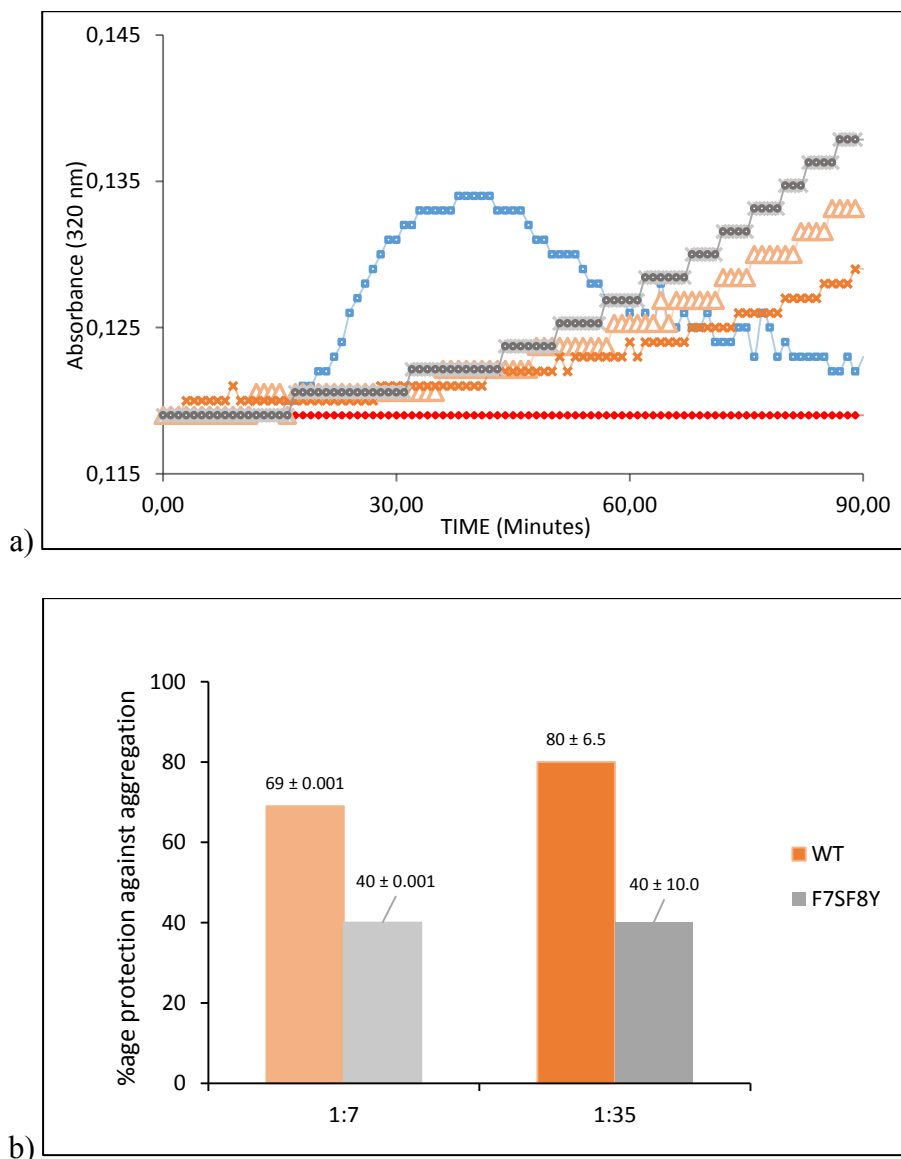


Figure 3.60 Light scattering plot (—♦—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—■—) F7SF8Y 1:7 and (—○—) F7SF8Y 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Trpv* sHSP 14.3 WT and its mutant variant *Trpv* sHSP 14.3 F7SF8Y, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

Protection effect displayed by the triple mutant variant with decrease in hydrophobicity at positions 5, 7 and 8 (I5TF7SF8Y), was not different than I5TF8Y (40%) at 1:7 substrate:chaperone ratio. However, at 1:35 substrate:chaperone ratio, protection from heat aggregation was 45%, which is somewhat higher than F7SF8Y mutant but less than the I5TF8Y mutant variant (Figure 3.61).

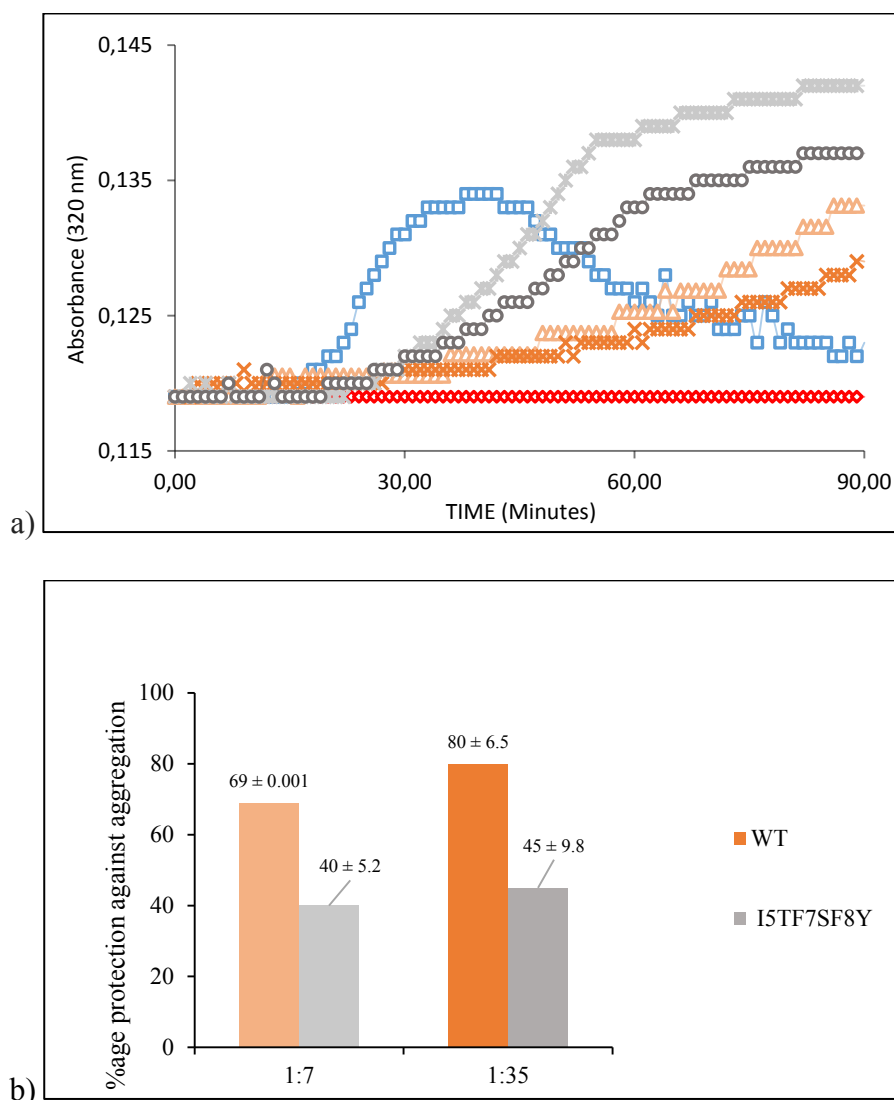


Figure 3.61 Light scattering plot (—◇—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—×—) I5TF7SF8Y 1:7 and (—○—) I5TF7SF8Y 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Tpv* sHSP 14.3 WT and its mutant variant *Tpv* sHSP 14.3 I5TF7SF8Y, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

Overall, F8Y mutation results in protection effect from heat aggregation similar to WT sHSP. However, when this mutation is combined with F7S mutation in the F7SF8Y double mutant variant, 3 and 2-fold decrease in the protection effect, at 1:7 and 1:35 substrate:chaperone ratio, respectively, was observed as compared to wildtype. This result may indicate the importance of phenylalanine residues adjacently located at position 7 and 8 for substrate binding. With 5-fold increase in sHSP concentration, substrate CS aggregation rate at 45°C was almost unchanged in case of the double mutants (F7SF8Y) and the triple mutant (I5TF7SF8Y). However, WT sHSP and I5T mutant variant protected CS from aggregation at 45°C, better at higher concentrations (700µg/mL) than low concentrations (140 µg/mL).

3.10.2 NTD Middle Part Mutants

Decreasing the hydrophobicity in the middle of the NTD of *Tpv* sHSP 14.3, at a highly conserved residue, M12, severely affected its chaperone activity. At 1:7 substrate:sHSP molar ratio, the mutation involving replacement of a highly conserved methionine, (M12T), with threonine, exhibited 40% protection against thermal aggregation which was around half of that of WT at similar, 1:7, ratios. However, this protective effect of the mutant completely vanished at high substrate:sHSP molar ratio (1:35). Moreover, increase in aggregation, as detected by an increase in light scattering, as compared to the control, was observed. This may imply that at high concentrations of M12T mutant sHSP, CS heat aggregation was enhanced to the extent even more than the case where the chaperone is absent (Figure 3.62). On the other hand, abolishing the highly conserved negative charge at position 11 and 22 by introducing hydrophobic residues at these positions, (E11V and E22G mutations) resulted in decrease in the protective effect (to more than half of WT) against thermal aggregation at 1:35 substrate:sHSP molar ratio,(Figure 3.63 and 3.64). When chaperone concentration was decreased, (*i.e.*, 1:7 substrate:sHSP molar ratio) the protective effect of variant E11V almost remained unaltered (Figure 3.63), while E22G variant's protective effect was even further decreased, which was about

20%, as compared to the control (*i.e.*, absence of chaperone) (Figure 3.64). This result may suggest that increased hydrophobicity in the middle part of the NTD does not make an additional contribution to the chaperone function of the *Tpv* sHSP 14.3. Substituted charged residues, however, might be important for the structure and/or function of the sHSP.

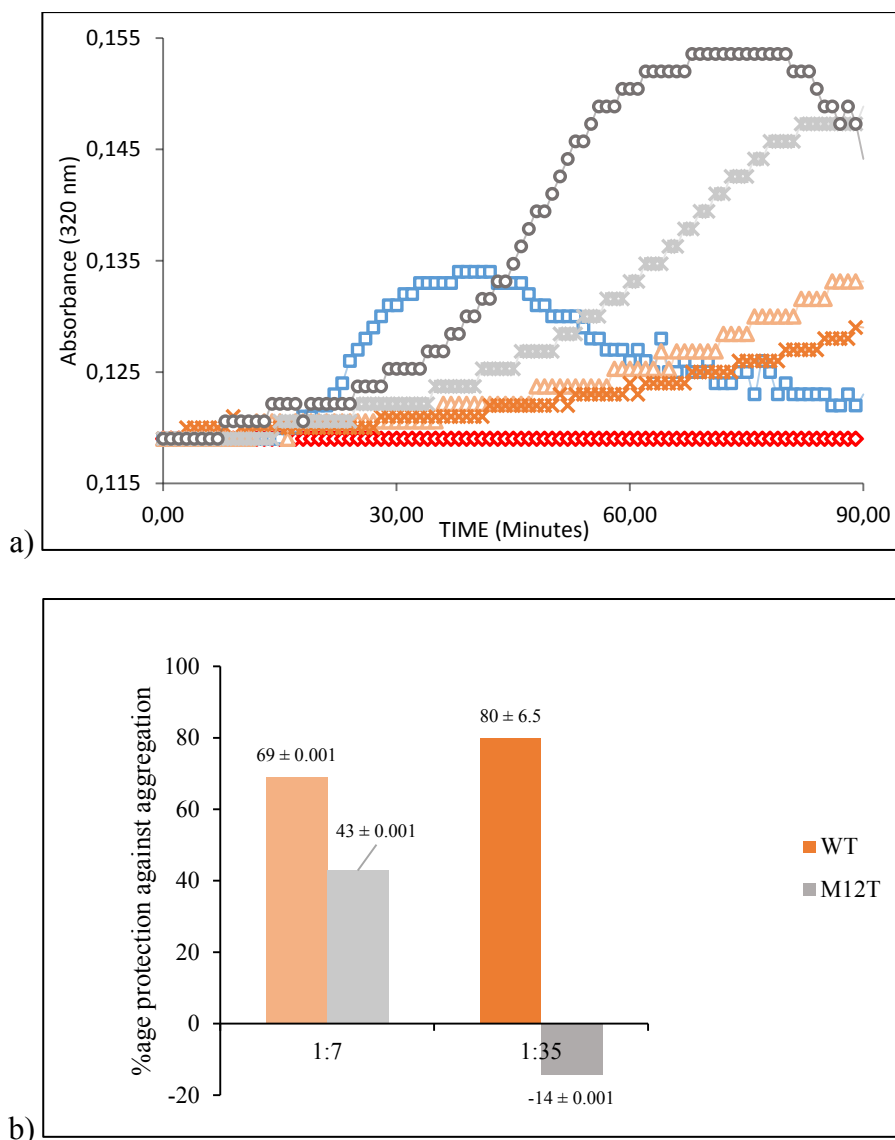


Figure 3.62 Light scattering plot (—◇—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—×—) M12T 1:7 and (—○—) M12T 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Tpv* sHSP 14.3 WT and its mutant variant *Tpv* sHSP 14.3 M12T, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

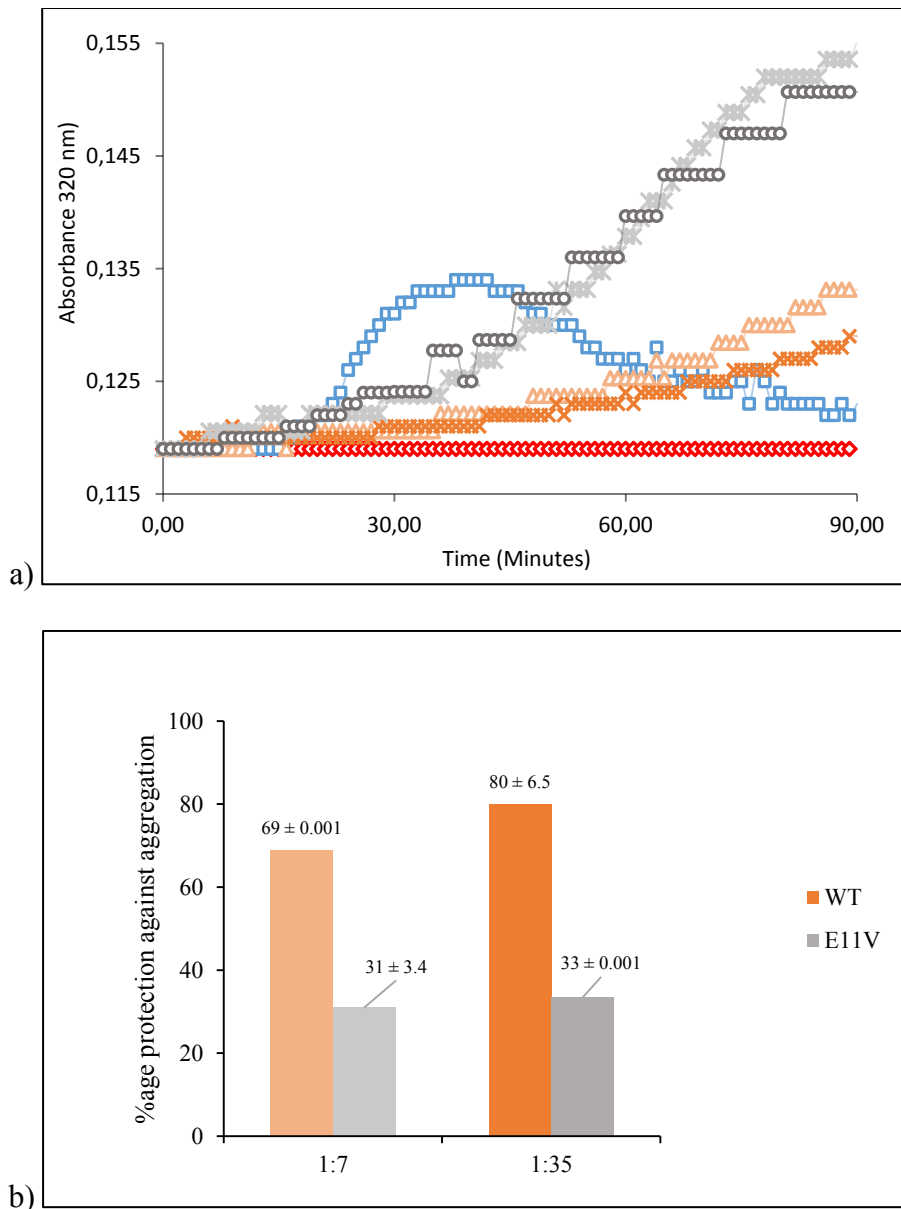


Figure 3.63 Light scattering plot (—◇—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—★—) E11V 1:7 and (—○—) E11V 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Tpv* sHSP 14.3 WT and its mutant variant *Tpv* sHSP 14.3 E11V, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

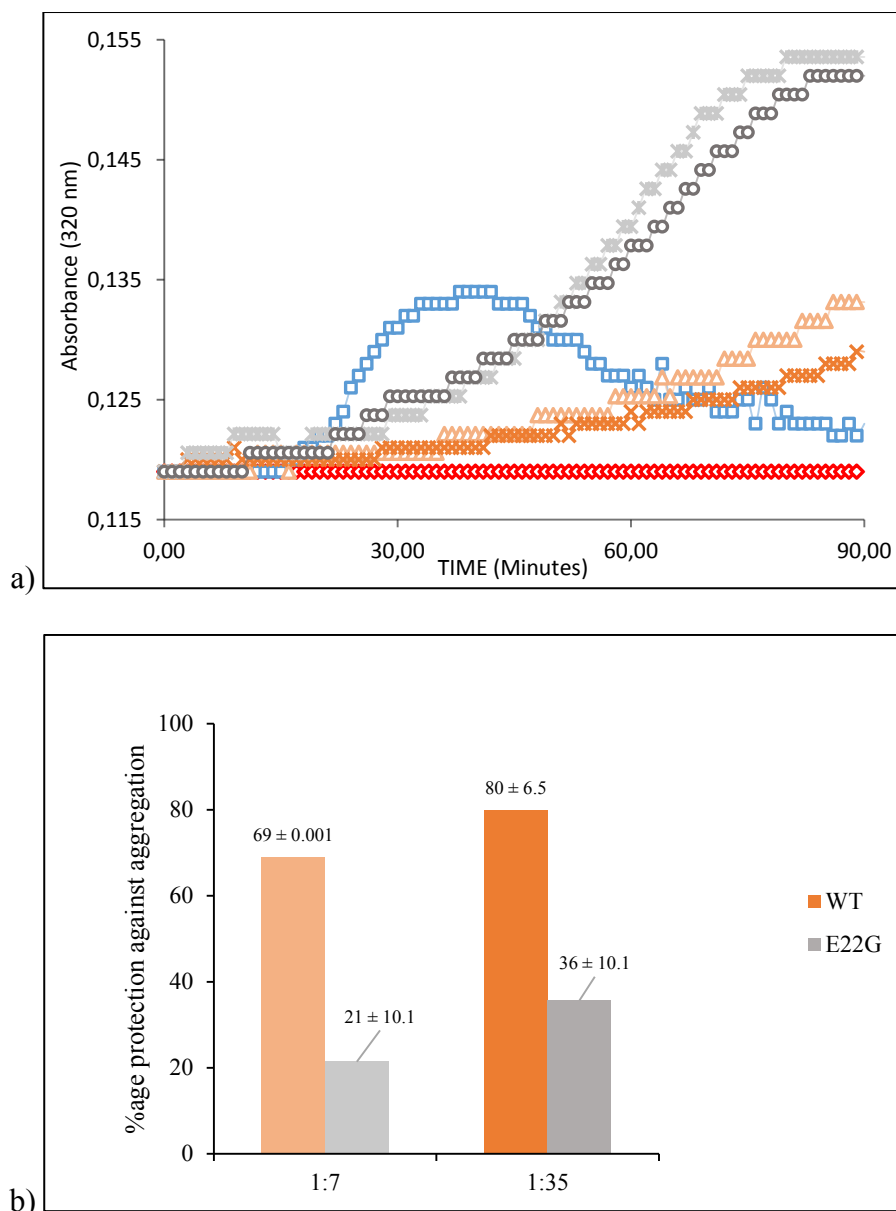


Figure 3.64 Light scattering plot (—◇—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—×—) E22G 1:7 and (—○—) E22G 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Tpv* sHSP 14.3 WT and its mutant variant *Tpv* sHSP 14.3 E22G, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

3.10.3 NTD Distal Part Mutants

It was observed that, when hydrophilicity is increased simultaneously at multiple positions, at the distal end of NTD, (F26YI27T and V23GF26YI27T), it resulted in increase in chaperone activity. These variants displayed similar protection against thermal aggregation of the CS, at both molar ratios studied (1:7 substrate:sHSP molar ratio and 1:35 substrate:sHSP molar ratio) and at similar rates for both double and triple mutant, (Figure 3.65 and 3.66). When compared to WT, the protection effect was higher than WT, and increase in concentration of the chaperone does not affect the protective effect, positively or negatively, significantly.

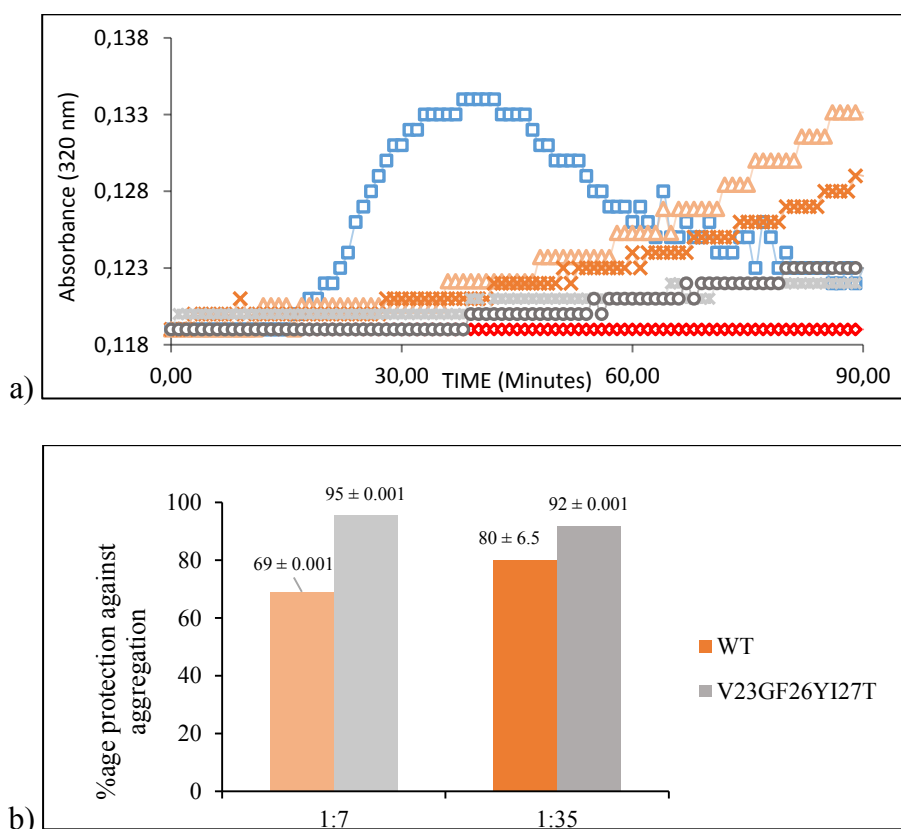


Figure 3.65 Light scattering plot (—♦—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—+—) V23GF26YI27T 1:7 and (—○—) V23GF26YI27T 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *ThpV* sHSP 14.3 WT and its mutant variant *ThpV* sHSP 14.3 V23GF26YI27T, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

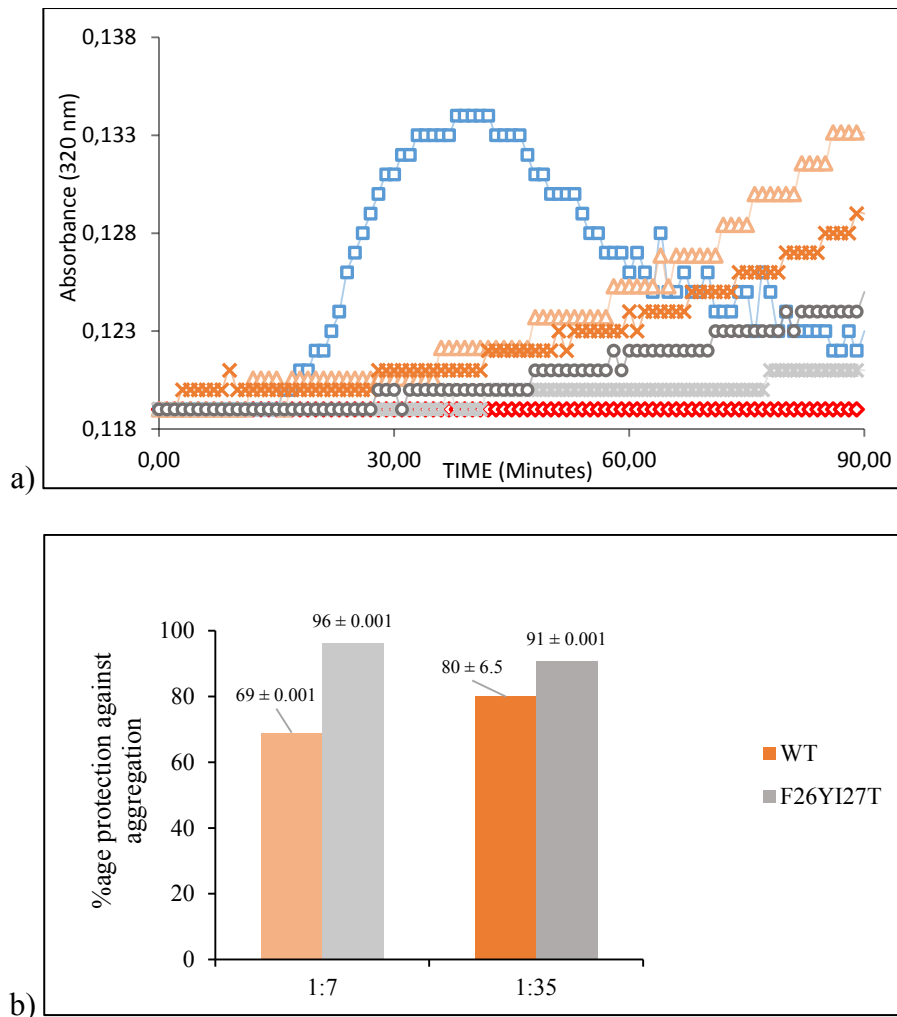


Figure 3.66 Light scattering plot (—♦—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—×—) F26YI27T 1:7 and (—○—) F26YI27T 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Tpv* sHSP 14.3 WT and its mutant variant *Tpv* sHSP 14.3 F26YI27T, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

Moreover, the two variants, which involve mutation at the junction of NTD with ACD, V31G and V31IL33I, both provided protection to the CS at higher temperature, where, this protective effect was higher than the WT at 1:7 substrate:chaperone molar ratios, and lower than WT when the chaperone concentration was increased, (Figure 3.67 and 3.68), suggesting an increase in

hydrophobicity at positions 31 and 33 simultaneously or lowering at position 31 alone, might play a role to increase the chaperone activity of the sHSP.

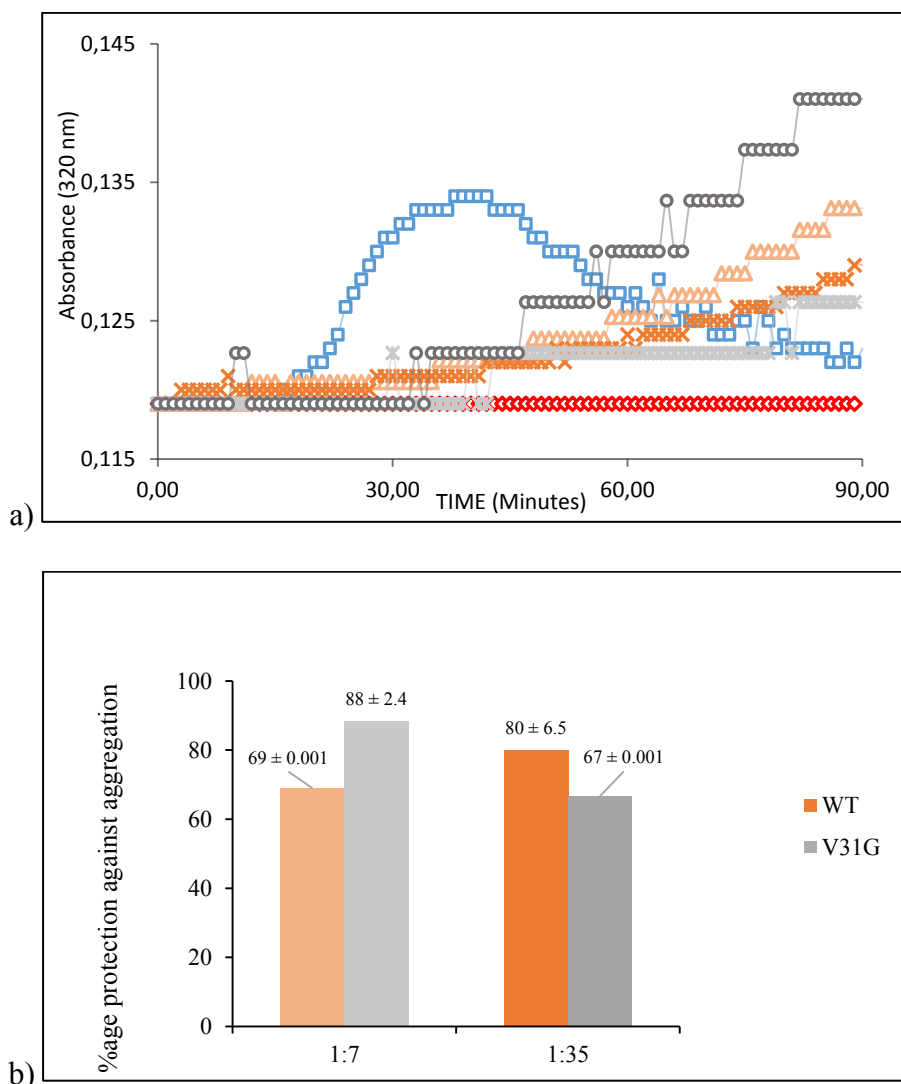


Figure 3.67 Light scattering plot (—◇—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—×—) V31G 1:7 and (—○—) V31G 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Tpv* sHSP 14.3 WT and its mutant variant *Tpv* sHSP 14.3 V31G, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

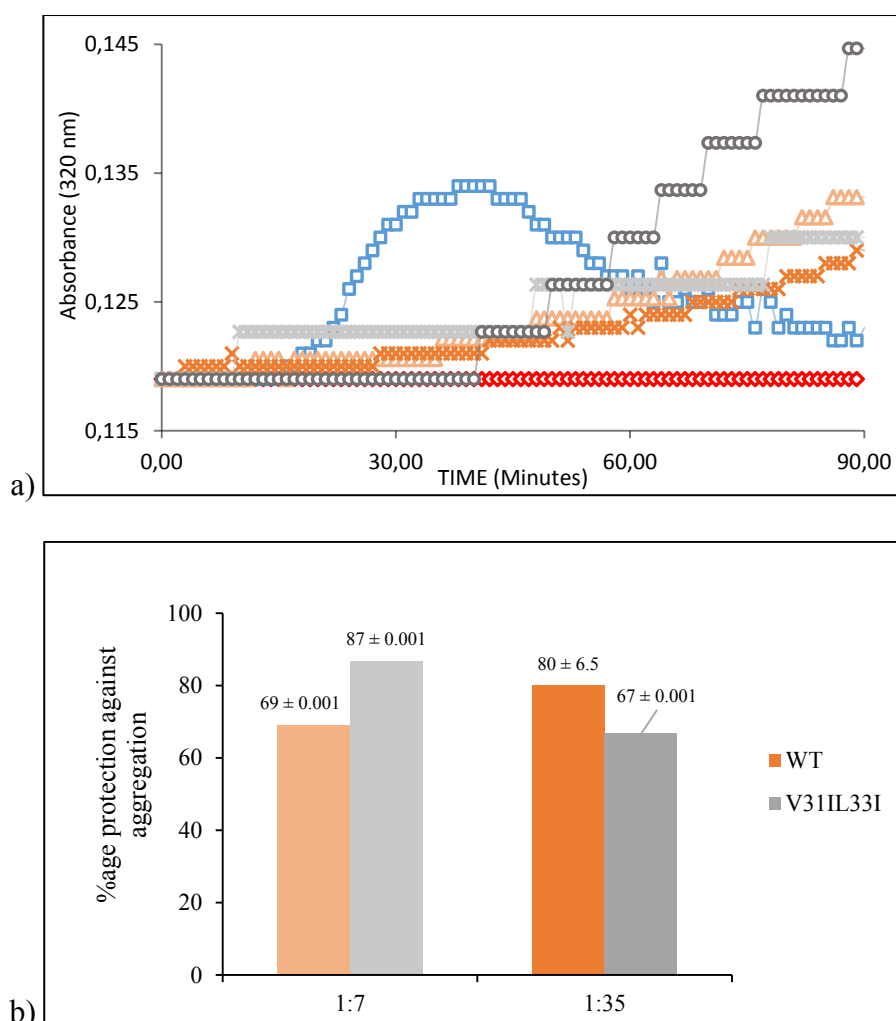


Figure 3.68 Light scattering plot (—◇—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—×—) V31IL33I 1:7 and (—○—) V31IL33I 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Tpv* sHSP 14.3 WT and its mutant variant *Tpv* sHSP 14.3 V31IL33I, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

An interesting fact observed was that in addition to decreasing hydrophobicity at position 31, if hydrophobicity was decreased at position 8 simultaneously, the protective effect of the resultant variant F8YV31G became higher than the WT at 1:7 substrate:sHSP ratio, approaching complete protection, (Figure 3.69), although F8Y mutation alone, had similar effect like WT at the same (1:7) substrate:sHSP ratio, (Figure 3.58). However, when the concentration of the chaperone was

increased, this variant exhibited protective effect similar to WT and F8Y, while the protective effect was more than V31G mutant alone.

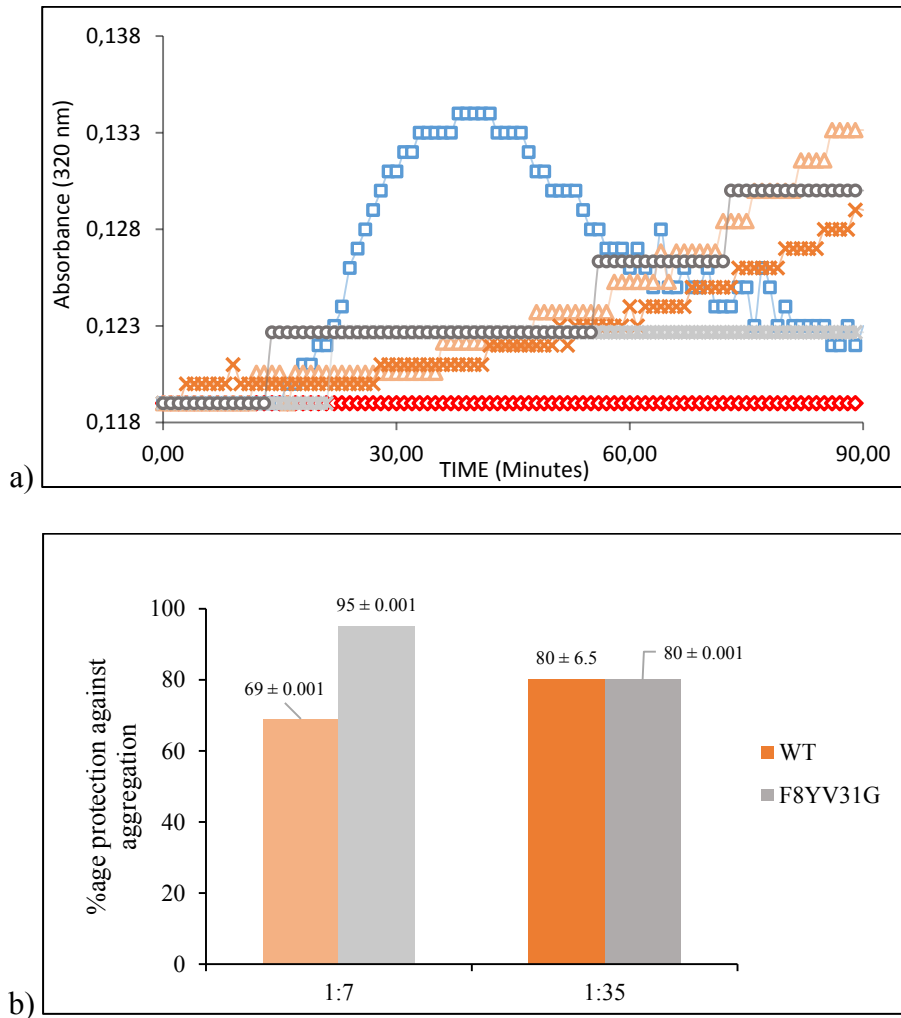


Figure 3.69 Light scattering plot (—◇—) Blank, (—□—) CS control, (—△—) WT 1:7 (—×—) WT 1:35, (—■—) F8YV31G 1:7 and (—○—) F8YV31G 1:35 (a) and bar chart (representing the slopes) (b) of the percentage protection of citrate synthase against heat induced aggregation in the presence of *Tpv* sHSP 14.3 WT and its mutant variant *Tpv* sHSP 14.3 F8YV31G, at 45°C. The percentage protection of control (in absence of any chaperone) is taken as zero in the bar chart. Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

There are reports that the distal part of the NTD, which is close to ACD, is involved in formation and stabilization of dimers (Sun and MacRae, 2005), or tetramers (Weeks *et al.*, 2014) and responsible for the maintenance of the globular structure in some sHSPs (AlIbpA) (Chernova *et al.*, 2020). Thus all the NTD distal part mutants

of *Tpv* sHSP 14.3 that involve decreased or increased hydrophobicity at specific positions led to equivalent or higher chaperone activity than the WT. This may suggest that as a result of such mutations highly stable oligomeric complexes are formed after the interaction of the sHSP with unfolding substrate. This could be the reason of better protection of the CS at suboptimal temperatures by these mutant sHSP variants .

3.11 Analysis of Oligomeric States by Native PAGE

Proteins were prepared in a non-reducing non-denaturing sample buffer, which maintains the proteins' secondary structure and native charge density, and then they are separated according to charge to mass ratio.

The native gel result in the Figure 3.70 shows that while four out of five standard proteins entered the gel, Ribonuclease as well as the two test samples could not enter the gel. The reason may be explained as follows: At the pH of the Running buffer (pH 8.48), the standard proteins Ferritin (pI = 5, 440 KDa), Catalase (pI = 5.42, 232 KDa), Aldolase (pI = 6.4, 158 KDa) and Ovalbumin (pI = 5.19, 44 KDa) were negatively charged. Therefore, they could enter the gel and travel through it, at a speed, depending on their MW. However, Ribonuclease (pI = 9.6, 13.7 KDa) and both test samples, *Tpv* sHSP 14.3 and its I5T variant (both having a theoretical pI = 8.504) were either positively charged (former) or possessed no charge (the latter), For this reason, they were unable to enter the gel and migrate towards anode.

To avoid the problem arising from the pI and net charges of proteins, it was decided to change the gel electrophoresis system and to shift to Blue Native gel-electrophoresis (BN-PAGE). In BN-PAGE system any sample protein prepared carries a net negative charge by virtue of binding with G-250.

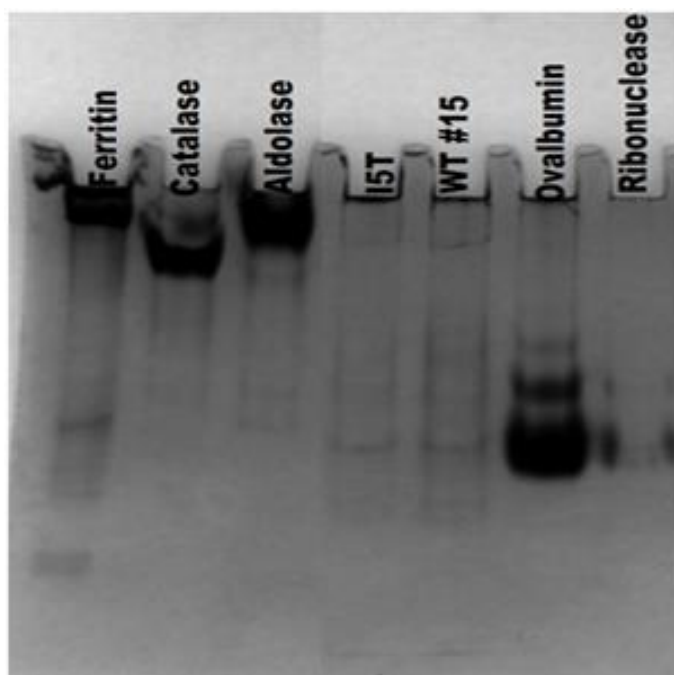


Figure 3.70 Native Gel Electrophoresis of *Tpv* HSP 14.3 WT and its I5T variant. Five pure proteins (Ferritin (440 KDa), Catalase (232 KDa), Aldolase (158 KDa), Ovalbumin (44 KDa) and Ribonuclease (13.7 KDa)) were used as standard proteins.

3.12 Analysis of Oligomeric States by BN-PAGE

3.12.1 Buffer Exchange by Ultrafiltration

First of all, to avoid the protein aggregation in the sample gel during electrophoresis, buffer exchange of protein samples from lysis buffer (containing 30 mM NaCl) to sample buffer (containing 50 mM NaCl) was performed. For this purpose, those samples were selected for which the protein of interest was expressed in very high concentration along with least background or host proteins present. The buffer exchange was performed by virtue of ultrafiltration whereby, desalting as well as concentration of the samples occurred simultaneously. The filter units used for this purpose were Pierce protein concentrator 30K cutoff from Thermo, Amicon 10K cutoff and Centriscart 5K cutoff, from Sartorius. The protocol as mentioned in the respective manual was followed and the resultant supernatant was visualized by

SDS-PAGE, (Figure 3.71). Highly concentrated protein with low background was obtained for almost all samples. The concentration of the proteins were measured by Picodrop in order to load a known and constant amount of protein of each sample, each time on the gel.

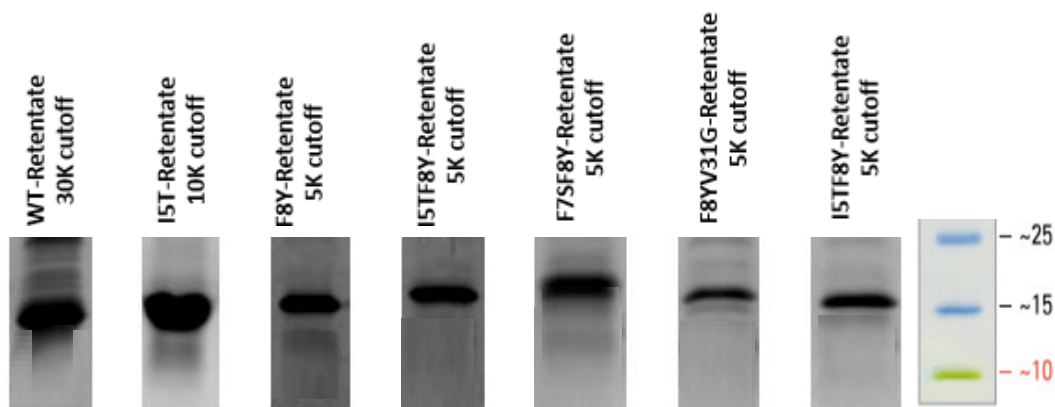


Figure 3.71 SDS-PAGE image of *Tpv* sHSP 14.3 WT and its NTD mutant proteins (I5T, F8Y, I5TF8Y, F7SF8Y, I5TF7SF8Y and F8YV31G) after buffer exchange to sample buffer, by ultrafiltration to be used in BN-PAGE study.

3.12.2 Result of BN-PAGE

The basic principle of BN-PAGE is Coomassie Blue G-250 binds to proteins imparting a negative charge to them, and the separation is according to size in acrylamide gradient gels.

This experiment was performed to resolve the oligomeric species of *Tpv* sHSP 14.3 WT and its above mentioned heat stable N-terminal domain variants (I5T, F8Y, I5TF8Y, F7SF8Y, I5TF7SF8Y and F8YV31G). The results are shown in Figure 3.72 and 3.73. The BN-PAGE picture of the standard proteins is shown in appendix D, Figure 13. In general, the proteins showed presence of two groups of species: one having smaller species, and the other consisting of larger order oligomers.

It was observed that, similar to WT, variants such as I5T, F8Y, I5TF8Y, F7SF8Y and I5TF7SF8Y, possess fewer smaller species at room temperature, *i.e.*, between monomer and/or dimer and 12-mer. However, the variant F8YV31G possessed many smaller species, namely, dimer and species smaller than 12-mer that might be

tetramer and hexamer. On the other hand, except F7SF8Y, the other variants and the WT, showed presence of larger oligomeric species around 60-mer or beyond. However, F7SF8Y exhibited presence of 24-mer as a dominant oligomeric specie, besides dimeric form. The single mutant I5T and the triple mutant I5TF7SF8Y display oligomeric specie spread on the gel, namely the 60-mer, 36-mer and dimer and lack 24-mer oligomer. All the rest of the mutants and the *Tpv* sHSP 14.3 WT show a clear band corresponding to the molecular weight equivalent to 24-mer protein, which is a striking characteristic multimeric form of the most of the archaeal sHSPs (Haslbeck *et al.*, 2008). The analysis of the oligomeric states after heat stress was also studied for comparison. For this purpose, the samples were heated (70°C for 10 minutes) before loading on to BN-PAGE. It was observed that, with the exception of F7SF8Y and F8YV31G variants, the variety of the larger oligomeric species mostly increased while status of smaller oligomers remained unchanged after heating. The reason could be that the sHSPs at room temperature may be found in the form of very high order oligomers. When heated, these large protein assemblies can be converted to slightly lower order oligomers, and thus their distribution all over the gel can be observed.

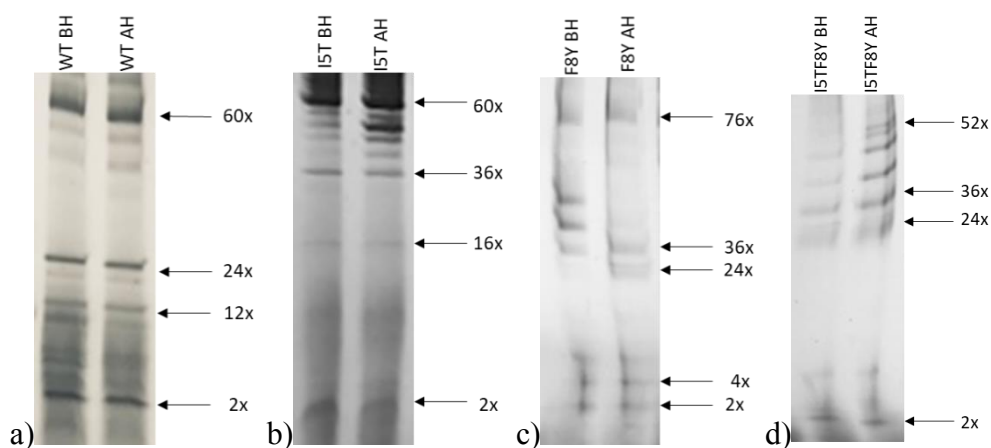


Figure 3.72 Resolution of the oligomeric forms of the *Tpv* sHSP 14.3 WT (a) and its various NTD variants; *Tpv* sHSP 14.3 I5T (b); *Tpv* sHSP 14.3 F8Y (c) and *Tpv* sHSP 14.3 I5TF8Y (d), under native conditions by BN-PAGE. For the calculation of molecular weight of unknown oligomers, ferritin (440 KDa), Catalase (232 KDa), Aldolase (158 KDa) and Ribonuclease (13.7 KDa) were used as standards. BH: Before heat, AH: After heat.

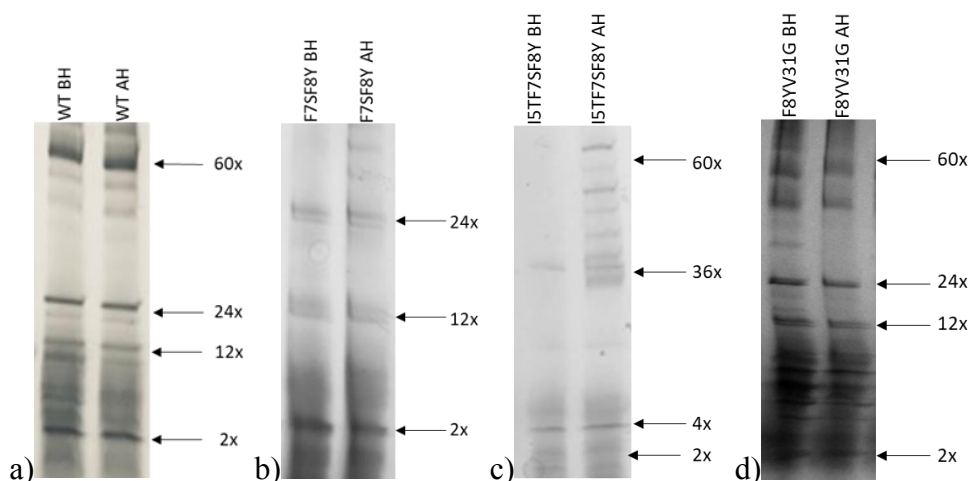


Figure 3.73 Resolution of the oligomeric forms of the *Tpv* sHSP 14.3 WT (a) and its various NTD variants; *Tpv* sHSP 14.3 F7SF8Y (b); *Tpv* sHSP 14.3 I5TF7SF8Y (c) and *Tpv* sHSP 14.3 F8YV31G (d) under native conditions by BN-PAGE. For the calculation of molecular weight of unknown oligomers, ferritin (440 KDa), Catalase (232 KDa), Aldolase (158 KDa) and Ribonuclease (13.7 KDa) were used as standards. BH: Before heat, AH: After heat.

3.13 Bioinformatics Section (3-D Molecular model Study)

3.13.1 Generation of Monomer and Dimer structure

First of all the monomer structure of *Tpv* sHSP 14.3 was generated using homology modelling with the help of the software Easy Modeller 4.0, as described in Chapter 2. The monomer structure in ribbon model is displayed in Figure 3.74. The grey and off-white colors indicate the NTD and CTD respectively, while the dark black middle part is the, structure and sequence wise, highly conserved domain, the ACD of the *Tpv* sHSP 14.3. The terminal points, beginning of NTD and the CTD, as a whole are unstructured. Most of the NTD acquires a helical form while the junction of NTD and ACD is characterized by the presence of a coil like structure. The ACD has eight beta-sheets, numbered from $\beta 2$ to $\beta 9$. All the β -sheets are stacked and present in two parallel structures where $\beta 2$, $\beta 3$, $\beta 8$ and $\beta 9$ are on one side while $\beta 4$, $\beta 5$ and $\beta 7$ are on the other one. The $\beta 6$ forms a coil like structure that overlaps the NTD, and in this form contributes to dimer interaction. In the dimer, the $\beta 6$ of one

monomer overlaps the β -sheets stack of another monomer and form interactions with the $\beta 2$ of the other monomer and *vice versa*, Figure 3.75. All the targeted positions for mutagenesis in this study for either single, double or triple mutations are indicated as stick model and labelled in the Figure 3.76. It shows that most of the mutations we designed are present either in the beginning or towards the carboxyl end of the NTD. This is due to the fact the, the NTD is more hydrophobic toward its amino and carboxyl ends, as compared to the central section. The middle of NTD contains four out of the total five charged residues of the NTD. The residue isoleucine at position 5 is present in the initial unstructured part of NTD. The two phenylalanine residues at position 7 and 8 form the first turn of the helical structure while E11 and M12 contribute to the second turn of the helix. The residues E22 and V23 are located at the last turn of the NTD helix while F26 and I27 form the coil-like structure where the NTD joins to the ACD. Valine at position 31 is present in the initial part of the $\beta 2$ sheet in the ACD (Figure 3.76).

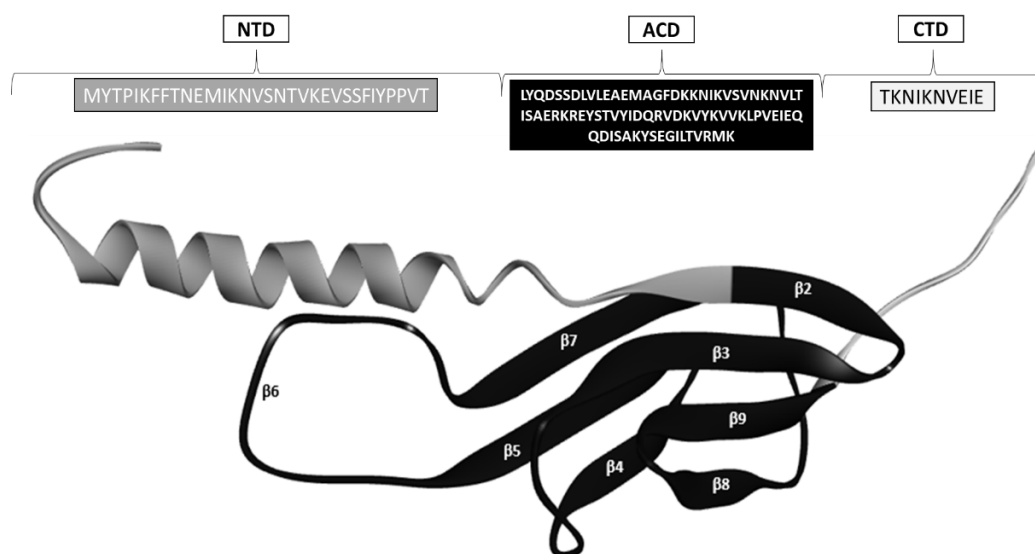


Figure 3.74 Monomer structure developed using Easy Modeller 4.0 of *Tpv* sHSP 14.3 WT. The NTD is highlighted in grey color while the CTD is present as off-white. All β -sheets of the ACD are labelled in white against a dark background of the sheets itself. The 124 amino acid sequence written, was experimentally first determined in our lab (Kocabiyik and Aygar, 2012), while National Centre for Biotechnology Information (NCBI) annotated from genome sequencing.

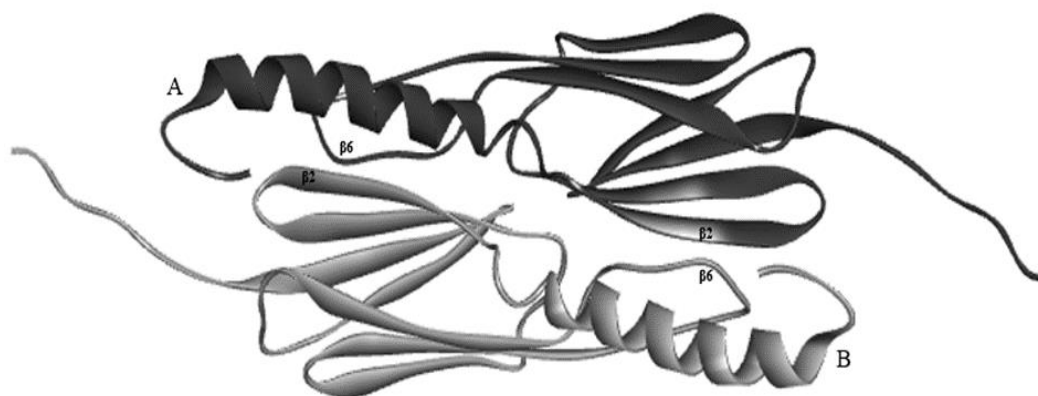


Figure 3.75 Top view of a dimer of *Tpv* sHSP 14.3 WT indicating the monomer-monomer interaction between the $\beta 2$ and $\beta 6$ of each monomer. The dark color represent one monomer, labeled A, while the second monomer is represented in grey color and labelled as B.

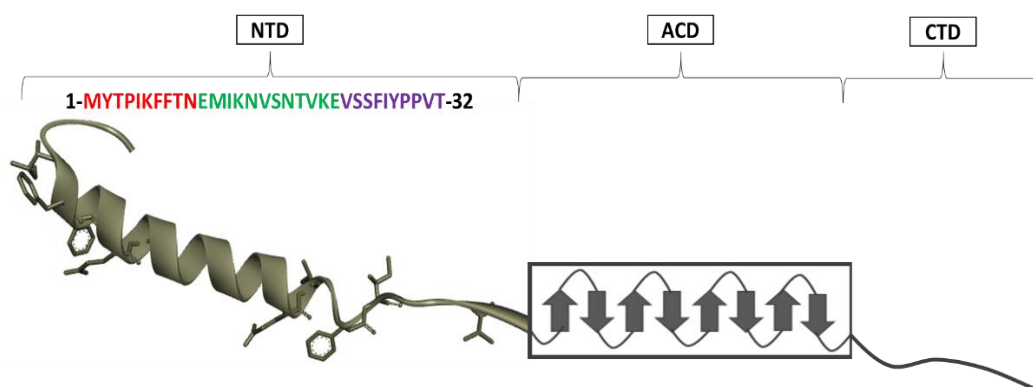


Figure 3.76 Mutation points in the NTD of *Tpv* sHSP 14.3 WT shown as stick model on a ribbon structure. The ACD and CTD are present as schematic diagrams. The NTD is divided into proximal (indicated by red letters), middle (indicated by green letters) and distal part (indicated by purple letters).

3.13.2 Inter and Intra-molecular interactions in a dimer

In this section, the inter-molecular and intra molecular interactions of the mutants at the mutation sites were studied and compared against the WT for those specific positions, within a dimer. This was done to identify the reasons underlying the similar or different behavior of the mutants from the WT. Below is a brief description of the bonds lost or gained in the mutants. Related Tables with detailed information

are given in appendix (where blue is for one monomer and pink for the other monomer in a dimer).

3.13.2.1 Proximal NTD mutants

3.13.2.1.1 I5T variant

For the position five, where isoleucine is present in WT, substitution was done to a less hydrophobic threonine. In the WT, I5 forms two intramolecular hydrogen bond, within a dimer, one in each monomer with threonine at position 3. When the isoleucine is replaced with threonine, these hydrogen bonds are lost and new intramolecular hydrogen bonds were formed with phenylalanine at position 8, (Figure 3.77 and 3.78; Appendix E, Table 2). This leads to the turn of the coil away from the center of the dimer (Figure 3.78).

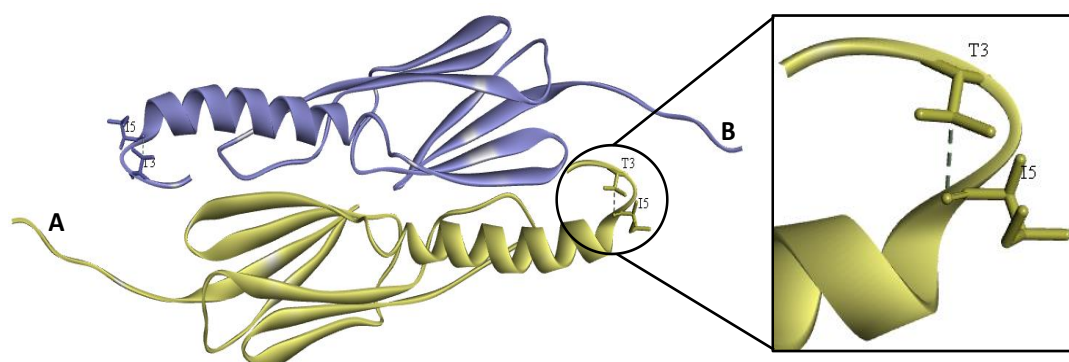


Figure 3.77 3-D image of dimer of *Tpv* sHSP 14.3 WT showing the intramolecular hydrogen bond formed by Ile5 with Thr3 in the maintenance of the proximal NTD coil structure/shape.

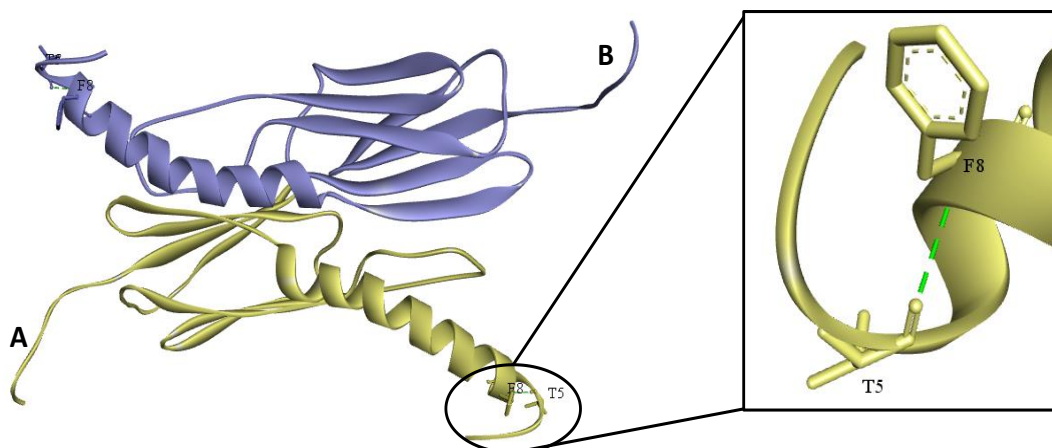


Figure 3.78 3-D image of dimer of *Tpv* sHSP 14.3 I5T variant showing the intramolecular hydrogen bond formed by Thr5 with Phe8 leading to the twisting of coil away from the center of the dimer.

3.13.2.1.2 F8Y variant

In the WT (Figure 3.79), the phenylalanine at position 8 forms one hydrophobic interaction with the methionine at position 12, one electrostatic interaction with glutamic acid at position 11 and two hydrogen bonds, one with Met12 and the other with Glu11, in the monomer, within a dimer. When F8 was replaced by Y (Figure 3.80), the intramolecular hydrophobic and electrostatic interactions mentioned above are lost and similar interactions are formed with new amino acids; isoleucine 5 (hydrophobic interaction) and methionine 1 (Electrostatic interaction), (Appendix E, Table 3). However, for the intramolecular hydrogen bonds, one of them is retained in the mutant with a slight decrease in bond length, and a new one is formed with Met1, suggesting a shift in the helix formed.

3.13.2.1.3 I5TF8Y variant

When the above mentioned single mutations are combined in a mutant by replacement of I5 and F8, simultaneously, with hydrophilic residues threonine and tyrosine, respectively. The result is complete loss of intra molecular hydrophobic interaction at position 8 in both the monomers, (Appendix E, Table 4). While the

intramolecular electrostatic interactions are retained as WT (opposed to F8Y alone) and the number of intramolecular hydrogen bonds increase to more than double (seven in mutant as compared to three in WT) in the mutant's monomer, within a dimer, at the positions of mutations (Figure 3.81 and 3.82).

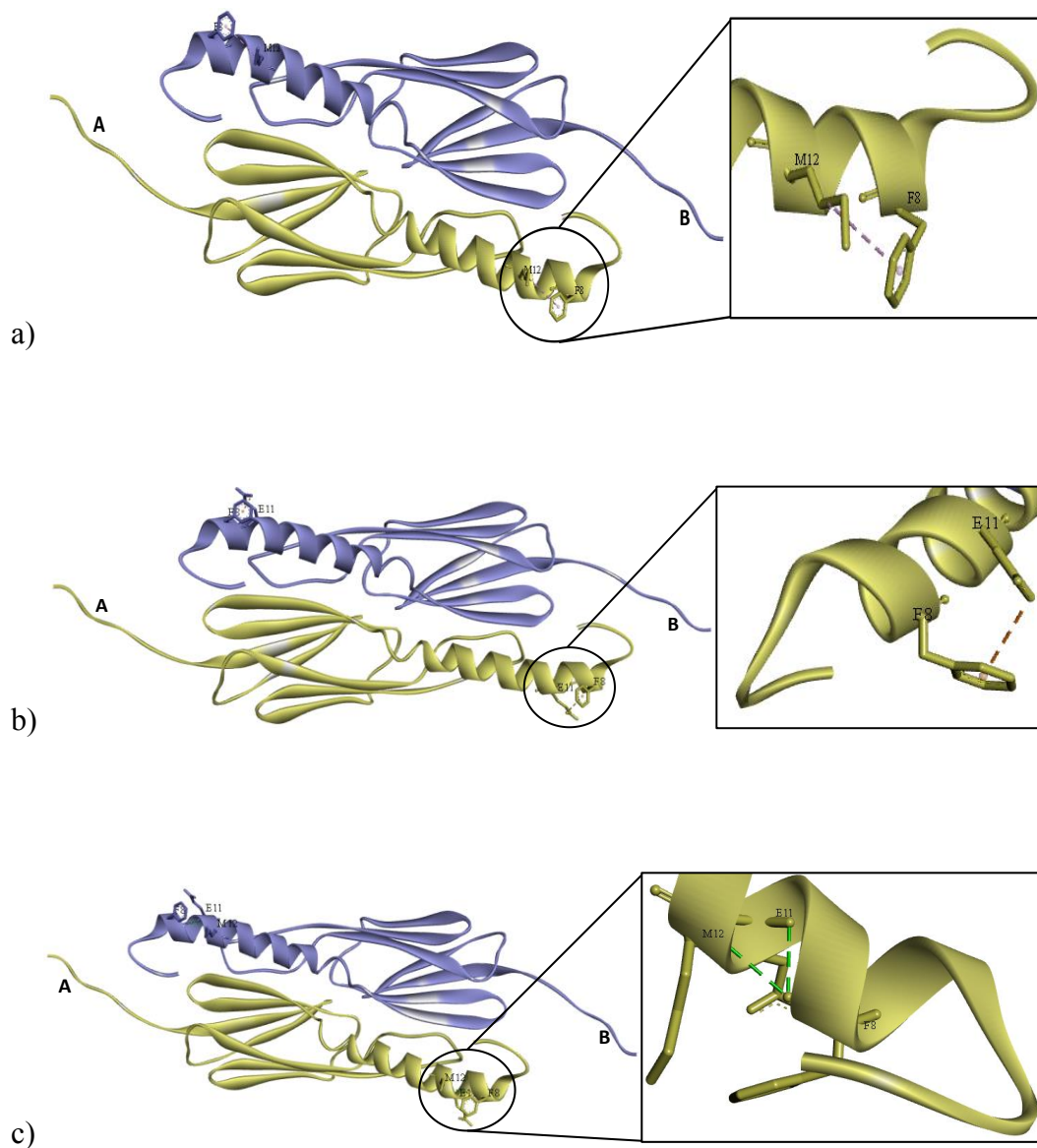


Figure 3.79 3-D images of dimer of *Tpv* sHSP 14.3 WT showing the (a) hydrophobic, (b) electrostatic and (c) hydrogen bonds formed by F8 in the formation and maintenance of the first helical turn.

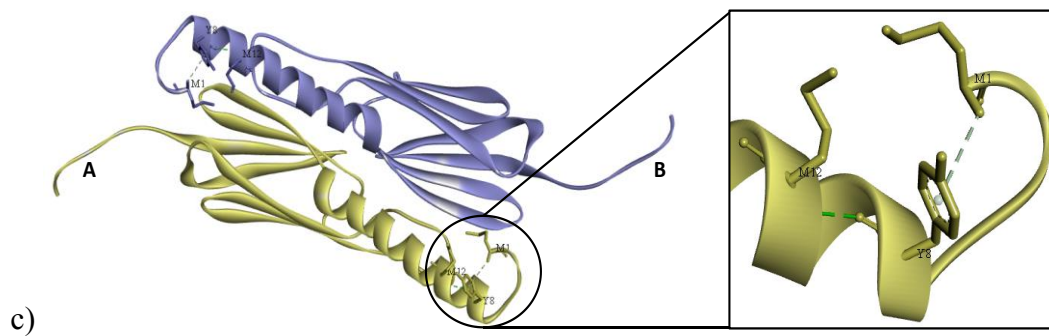
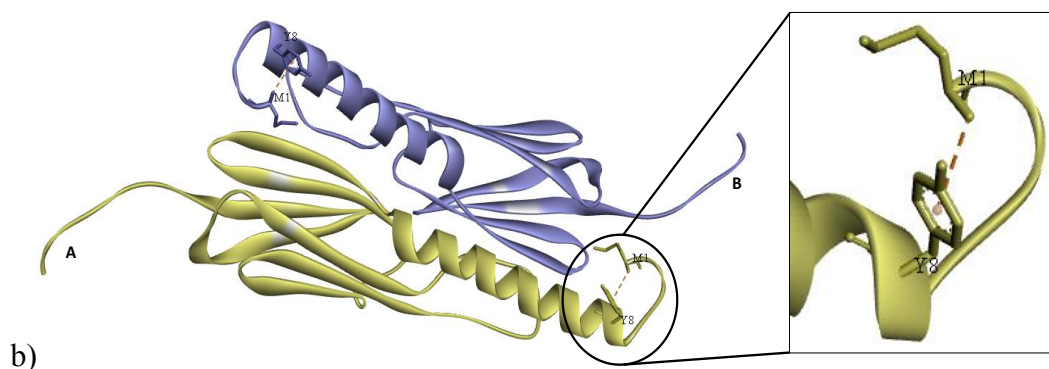
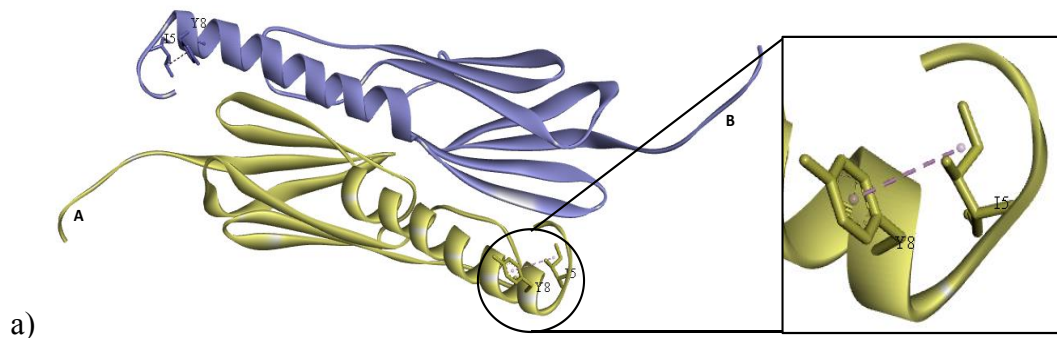


Figure 3.80 3-D images of dimer of *Tpv* sHSP 14.3 F8Y variant showing the (a) hydrophobic, (b) electrostatic and (c) hydrogen bonds formed by Tyr8 and the opening/disordering of the first helical turn.

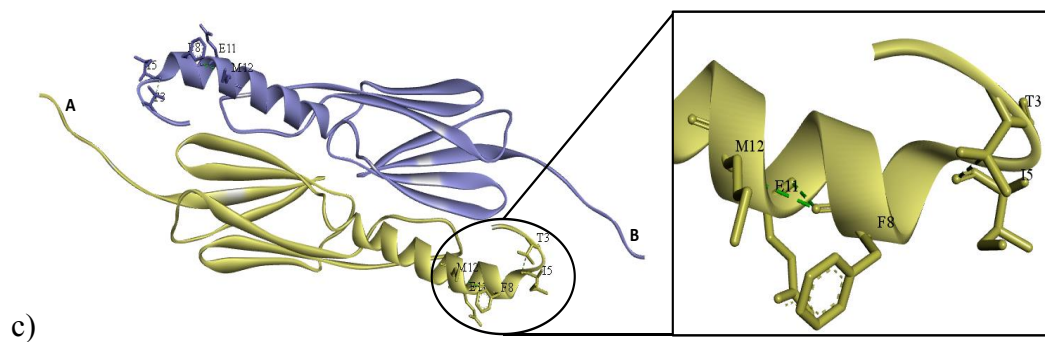
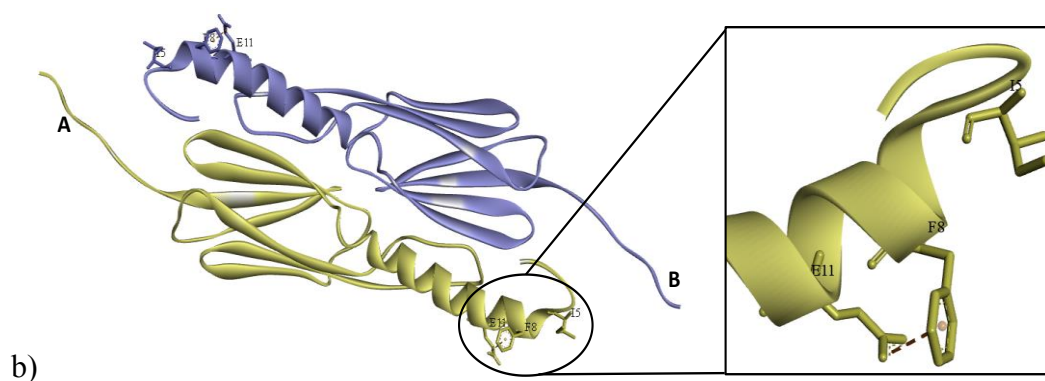
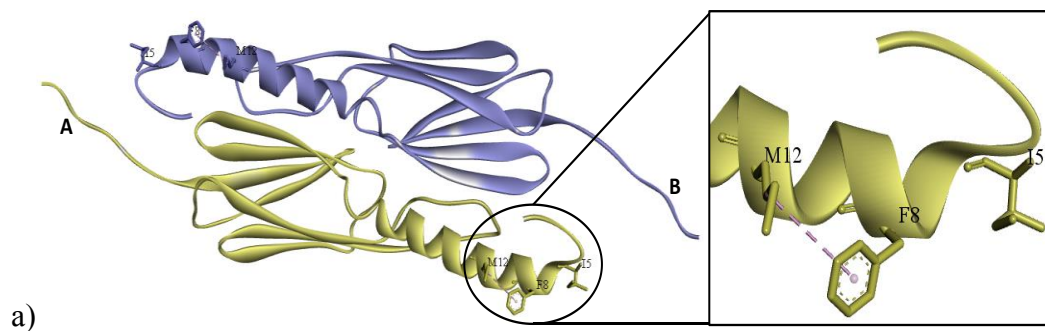


Figure 3.81 3-D images of dimer of *Tpv* sHSP 14.3 WT showing the (a) hydrophobic, (b) electrostatic and (c) hydrogen bonds formed by I5 and F8.

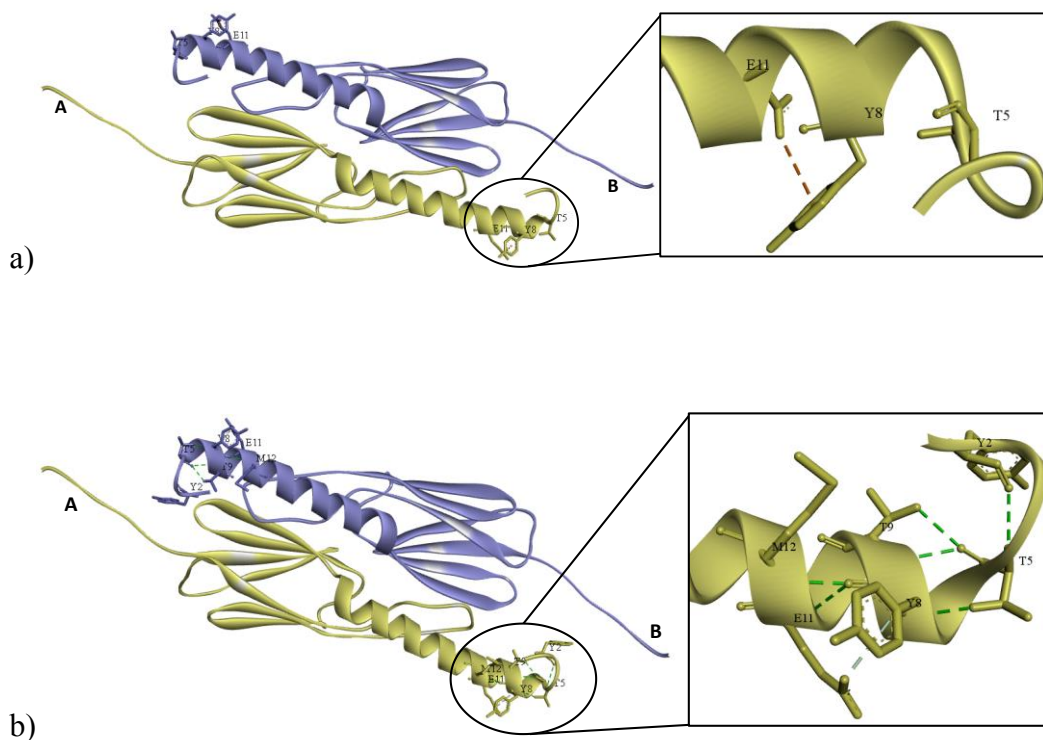


Figure 3.82 3-D images of dimer of *Tpv* sHSP 14.3 I5TF8Y variant displaying the (a) electrostatic and (b) hydrogen bonds formed by Thr5 and Tyr8 in the double mutant. The coil of proximal NTD evidently shows a different twist as compared to the WT.

3.13.2.1.4 F7SF8Y variant

When two adjacent highly hydrophobic amino acids (phenylalanine residues) are replaced with hydrophilic serine and less hydrophobic tyrosine respectively, at position 7 and 8, the first turn of the alpha helix of NTD becomes more hydrophilic.

This double mutation results in loss of electrostatic interactions in the mutant as well as one of the hydrogen bonds of phenylalanine at position 8, (Appendix E, Table 5). Besides, there is formation of an extra hydrophobic interaction (similar to F8Y single mutant), when compared to WT (Figure 3.83 and 3.84).

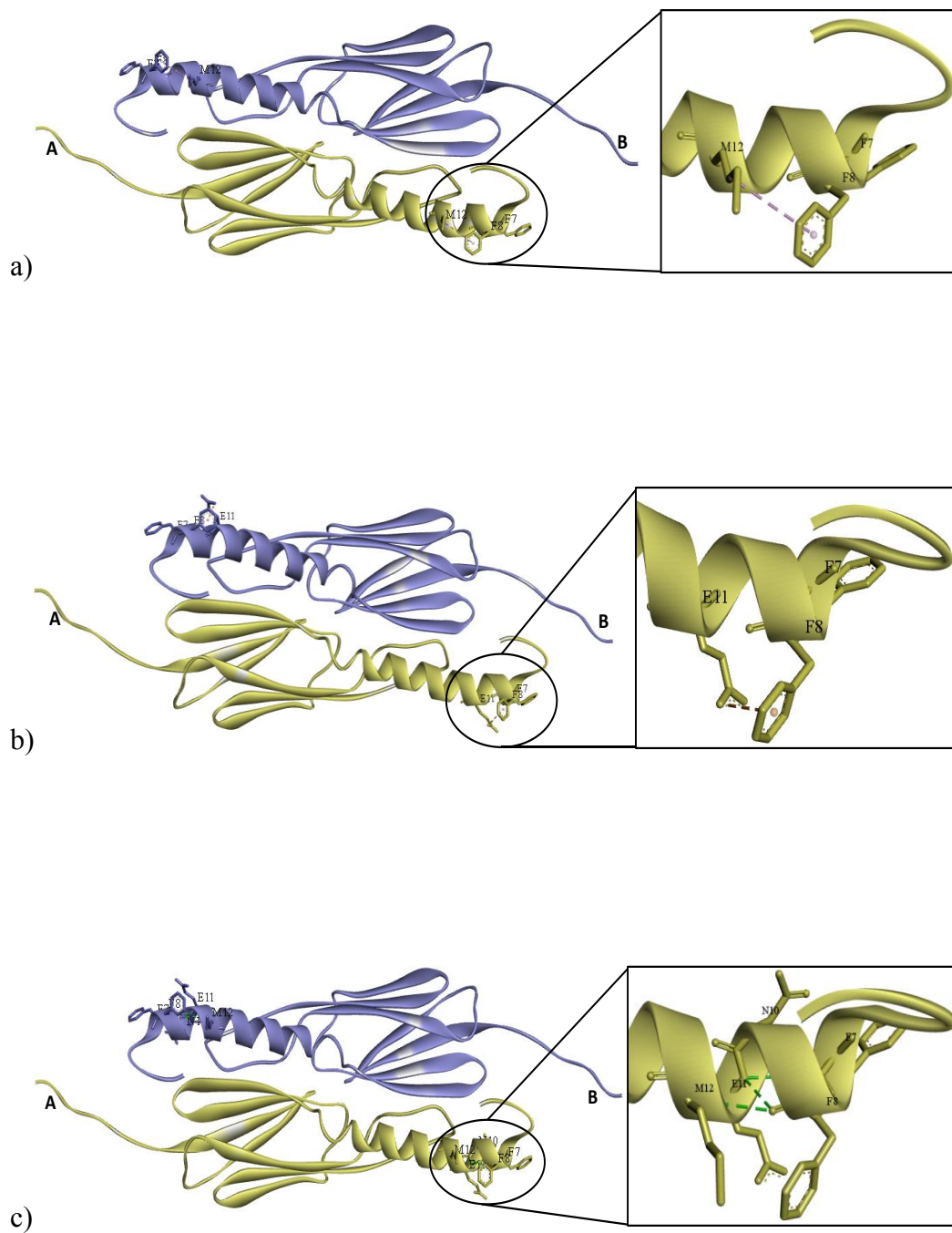


Figure 3.83 3-D images of dimer of *Tpv* sHSP 14.3 WT showing the (a) hydrophobic, (b) electrostatic and (c) hydrogen bonds formed by phenylalanine residues present at position 7 and 8.

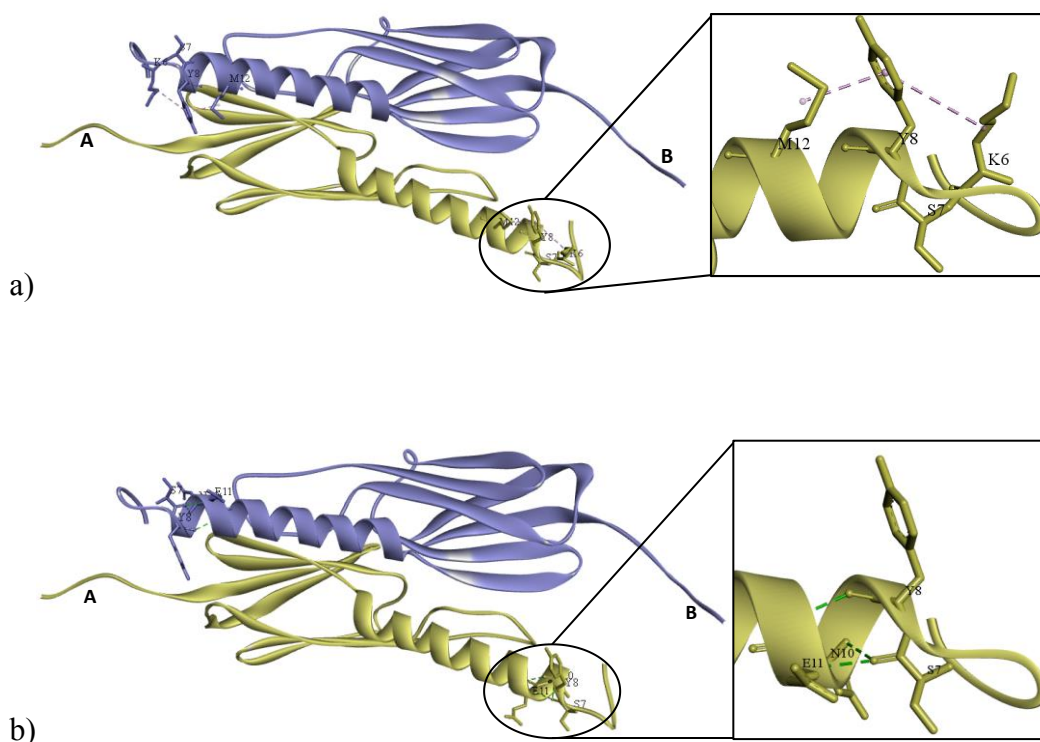


Figure 3.84 3-D images of dimer of *Tpv* sHSP 14.3 F7SF8Y variant displaying the (a) hydrophobic and (b) hydrogen bonds formed by Ser5 and Tyr8 in the double mutant.

3.13.2.1.5 I5TF7SF8Y variant

Substituting the three positions, I5, F7 and F8, simultaneously with hydrophilic amino acids, retains the hydrophobic interactions (opposed to I5TF8Y) but with shorter distance as compared to that of WT. On the other hand, electrostatic interactions of Phe8 are lost (similar to F7SF8Y), (Appendix E, Table 6).

Also, there is an additional hydrogen bond formed within each monomer of the variant dimer (Figure 3.85 and 3.86). This combination of a set of three simultaneous mutations has led to increase in the disorder region of the NTD as evident from the increased length of the coil (Figure 3.86), and as also mentioned in IDR section. Moreover, the coil moves away from the center of the dimer, as opposed to WT.

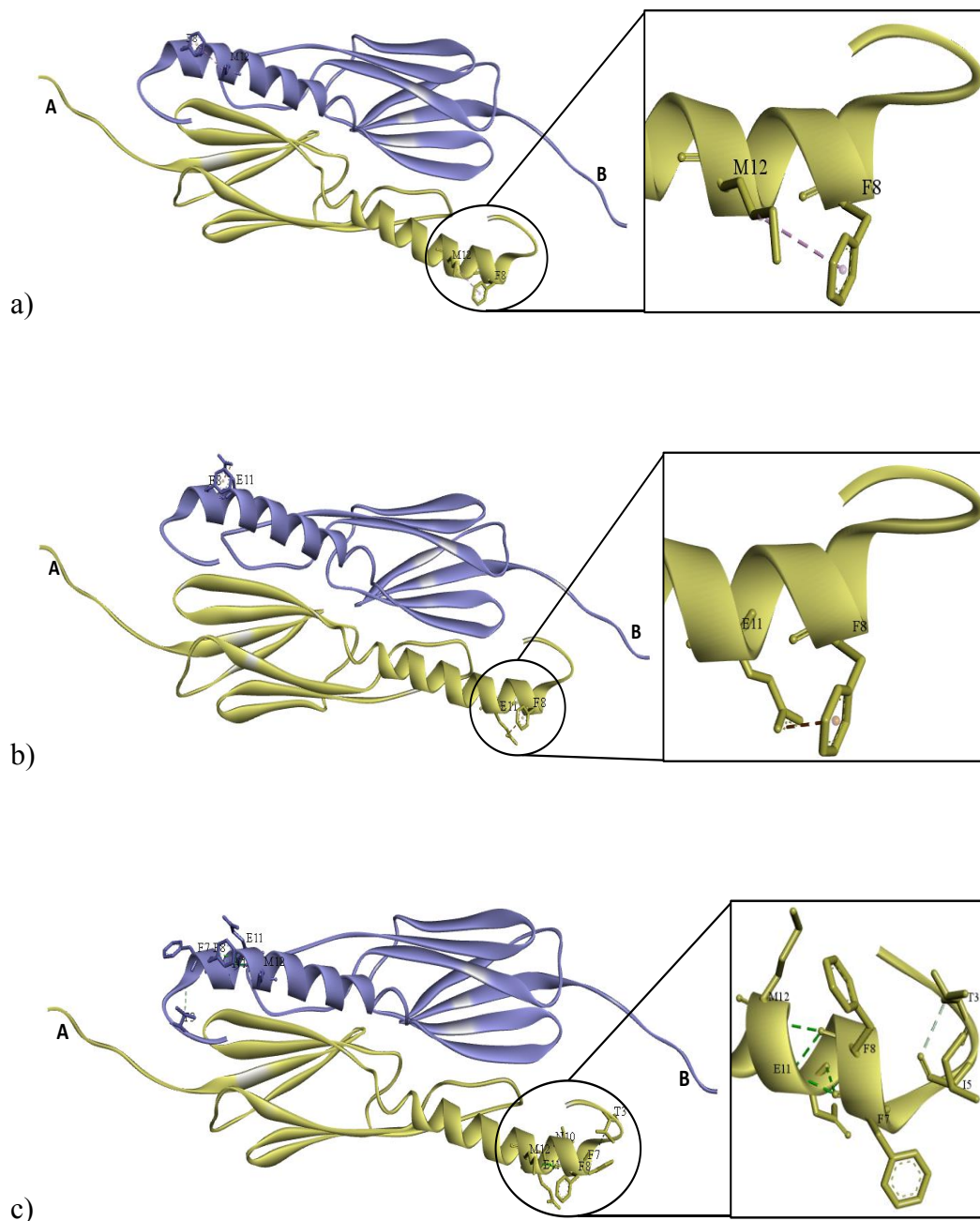


Figure 3.85 3-D images of dimer of *Tpv* sHSP 14.3 WT showing the (a) hydrophobic, (b) electrostatic and (c) hydrogen bonds formed by threonine and phenylalanine residues present at position 5, 7 and 8, respectively. The two monomers are represented by different colors, bronze (monomer A) and blue (monomer B).

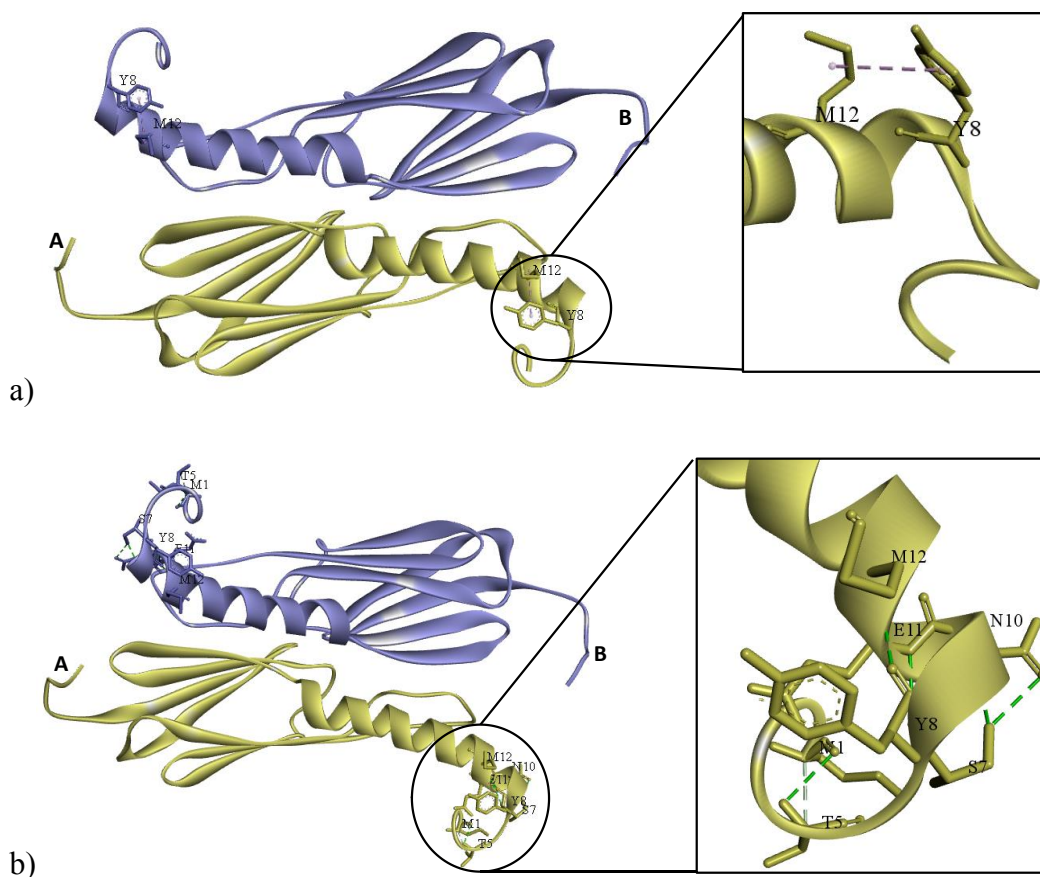


Figure 3.86 3-D images of dimer of *Tpv* sHSP 14.3 I5TF7SF8Y variant displaying the (a) hydrophobic and (b) hydrogen bonds formed by Thr5, Ser5 and Tyr8 in the triple mutant. The increase in length of coil and its twist away from center of dimer is evident.

3.13.2.2 Middle NTD Mutants

3.13.2.2.1 E11V variant

When E11 is replaced with hydrophobic valine, the intramolecular electrostatic interactions which E11 makes with phenylalanine at position 8 (Figure 3.87a) are lost in both the monomers, in the variant dimer. However, the intramolecular hydrogen bonds (5 in each monomer) are more or less, same in both, the WT (Figure 3.87b) and mutant (Figure 3.88), but the distance of these hydrogen bonds is increased in the mutant, suggesting that the helix might become loose or open, at this position, (Appendix E, Table 7). The change in the bond length might be the reason

of the bending of the helical structure towards the center of the dimer, when considering its top view (Figure 3.88).

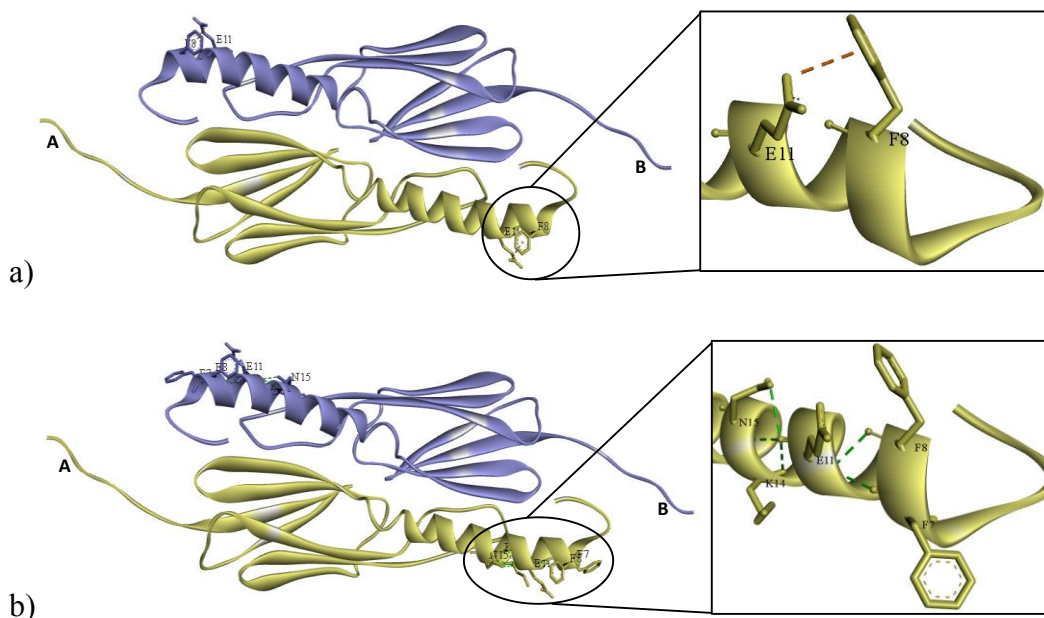


Figure 3.87 3-D images of dimer of *Tpv* sHSP 14.3 WT displaying the (a) electrostatic and (b) hydrogen bonds formed by E11 within a monomer in a dimer.

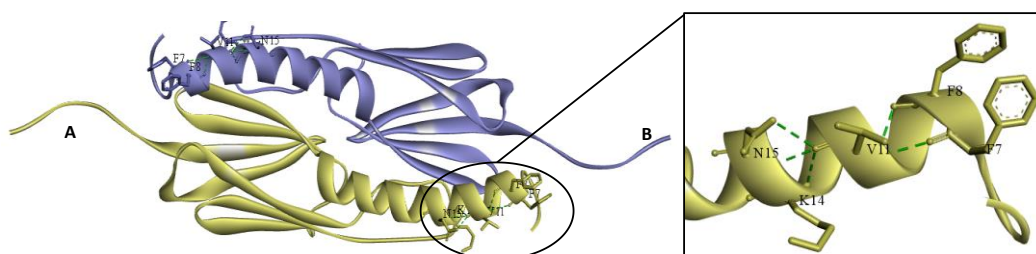


Figure 3.88 3-D image of dimer of *Tpv* sHSP 14.3 E11V variant displaying the hydrogen bonds formed by Val11 in the E11V mutant.

3.13.2.2.2 E22G variant

Considering the interactions, which glutamic acid at position 22 shows with neighboring amino acids, 6 intramolecular hydrogen bonds are formed by E22 in WT (Figure 3.89) with 2 hydrogen bonds with serine at position 25, in each monomer, (Appendix E, Table 8). When this position (E22) is made hydrophobic by replacing it with glycine, one of the hydrogen bond with serine 25 is lost in each

monomer (Figure 3.90), which might be due to change in the nature of the amino acid, at position 22, from negatively charged Glu to hydrophobic Gly. This mutation (and probably the loss of one hydrogen bond with Ser25), in contrast to E11V, may cause a structural perturbation that bends the whole helix away from the center of the dimer (top view, Figure 3.90)

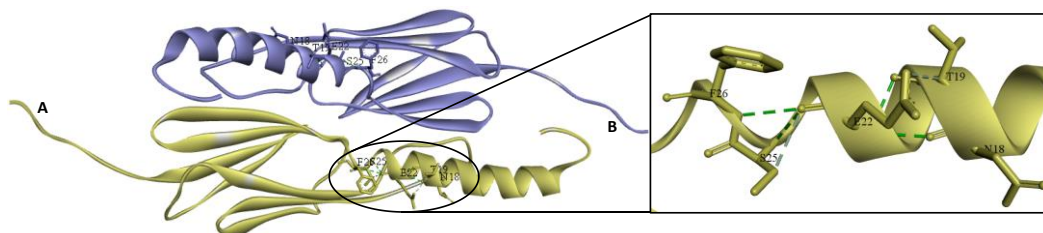


Figure 3.89 3-D images of dimer of *Tpv* sHSP 14.3 WT displaying the hydrogen bonds formed by Glu22 in each monomer within a dimer.

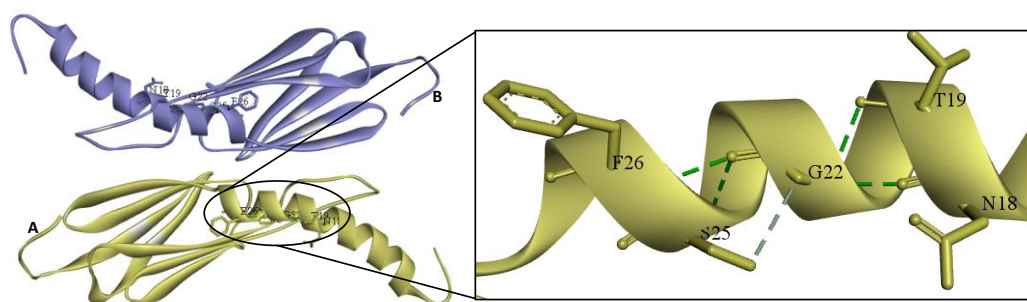


Figure 3.90 3-D images of dimer of *Tpv* sHSP 14.3 E22G variant displaying the hydrogen bonds formed by Gly22 in the E22G mutant.

3.13.2.2.3 M12T variant

Considering the intramolecular interactions, methionine at position 12 shows two hydrophobic interactions with phenylalanine at position 8, one in each monomer of the dimer. These interactions disappear, (Appendix E, Table 9), once methionine is replaced with hydrophilic threonine. However, the hydrogen bonds which Met12 forms are also formed by Thr12, in the M12T mutant, with similar bond-length. Thus M12 in the WT shows two non-covalent bond interactions with F8 of the same monomer; a hydrophobic bond (Figure 3.91a) and a hydrogen bond (Figure 3.91b). In the variant (Figure 3.92), since the hydrophobic interaction between residues 8

and 12 is lost, the interaction between residues 8 and 12 decreases/weakens which (might result in a loose helix at this position and) may lead to formation of lower order oligomers and mostly dimers. They could be more prone to bind the cellular proteins and be removed in centrifugation. This could be the apparent reason for less concentration of expressed protein in the cell free extract before any heat treatment. The bend of the helix in the variant M12T, is similar to E22G, away from the dimer center. In addition to this, the mutation M12T leads to increase in the length of the disordered coil in the proximal NTD (Figure 3.92).

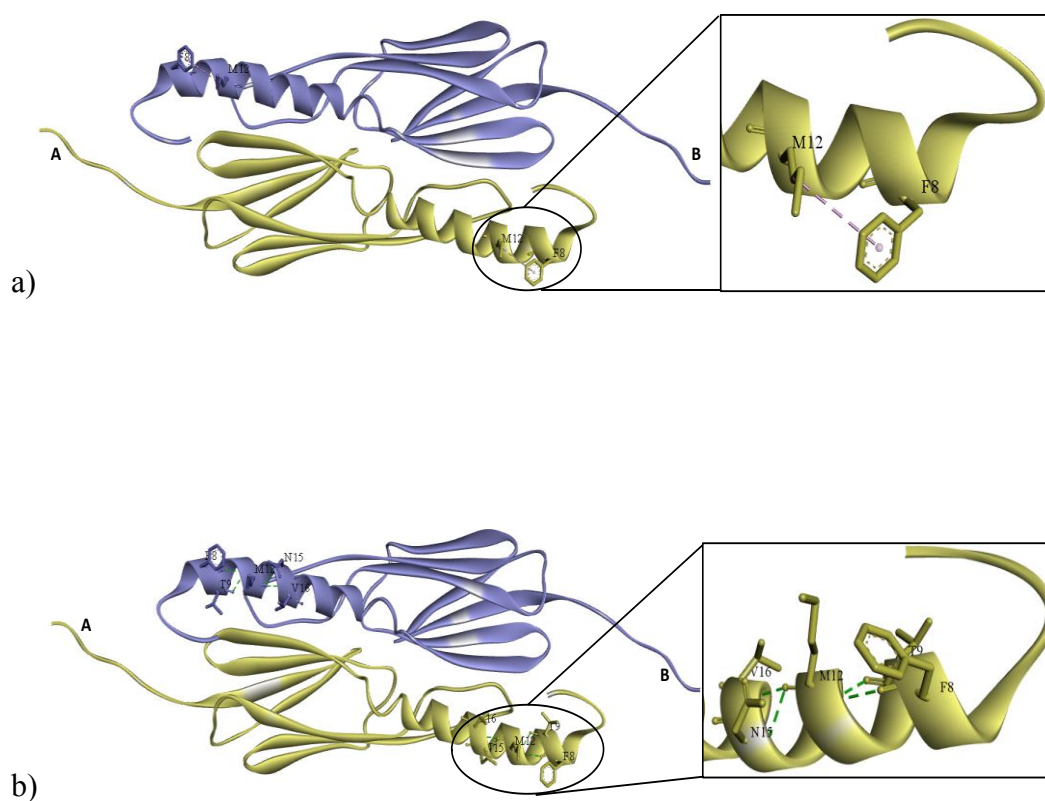


Figure 3.91 3-D images of dimer of *Tpv* sHSP 14.3 WT displaying the (a) hydrophobic and (b) hydrogen bonds formed by Met12 in each monomer within a dimer.

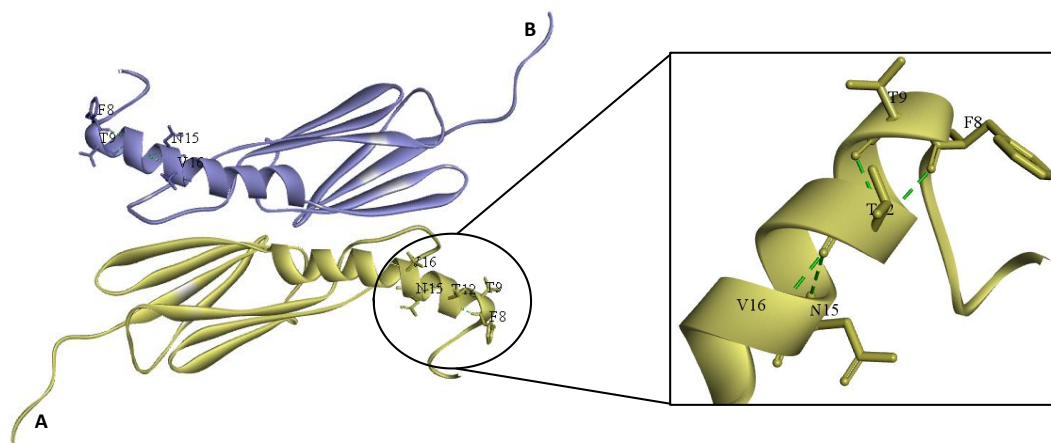


Figure 3.92 3-D image of dimer of *Tpv* sHSP 14.3 M12T variant showing hydrogen bonds formed by Thr12 in each monomer within a dimer.

3.13.2.3 Distal NTD Mutants

3.13.2.3.1 F26YI27T variant

Considering interactions of the substituted amino acids and their corresponding residues in WT, the substitution of two side by side/parallel amino acids such that, the hydrophobicity at this position is decreased considerably, resulted in almost doubling the number of intramolecular hydrogen bonds formed by amino acid at position 27 (Threonine) in the mutant, (Appendix E, Table 10). Also loss of the intermolecular hydrophobic interaction formed between isoleucine at position 27 of both the monomers (in WT, Figure 3.93) which may result in loosely bound dimer at this position. However, the Tyrosine at position 26 in the mutant forms intramolecular hydrophobic interactions with valine at position 23, whereas these interactions are not seen in the WT dimer. Moreover, as seen in 3-D image (Figure 3.94), this double mutation leads to the structural change from coil (in the WT) to helix (in the variant).

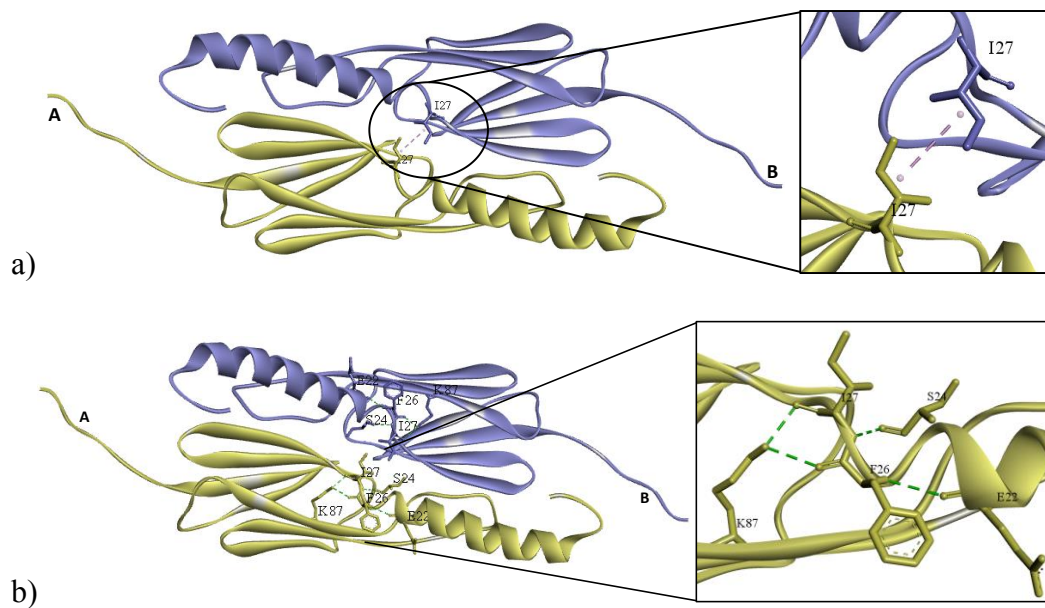


Figure 3.93 3-D images of dimer of *Tpv* sHSP 14.3 WT displaying the (a) intermolecular hydrophobic and (b) intramolecular hydrogen bonds formed by Phe26 and Ile27 within a dimer.

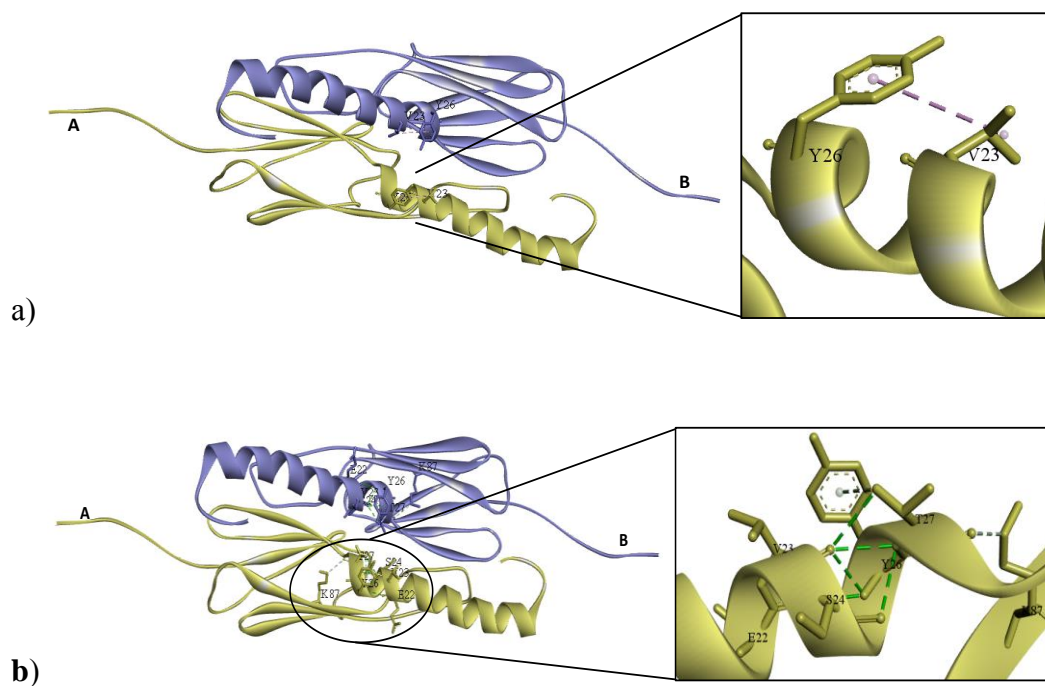


Figure 3.94 3-D images of dimer of *Tpv* sHSP 14.3 F26YI27T displaying the (a) intramolecular hydrophobic and (b) intramolecular hydrogen bonds formed by Tyr26 and Thr27 within a dimer.

3.13.2.3.2 V23GF26YI27T variant

Along with position 26 and 27, if position 23 is also made less hydrophobic, by replacing Valine with Glycine, the intramolecular hydrogen bonds formed by Phe26 and Ile27 with Lys87 in the WT (Figure 3.95b), are lost in the variant (Figure 3.96b), (Appendix B, Table 11). However, the mutated residues at position 26 and 27 now formed new bonds with Gly23 (which are not found in the WT). Overall, the WT and mutant possessed similar number of intramolecular hydrogen bonds. However, in contrast to the double mutant (F26YI27T), intermolecular hydrophobic interaction (Figure 3.95a) was replaced by intermolecular hydrogen bond at same position, *i.e.*, 27, in the mutant (Figure 3.96a). Moreover, dissimilar to F26YI27T, this triple mutation increases the disordered length of the proximal NTD which also moves away from the center of the dimer (Figure 3.96).

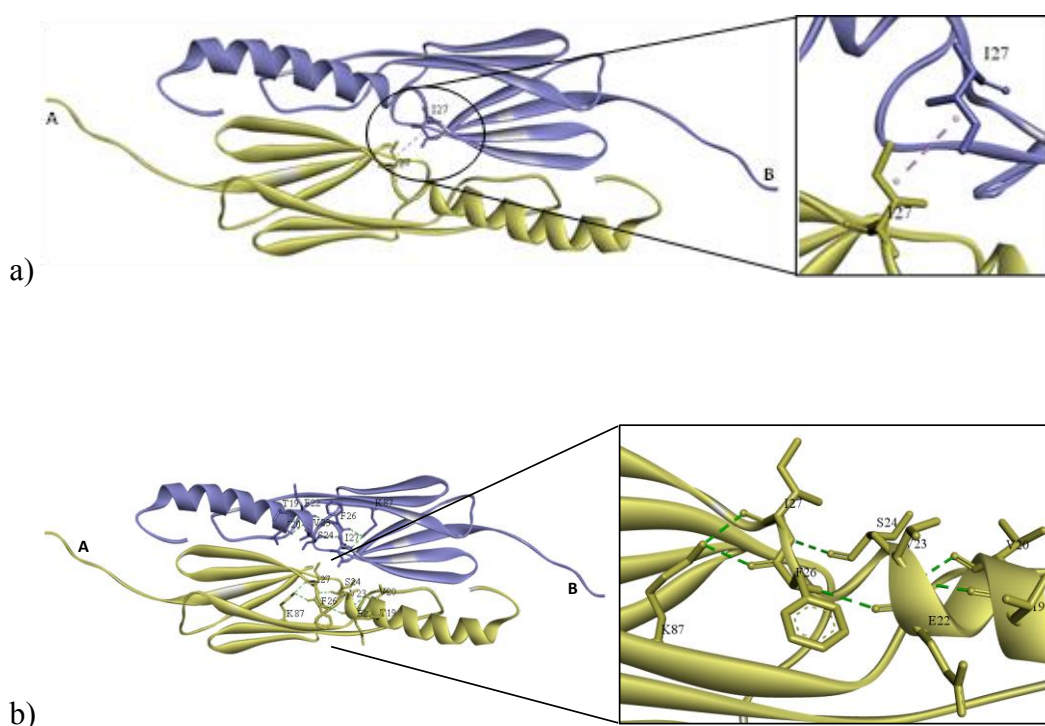


Figure 3.95 3-D images of dimer of *Tpn* sHSP 14.3 WT displaying the (a) intermolecular hydrophobic bond formed by Ile27 and (b) intramolecular hydrogen bonds formed by Val23, Phe26 and Ile27 within a dimer.

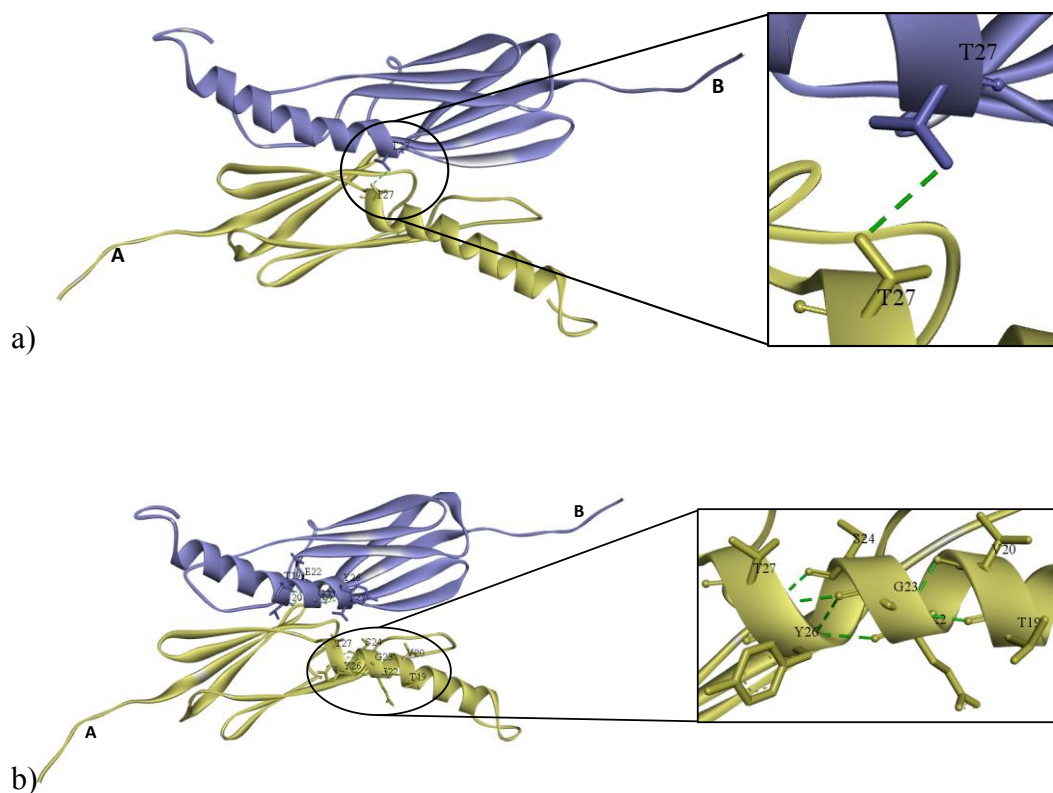


Figure 3.96 3-D images of dimer of *Tpv* sHSP 14.3 V23GF26YI27T displaying the (a) intermolecular hydrogen bond formed by Thr27 and (b) intramolecular hydrogen bonds formed by Gly23, Tyr26 and Thr27 within a dimer.

3.13.2.3.3 V31G variant

Analysis of the 3-D structure of WT and V31G mutant variant show that, decreasing the hydrophobicity at position 31, by replacing valine with glycine, results in loss of intramolecular hydrophobic interactions of residue 31 with Ala44 in the variant (Figure 3.97 and 3.98).

However, the intermolecular hydrogen bonds remained more or less, same (Appendix E, Table 12). Moreover, this single mutation alone has impacts on the proximal NTD by increasing the length of the disordered coil and transforming some part of the coil present at the junction of NTD and ACD (in WT) into structured helix (Figure 3.98).

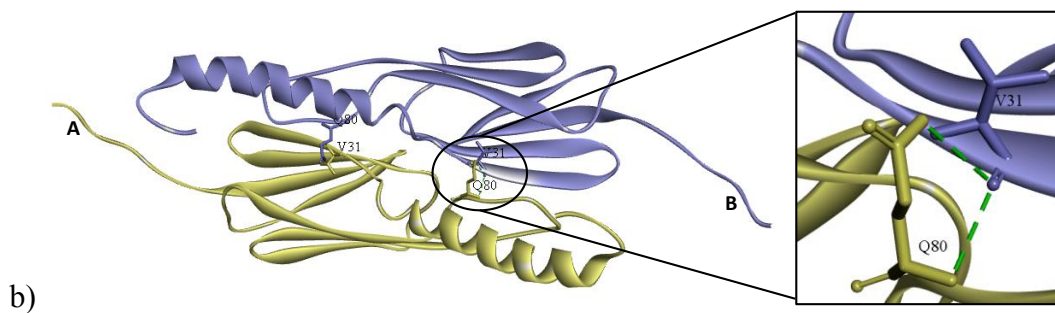
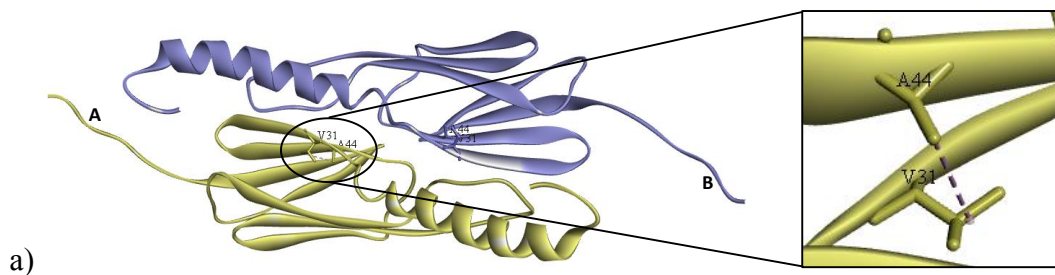


Figure 3.97 3-D images of dimer of *Tpv* sHSP 14.3 WT displaying the (a) intramolecular hydrophobic bond and (b) intermolecular hydrogen bonds formed by Val31 within a dimer.

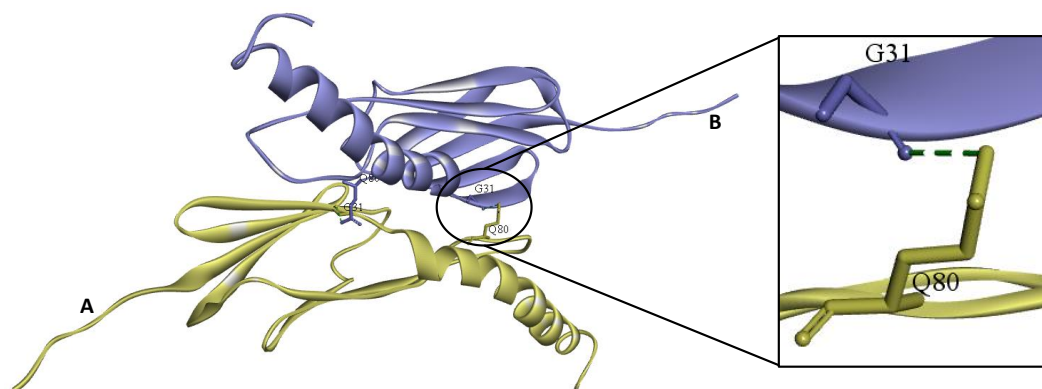


Figure 3.98 3-D image of dimer of *Tpv* sHSP 14.3 V31G displaying the intermolecular hydrogen bonds formed by Gly31 within a dimer.

3.13.2.3.4 V31IL33I variant

Increase in the hydrophobicity at position 31 is attempted along with position 33 (V31IL33I), resulted in similar number of intermolecular hydrophobic interactions.

Although the hydrophobic interaction of position 31 with Pro92 was lost, a new interaction was formed between position 31 and Leu42.

Also, new intramolecular hydrogen bonds were formed in the double mutant V31I/L33I (Figure 3.100b). The results along with distances are shown in Appendix E, Table 13. The intermolecular hydrogen bonds within the dimer increased in the mutant (Figure 3.100c) as compared to WT (Figure 3.99b).

In contrast to V31G mutant, V31I/L33I did not have an impact on proximal NTD. Instead, orientation of the CTD has changed by switching outwards.

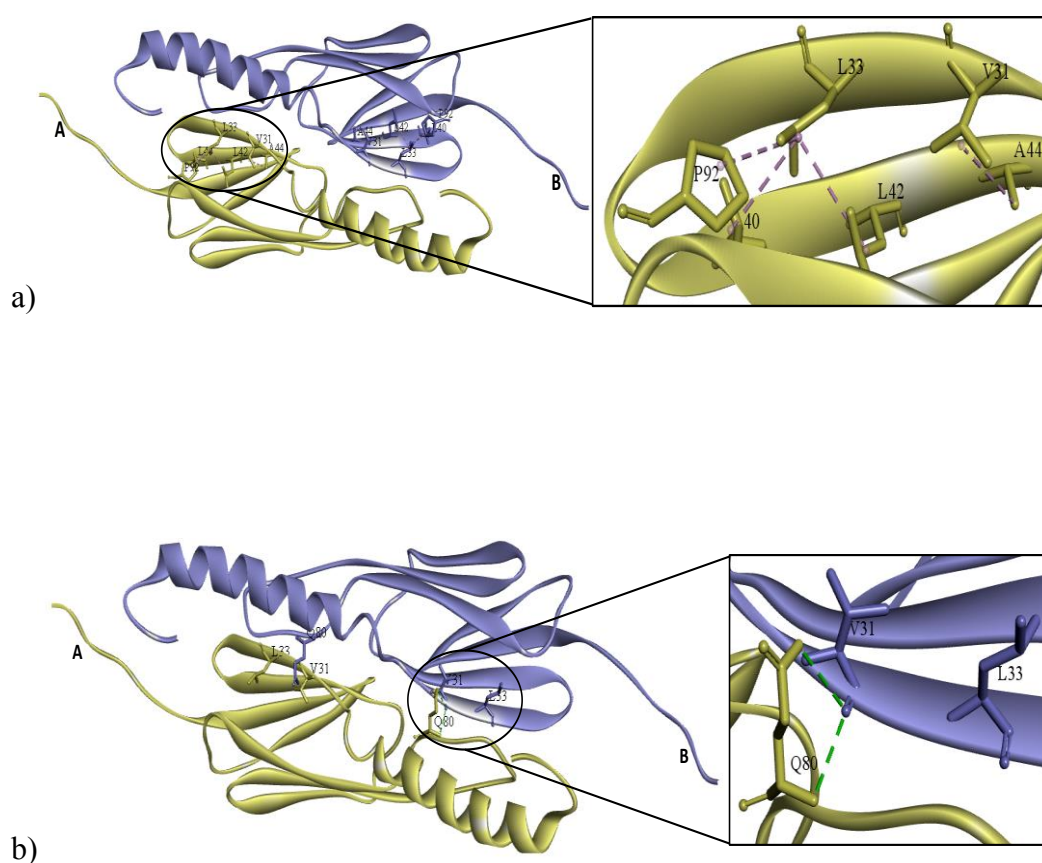


Figure 3.99 3-D images of dimer of *Tpv* sHSP 14.3 WT displaying the (a) intramolecular hydrophobic bonds and (b) intermolecular hydrogen bonds formed by Val31 and Leu33 within a dimer.

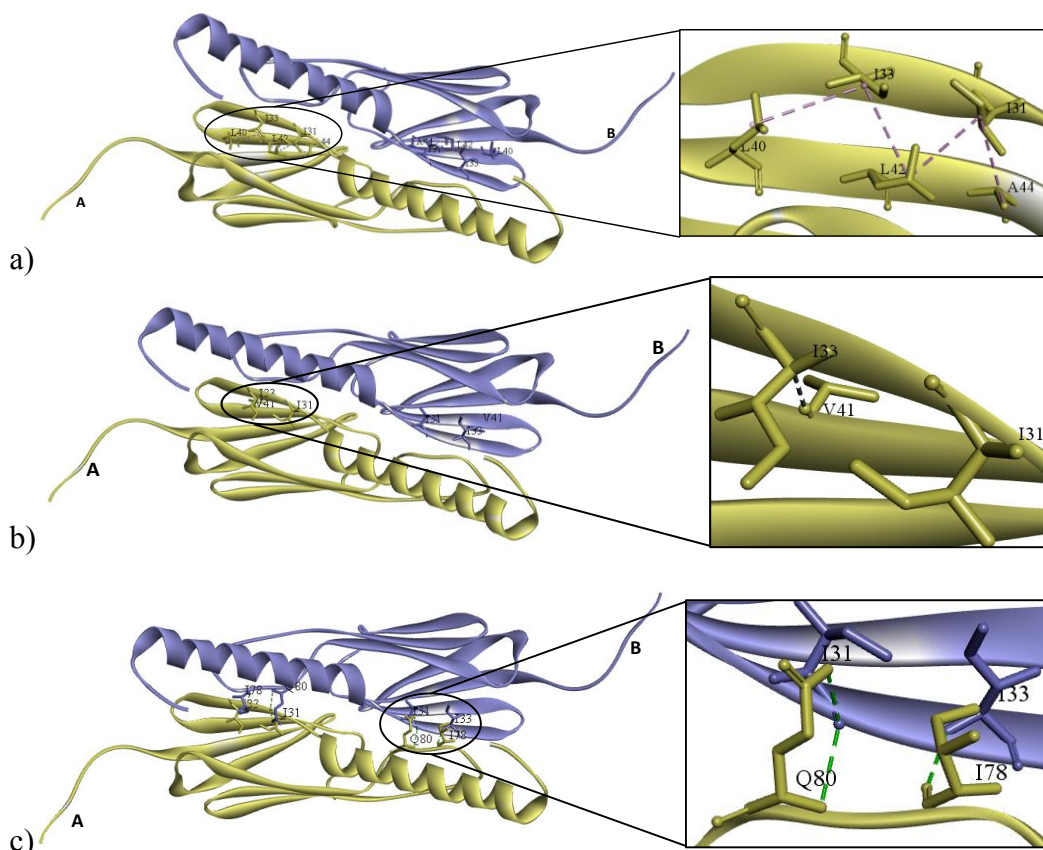


Figure 3.100 3-D images of dimer of *Tpv* sHSP 14.3 V31IL33I displaying the (a) intramolecular hydrophobic bonds and (b) intramolecular hydrogen bonds and (c) intermolecular hydrogen bonds formed by Ile31 and Ile33 within a dimer.

3.13.2.3.5 F8YV31G variant

The double mutant F8YV31G lost its intramolecular hydrophobic interactions with alanine at position 44. This mutant possessed the intermolecular hydrogen bonds with the same residue as in WT, *i.e.*, glutamine 80 (although the atoms involved are different), (Appendix E, Table 14). On the other hand, the single mutation F8Y resulted in formation of intramolecular hydrophobic and electrostatic interactions with new residues (Figure 3.80). Introduction of another distal mutation V31G along with F8Y, results in loss of hydrophobic interaction at position 8 (in the mutant) while the electrostatic interactions are retained, similar to WT, but with increase in the length of the interaction (Figure 3.102). Also, the intramolecular hydrogen bonds, at position 8, in the double mutant F8YV31G are exactly similar to WT.

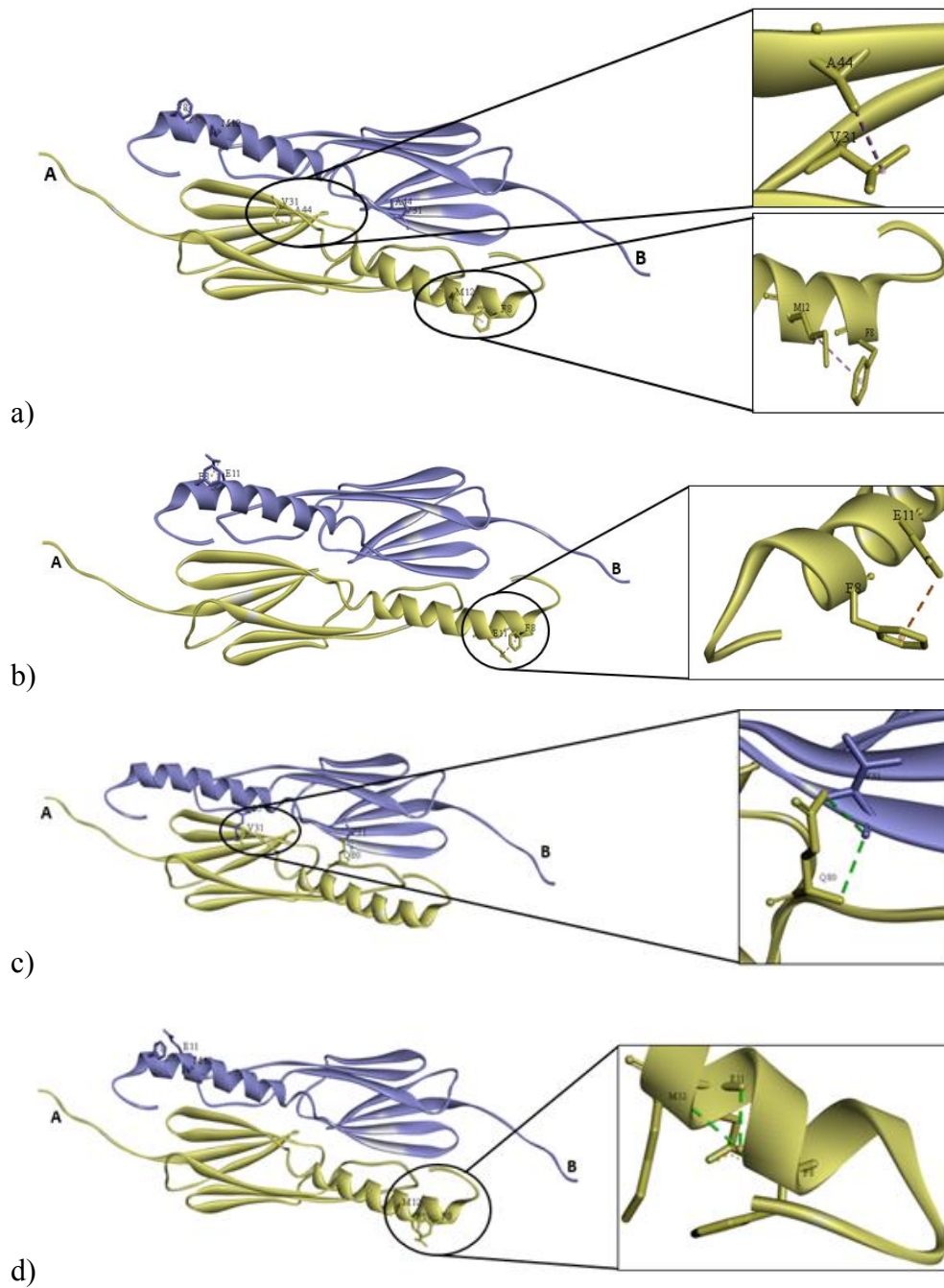


Figure 3.101 3-D images of dimer of *Tpv* sHSP 14.3 WT displaying the (a) intramolecular hydrophobic bonds, (b) intramolecular electrostatic, (c) intermolecular hydrogen and (d) intramolecular hydrogen bonds formed by Phe8 and Val31 within a dimer.

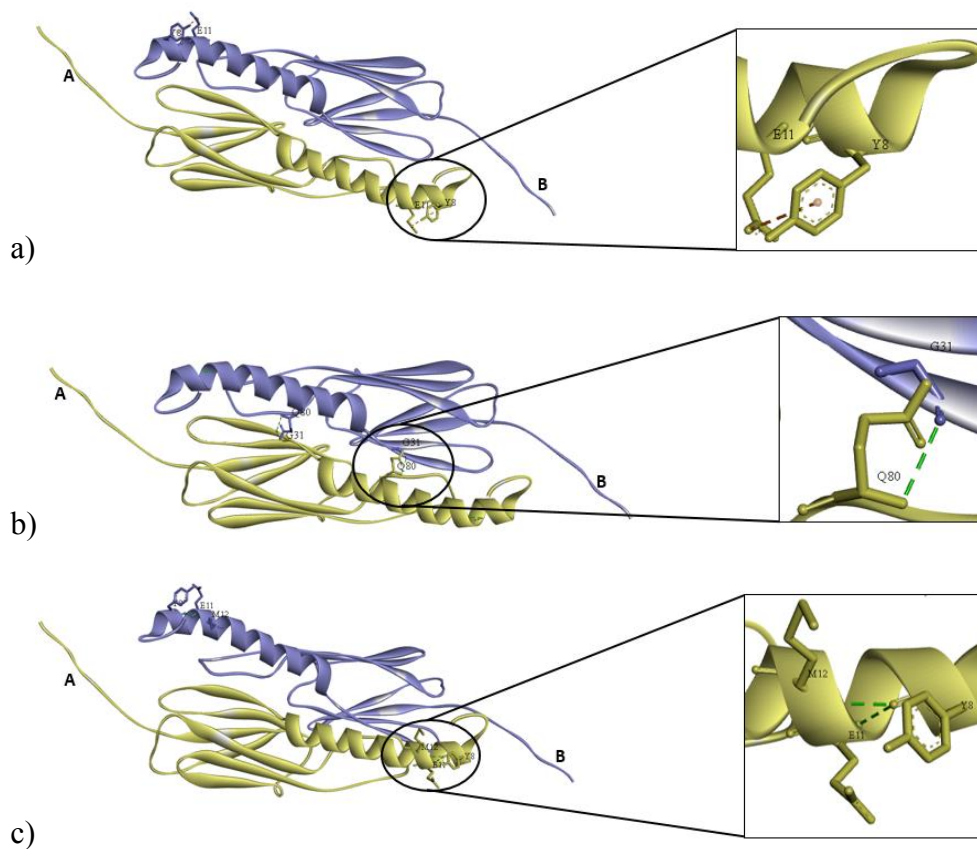


Figure 3.102 3-D images of dimer of *Tpv* sHSP 14.3 F8YV31G displaying the (a) intramolecular electrostatic, (b) intermolecular hydrogen and (c) intramolecular hydrogen bonds formed by Tyr8 and Gly31 within a dimer.

3.13.3 ZipperDB and Rosetta Energy

The Rosetta Energy calculation and analysis was done by using the database web service available at <https://services.mbi.ucla.edu/>. It identifies the fibril forming segments within a protein by 3-D profile method. The Rosetta energy is in Kcal/mol and it is the energy of two beta-strands with 6 residues. The cut-off value for the Rosetta Energy is at -23. The energy ≥ -23 indicates that the particular protein segment possess the tendency to form fibril structure. If an energy value crosses the threshold of -23 Kcal/mol, the probability of that region in forming β -amyloid fibrils increases. In addition to Rosetta Energy, this program also calculates the shape complementarity, area of interface and C-score for each of six residue fibril in the whole protein.

Shape complementarity means the steric zipper interface and values vary between 0.0 and 1.0, where higher values increase the probability of fibril formation. Area of interface dictates the solvent accessible area, and increase in this decreases the fibril formation chances.

Combining the effect of Rosetta energy, area of interface and shape complementarity, a score is given to each peptide, known as composite score (C-score). The lower the composite score, the lower the propensity of the peptide to form fibril (Kuhlman and Baker, 2000; Nelson *et al.*, 2005; Thompson *et al.*, 2006; Sawaya *et al.*, 2007).

The graphical representation of Rosetta Energy of the whole protein *Tpv* sHSP 14.3 is given in Figure 3.103. The y-axis displays the energy value for each of the six amino acids peptide. The colors of the peptide vary from red to blue and depend upon the energy value of the peptide, from lower to higher energy. When viewing the overall picture of the WT protein *Tpv* sHSP 14.3, the ACD, particularly the part of ACD lying close to the NTD is more inclined to forming β -amyloid fibrils. This is probably balanced by the rest of the protein which shows less incline towards forming fibrils or aggregates, (Figure 3.103).

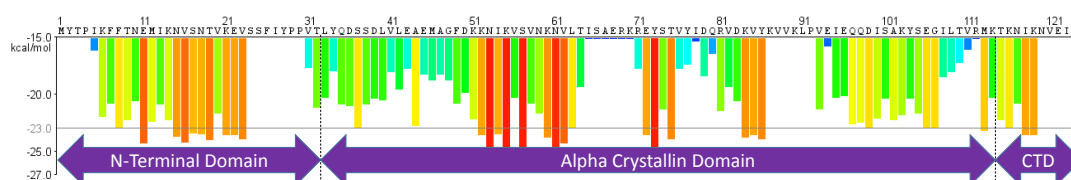


Figure 3.103 Graphical representation of Rosetta energy of *Tpv* sHSP 14.3 WT protein. The y-axis shows the energy of each six amino acid protein, in terms of Rosetta energy. The bars are colored from blue to red range according to energy, where red represents the highest energy and blue the least energy.

All the numerical output of ZipperDB is given in Table 3.12 – 3.14. The ZipperDB results show that mutations in the proximal part of NTD of *Tpv* sHSP 14.3, those which intend to decrease the hydrophobicity does not cause a significant alteration in the Rosetta Energy of the protein molecule. Reduction in Rosetta Energy of any

sequential 2 – 6 peptide, below -23 Kcal/mol causes an increase in propensity to form fibrils by the protein in that particular region.

Considering the NTD alone, of *Tpv* sHSP 14.3 WT, (Figure 3.104a), the initial region and the region lying close to ACD shows low Rosetta Energy while that lying in the middle have high energy which at positions 11, 15 – 19, 21 and 23 have passed the threshold of -23 Kcal/mol. Still the Rosetta Energy in this section is higher than that of ACD proximal region, but lower than the ACD 51 – 62 residues, (Figure 3.103). First, if we consider all the mutations planned that lie in the proximal section, *i.e.*, initial 10 residues of the NTD, (Figure 3.104b – 3.104f), show that two variants F7SF8Y and I5TF7SF8Y show the highest increase in Rosetta Energy, in at least two hexapeptides. This may indicate that hydrophobicity at this position, probably F7 residue, is important for maintaining the structure of the protein molecule as a soluble entity and preventing it from forming β -amyloid fibrils. This is also indicated in the Table 3.12, where the increase in energy of the 6 residue peptides in both the mutants (as compared to WT) is highlighted in yellow. It indicates the increasing of the overall score, C-score, which is also an indication that the ability to maintain the protein solubility by this region of the peptide is decreased. The C-score in both the mutants, F7SF8Y and I5TF7SF8Y, also show an increasing trend, particularly in the former, where the peptide ${}^5\text{IKSYTN}_{10}$ show an increase in C-score by a value of 10. Another peptide ${}^7\text{SYTNEM}_{12}$ show an increase in C-score by a value of 4.5. Both the mutants showed decrease in ability to protect CS from aggregation and its activity at sub-optimal temperature which might be due to this increase in fibril formation propensity of this region of the protein. Besides, decrease in hydrophobicity at position 7 and 8 could severely interfere with substrate interaction. On the other hand, in other variants I5T (Figure 3.104b), F8Y (Figure 3.104c) and their double mutant I5TF8Y (Figure 3.104d), very slight increase in Rosetta Energy is observed around the mutating residue, and it does not increase beyond the threshold of -23 Kcal/mol. Increase in the C-score for these mutant variants also was very low. This is also supported by the fact that high amount of expressed protein is found in the soluble part of cell lysate, stability at temperatures as high as 80°C, as well as the

protection effect against heat induced aggregation of CS (and its activity) of these variants (I5T, F8Y and I5TF8Y) was comparable to WT.

Considering the three single mutants, where the mutation point is present in the central region of the NTD of *Tpv* sHSP 14.3, *i.e.*, E11V (3.104g), E22G (3.104h) and M12T (3.105b), all the three mutations resulted in drastic increase in the Rosetta Energy and C-Score of the region around the mutation point. The values of Rosetta Energy and C-Score, before and after the mutation are listed in Table 3.13 for these mutants, where those values that pass the threshold (for Rosetta Energy) of forming fibrils are marked in yellow. For the mutant M12T, the last three segments (hexapeptide) that include the mutation point, showed drastic increase in C-score, where the second last segment recorded highest increase of 9.501. Also, in the mutant, out of six hexapeptide segments including the mutation point, three passed the threshold value of Rosetta Energy (-23 Kcal/mol), whereas in WT, only one out of six of these segments passed the -23 Kcal/mol value.

Similar, rather drastic, results are seen in the variant E11V. Prior to mutation, one out of six hexapeptide possessed Rosetta Energy values passing the threshold whereas mutation resulted in five out of the six hexapeptide to have energy values passing the threshold. In case of C-score, besides a single segment, all other peptide segments, in the variant, showed increase in C-score, with the highest increase (4.318) being in the same peptide that showed highest increase in Rosetta Energy (2.7 Kcal/mol). In the variant E22G, the last segment/hexapeptide that includes the mutation point of 22 showed severe change in both Rosetta Energy (increase by 3.6 Kcal/mol) and C-score (5.6). It was observed that in spite of exhibiting high transformation efficiency, (Table 3.8 and 3.10), and high plasmid concentration in the cells, (Figure 3.32 and 3.33), the protein amount was very low in the cell free extract, (Figure 3.38, 3.39 and 3.40). This might be explained in such that the protein expressed had increase tendency to form fibrils (evident by Rosetta Energy and C-score values) and, therefore, formed aggregates, which got removed during the centrifugation steps.

The mutants, F26YI27T and V23GF26YI27T, where the hydrophobicity gets reduced at adjacent residues near the carboxyl end of the NTD, there has been a slight increase in Rosetta Energy (Figure 3.105e and 3.105f) only, in the last hexapeptide segment, from 23.9 in the WT to 24.8 and 24.9 in the double and the triple mutant, respectively, otherwise the Rosetta Energy has decreased in the other segments, (Table 3.14). The C-score, which counts the Rosetta Energy, Shape complementarity and Area of Interface as a whole, except increase (by a value of 5.99) in C-score of the last hexapeptide segment of the double mutant, all the hexapeptide segments for the triple mutant show an decrease in their C-score values. Therefore, it shows that, decreasing the hydrophobicity at distal end of NTD of sHSP results in decrease in the propensity of fibril formation and could be the reason of better protective effect of these mutants against heat induced aggregation of CS.

For the three variants, V31G, F8YV31G and V31IL33I, all hexapeptide segments (except one for F8YV31G), showed increase in the C-score (varying from 0.1 to 8), (Table 3.14), along with a humble increase in Rosetta Energy. However, no value of Rosetta Energy of any peptide was so low as to pass the threshold of -23, (Figure 3.105c, 3.105d and 3.105g).

Since no drastic change in Rosetta Energy was observed for these mutants, it was expected that the mutant protein yield would not be severely affected due to aggregation, (Figure 3.37, 3.43 and 3.44). However, all three variants could not withstand the high temperature of 70°C, as opposed to WT. This decrease in high temperature withstanding stability, as compared to WT, could be attributed to this moderate increase in Rosetta Energy as well as the increase in C-score, by a value of 5 or more, in at least one hexapeptide segment of each of the mutants, in addition to other factors (gain/loss of inter and intra molecular interactions).

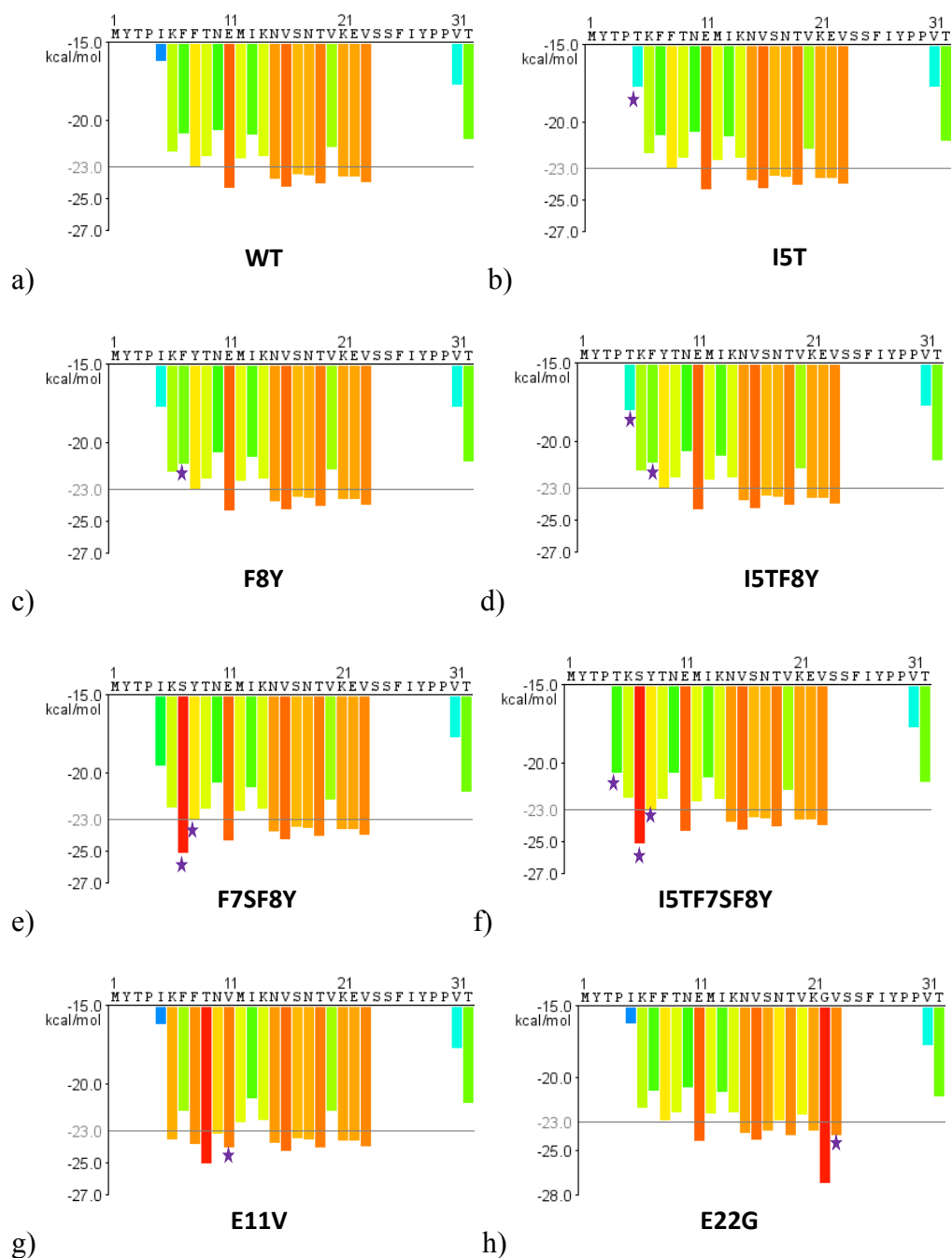


Figure 3.104 Rosetta Energy graph of each of *Tpv* HSP 14.3 WT and mutants as analyzed by ZipperDB for (a) WT, (b) I5T, (c) F8Y, (d) I5TF8Y, (e) F7SF8Y, (f) I5TF7SF8Y, (g) E11V and (h) E22G. Each histogram bar represent one hexapeptide starting at the indicated position in the sequence and is colored according to its Rosetta energy. Orange-red segments are those with higher Rosetta energies.

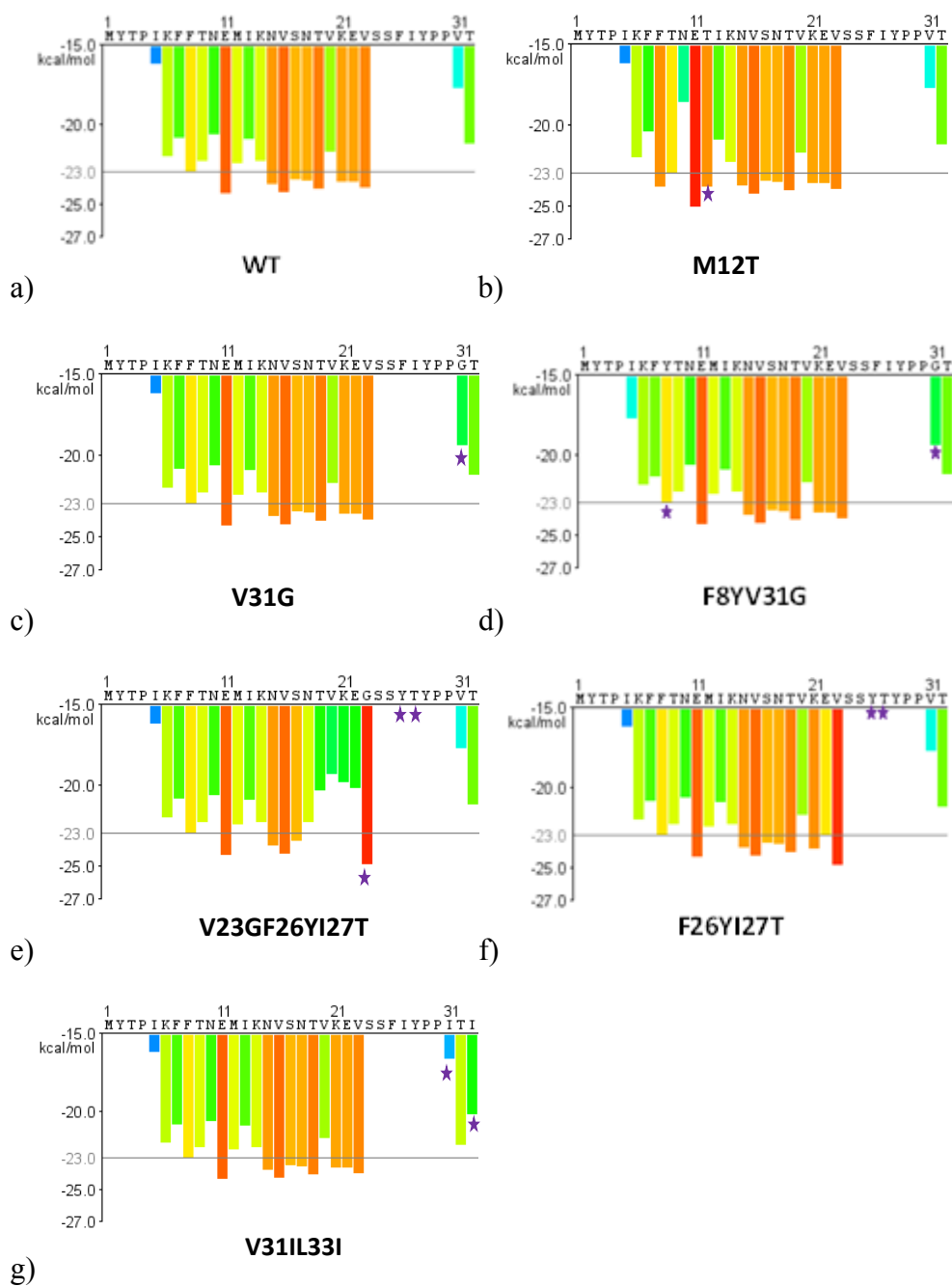


Figure 3.105 Rosetta Energy graph of each of *Tpv* HSP 14.3 WT and mutants as analyzed by ZipperDB for (a) WT, (b) M12T, (c) V31G, (d) F8YV31G, (e) V23GF26YI27T, (f) F26YI27T and (g) V31IL33I. Each histogram bar represent one hexapeptide starting at the indicated position in the sequence and is colored according to its Rosetta energy. Orange-red segments are those with higher Rosetta energies.

Table 3.12 Comparison of Rosetta Energy, Shape Complementarity and C-Score of *Tp ν* sHSP 14.3 WT and Mutants involving initial NTD residues

Mutant Name	WILD TYPE				MUTANT			
	Sequence	Rosetta Energy	Shape Complementarity	C-Score	Sequence	Rosetta Energy	Shape Complementarity	C-Score
I5T	IKFFTN	-16.2	0.8467	-28.901	TKFFTN	-17.7	0.8694	-30.741
F8Y	IKFFTN	-16.2	0.8467	-28.901	IKFYTN	-17.7	0.8467	-30.401
	KFFTNE	-22.0	0.8108	-34.162	KFYTNE	-21.8	0.8903	-35.154
	FFTNEM	-20.8	0.7112	-31.468	FYTNEM	-21.3	0.7112	-31.968
	FTNEMI	-22.9	0.7649	-34.374	YTNEMI	-22.9	0.7567	-34.251
I5TF8Y	IKFFTN	-16.2	0.8467	-28.901	TKFYTN	-18	0.8694	-31.041
	KFFTNE	-22	0.8108	-34.162	KFYTNE	-21.8	0.8903	-35.154
	FFTNEM	-20.8	0.7112	-31.468	FYTNEM	-21.3	0.7112	-31.968
	FTNEMI	-22.9	0.7649	-34.374	YTNEMI	-22.9	0.7567	-34.251
I5TF7SF8Y	IKFFTN	-16.2	0.8467	-28.901	TKSYTN	-20.6	0.7169	-31.354
	KFFTNE	-22	0.8108	-34.162	KSYTNE	-22.2	0.8903	-35.554
	FFTNEM	-20.8	0.7112	-31.468	SYTNEM	-25.1	0.7268	-36.002
	FTNEMI	-22.9	0.7649	-34.374	YTNEMI	-22.9	0.7567	-34.251
F7SF8Y	IKFFTN	-16.2	0.8467	-28.901	IKSYTN	-19.5	0.77	-39.3
	KFFTNE	-22	0.8108	-34.162	KSYTNE	-22.2	0.8903	-35.554
	FFTNEM	-20.8	0.7112	-31.468	SYTNEM	-25.1	0.7268	-36.002
	FTNEMI	-22.9	0.7649	-34.374	YTNEMI	-22.9	0.7567	-34.251

Table 3.13 Comparison of Rosetta Energy, Shape Complementarity and C-Score of *Tpv* sHSP 14.3 WT and Mutants involving central NTD residues

Mutant Name	WILD TYPE					MUTANT				
	Sequence	Rosetta Energy	Shape Complementarity	C-Score	Sequence	Rosetta Energy	Shape Complementarity	C-Score		
M12T	FFT NEM	-20.8	0.7112	-31.468	FFT NET	-20.4	0.7112	-31.068		
	FT NEM I	-22.9	0.7649	-34.374	FT NET I	-23.8	0.835	-36.325		
	T NEM IK	-22.3	0.8044	-34.366	T NET IK	-22.9	0.8044	-34.966		
	N EM IKM	-20.6	0.762	-32.03	N ET IKN	-18.6	0.8419	-39.078		
	E MI KNV	-24.3	0.8673	-37.309	E TI KNV	-25	0.854	-46.81		
	MI KNVS	-22.4	0.8745	-35.517	TI KNVS	-23.8	0.7549	-35.124		
E11V	KFF TNE	-22	0.8108	-34.162	KFF TNV	-23.5	0.8108	-35.662		
	FFT NEM	-20.8	0.7112	-31.468	FFT NVM	-21.7	0.7857	-33.486		
	FT NEM I	-22.9	0.7649	-34.374	FT NVM I	-23.8	0.843	-36.445		
	T NEM IK	-22.3	0.8044	-34.366	T NVM IK	-25	0.9123	-38.684		
	N EM IKN	-20.6	0.762	-32.03	N VM IKN	-23.1	0.762	-34.53		
	EM IKNV	-24.3	0.8673	-37.309	VM IKNV	-24	0.7862	-35.793		
E22G	SNT VKE	-23.4	0.7761	-35.041	SNT VKG	-23.6	0.7761	-35.242		
	NT VKE V	-23.5	0.6689	-33.533	NT VKG V	-22.9	0.9676	-37.414		
	TV KE VS	-24	0.7549	-35.324	TV KG VS	-23.9	0.7549	-35.224		
	V KE VSS	-21.7	0.7251	-32.577	V KG VSS	-22.5	0.7316	-33.474		
	K EV SSF	-23.6	0.748	-42.87	K GV SSF	-23.6	0.7292	-34.538		
	EV SSFI	-23.6	0.867	-45.305	GV SSFI	-27.2	0.8368	-39.752		

Table 3.14 Comparison of Rosetta Energy, Shape Complementarity and C-Score of *Tpv* sHSP 14.3 WT and Mutants involving NTD residues towards the carboxyl end central NTD residues

Mutant Name	WILD TYPE				MUTANT			
	Sequence	Rosetta Energy	Shape Complementarity	C-Score	Sequence	Rosetta Energy	Shape Complementarity	C-Score
V31G	VTLYQD	-17.7	0.8605	-30.608	GTLYQD	-19.4	0.7219	-35.578
V31H33I	VTLYQD	-17.7	0.8605	-30.608	ITLYQD	-16.6	0.925	-38.575
	TLYQDS	-21.2	0.7483	-32.425	TLYQDS	-22.1	0.7483	-33.325
	LYQDSS	-20.3	0.8456	-32.984	IYQDSS	-20.2	0.8647	-33.17
V23GF26YI27T	NTVKEV	-23.5	0.6689	-33.533	NTVKEG	-22.3	0.6689	-32.334
	TVKEVS	-24	0.7549	-35.324	TVKEG	-20.3	0.8467	-33.001
	VKEVSS	-21.7	0.7251	-32.577	VKEGSS	-19.3	0.7251	-30.177
	KEVSSF	-23.6	0.748	-42.87	KEGSSY	-19.8	0.8724	-32.886
	EVSSFI	-23.6	0.867	-45.305	EGSSYT	-20.2	0.8627	-33.141
	VSSFIY	-23.9	0.885	-37.175	GSSYTY	-24.9	0.7869	-36.703
F26YI27T	KEVSSF	-23.6	0.748	-42.87	KEVSSY	-23.8	0.747	-43.055
	EVSSFI	-23.6	0.867	-45.305	EVSSYT	-22.9	0.7829	-43.343
	VSSFIY	-23.9	0.885	-37.175	VSSYTY	-24.8	0.7241	-43.161
F8YV31G	IKFFT	-16.2	0.8467	-28.901	IKFYTN	-17.7	0.8467	-30.401
	KFFTNE	-22.0	0.8108	-34.162	KFYTNE	-21.8	0.8903	-35.154
	FFTNE	-20.8	0.7112	-31.468	FYTNE	-21.3	0.7112	-31.968
	FTNEMI	-22.9	0.7649	-34.374	YTNEMI	-22.9	0.7567	-34.251
	VTLYQD	-17.7	0.8605	-30.608	GTLYQD	-19.4	0.7219	-35.578

3.13.4 Intrinsic Disorder Region (IDR) Analysis

A study of intrinsic disordered regions in *Tpv* sHSP 14.3, especially its NTD region, was done since some regions of the sHSPs at low temperatures are considered to be disordered and lack a definite tertiary structure. There are four residues of each Proline and Phenylalanine present in the primary sequence of *Tpv* sHSP 14.3 WT, and three of each are present in the NTD alone. High representation of Proline and aromatic amino acids in the NTD is in consensus with it being unstructured and involved in substrate binding (Montfort, Slingsby and Elizabeth, 2002; Kriehuber *et al.*, 2010), and therefore, was expected to be intrinsically disordered like NTD of other sHSPs (Wu *et al.*, 2019), along with the CTD of the protein (V. Sudnitsyna *et al.*, 2012). The IDR of *Tpv* sHSP 14.3 protein was studied using two online servers, independently, for comparison and confirmation. The first tool employed was PONDR (Predictor of Naturally Disordered Regions), while the second one was PrDOS (Protein Disorder Prediction Server).

3.13.4.1 Analysis using PONDR program

In this approach the fractional composition of an amino acid, its hydrophathy as well as other attributes are considered, averaged and then used as input to make predictions. The output, also known as PONDR score (along the y-axis) in the graph, is between 0 and 1. The threshold set for this tool is 0.5 and any residue value that exceeds this threshold is considered to lie in the disordered region (Romero *et al.*, 2001). The output for the whole protein of *Tpv* sHSP 14.3 WT is shown in Figure 3.106. Of the many algorithms available at PONDR, the algorithm with success rate of predicting order and disorder of 78% and 60%, respectively, was used. According to this online tool, our protein of interest contains three disordered regions, one in the NTD (residue 1) and two in the ACD of the protein (residues 61 – 68 and residue 73 – 78).

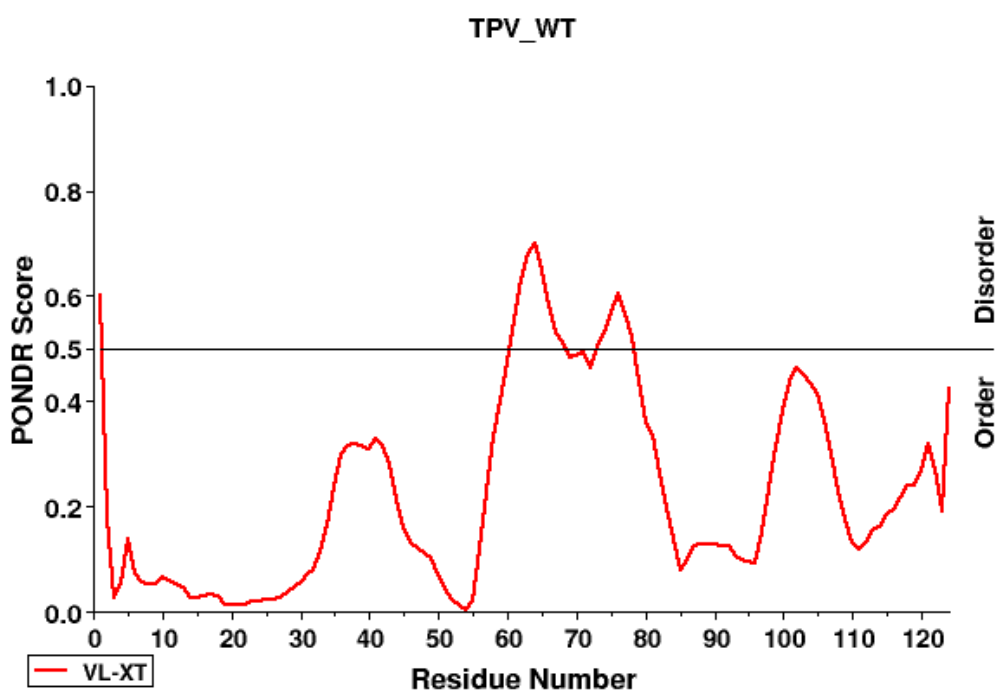


Figure 3.106 PONDR output of Tpv HSP 14.3 protein. The x-axis represent each amino acid in the protein monomer. PONDR score of each amino acid is given on y-axis.

3.13.4.1.1 PONDR output for NTD proximal mutant variants

By mutagenesis, the hydrophobicity of the proximal region of NTD was decreased by a set of point mutations, namely I5T, F8Y, I5TF8YT, F7SF8Y and I5TF7SF8Y. Thus the mean hydrophobicity of this region was disturbed/decreased in the variants, whereas the mean net charge remained same. Therefore, it was naturally expected that the equilibrium will be shifted towards increase in IDR in the NTD, since IDRs depends on the balance between net charge and mean hydrophobicity (Van Der Lee *et al.*, 2014). The graphical representation of the IDR of NTD of *Tpv* sHSP 14.3 WT and NTD proximal mutant variants is given in Figure 3.107, while the numerical value of each residue, known as PONDR score, is listed in Table (Appendix E, Table 15). Introduction of polarity at position five by replacing a highly hydrophobic residue isoleucine with threonine, results in 3% decrease in hydrophobicity in this region, (Appendix E, Table 1) and thus resulted in slight increase in the disordered region of the protein Figure 3.107b. Upon analysis, it is seen that the PONDR score

of the initial twelve residues increases after this mutation, Table 15 (Appendix E), however, in contrast to WT, which has only one residue, M1, lying in the disordered region, in mutant I5T, the first two residues' score pass the threshold of 0.5, and are, therefore, declared to be disordered. On the other hand, replacing the highly hydrophobic phenylalanine at position eight with a comparatively less hydrophobic tyrosine results in decrease in overall hydrophobicity of the NTD, similar to I5T mutant, with regard to percentage (Appendix E, Table 1). However, in contrast to I5T, the PONDR score of residues decreased, from residue number three to number 22. This decrease in score is slight and, therefore, the overall image of the graph does not depict a severe change, and is very much similar to that of WT, as seen in Figure 3.107c. Combining both the single mutations I5T and F8Y in the same protein, results in further decrease in the hydrophobicity of this region of the protein by more than 4.5%. However, when analyzing the IDR of I5TF8Y, the result is somewhat similar to the single mutant I5T, Figure 3.107d. The first two residues showed large increase in PONDR score, and are thus declared to be disordered, while the residues from 3 – 14 exhibit a slight increase in score. This increase tapers off until the 14th position/residue. However, decreasing the hydrophobicity even further (6.25%), at adjacent positions by replacing two phenylalanine residues, at position 7 and 8, by a polar residue serine and less hydrophobic tyrosine, respectively, resulted in further increase in PONDR score for residues 2 – 21, yet, similar to I5T and I5TF8Y, only the first two amino acids crossed the threshold of 0.5 and thus were found to lie in disordered region. The triple mutant, I5TF7SF8Y, shows a very different image than WT and all the other mutants in this group, Figure 3.107f. This substantial decrease in hydrophobicity (around 10%) resulted in a sharp rise in the length of disordered region. Following this mutation, the initial seven residues crossed the threshold of 0.5 to be declared as disordered. Not only this, the residues from 8 to 14 showed a decent increase in PONDR score. The increase was moderate after residue 14th and diminished off until residue 22.

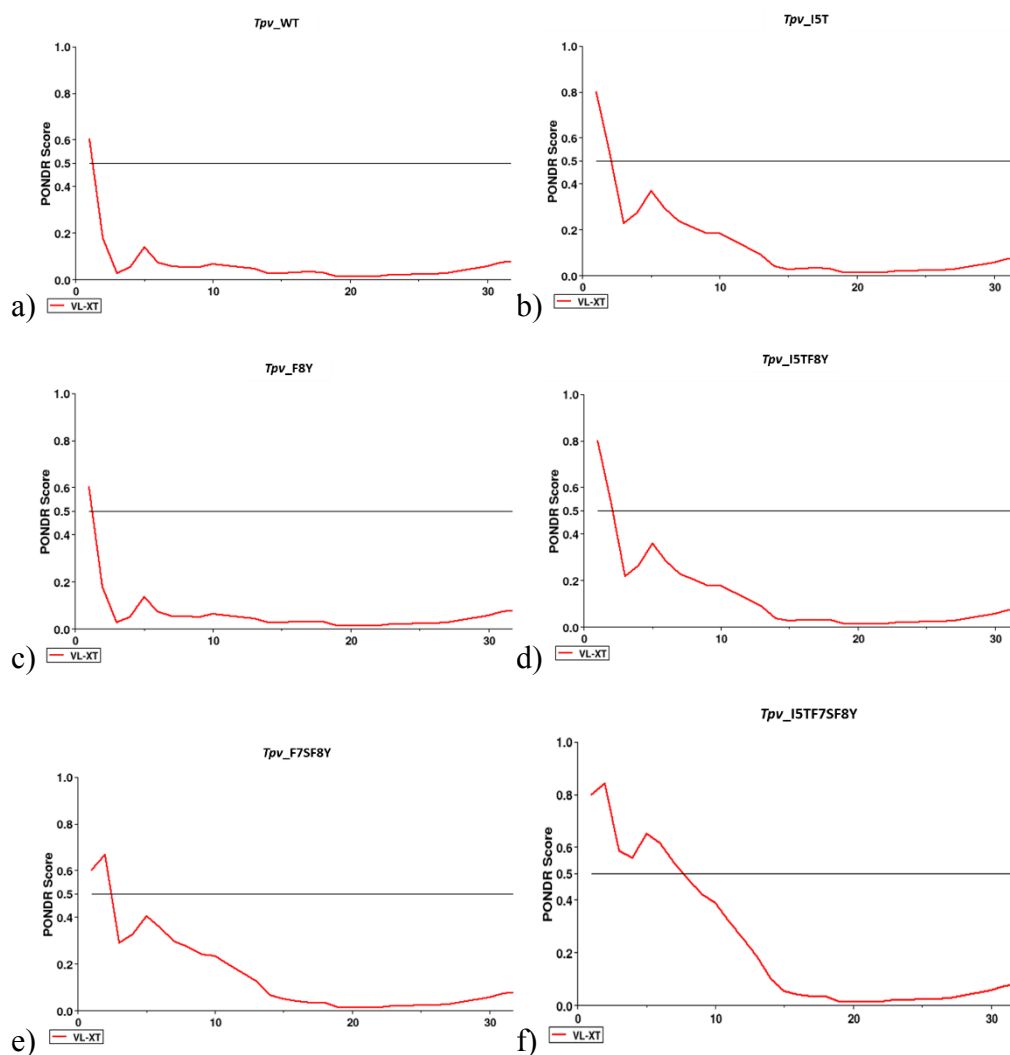


Figure 3.107 POND R score plots of NTD of *Tpv* HSP 14.3 WT (a), I5T (b), F8Y (c), I5TF8Y (d), F7SF8Y (e), and I5TF7SF8Y (f). The x-axis represents the residue number while the y-axis shows the POND R score of each residue. The grey line represents the threshold, above which is the disordered region and below, the ordered region.

3.13.4.1.2 POND R output for middle NTD mutant variants

In this group, two mutants include substituting the highly conserved negatively charge residue, E11 and E22, which results in decrease in the overall charge of the NTD (by 3.12%). Consequently there is an increase in the hydrophobicity of the NTD. Both the variants E11V and E22G show an increase in hydrophobicity around 3%, while the variant M12T exhibits a decrease in hydrophobicity by 3.12 %

(Appendix E, Table 1). The PONDR results (Figure 3.108 and appendix E, Table 16) indicate no drastic change in the disordered region of all the three mutants. Similar to WT, the first residue M1 is predicted to be disordered in all the three mutants. The PONDR score around the mutation point is slightly affected. For example, for E11V and M12T, there is a decrease in PONDR score from residue number 5 to residue number 25 (E11V) and 26 (M12T), where E11V shows higher decrease in the score. However, mutation E22G resulted in alteration in PONDR score of 29 residues (from residue number 8 to residue number 36), where residues 19 – 22 showed a slight increase and all the rest showed decrease in the PONDR score. None of the change was prominent enough to shift the ordered NTD region from ordered to disorder. Moreover, according to PONDR score, the remaining protein part also remained unaffected.

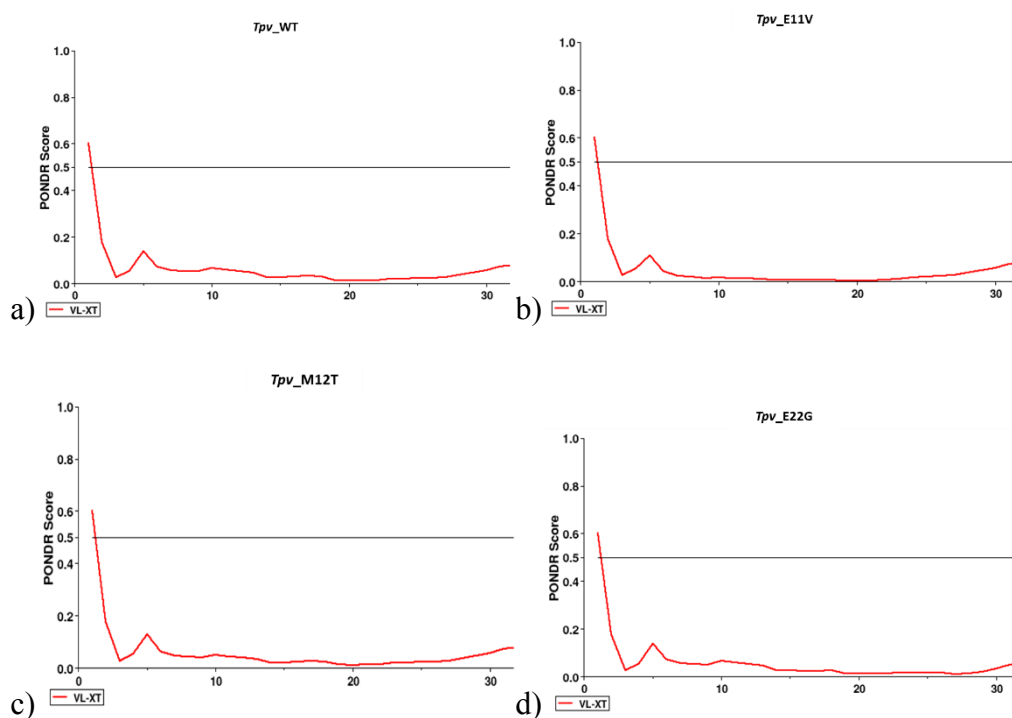


Figure 3.108 PONDR score plots of NTD of *Tpv* HSP 14.3 WT (a), E11V (b), M12T (c), and E22G (d). The x-axis represents the residue number while the y-axis shows the PONDR score of each residue. The grey line represents the threshold, above which is the disordered region and below, the ordered region.

3.13.4.1.3 PONDR output for distal NTD mutant variants

This group included five mutant variants, one being a mix of proximal and distal NTD mutation (F8YV31G). The two related mutants, F26YI27T and V23GF26YI27T, both showed decrease in hydrophobicity by 6.25 % and 9.38 %, respectively (Appendix E, Table 1). Both the mutants show increase in the PONDR score, almost throughout the length of NTD.

In case of V23GF26YI27T, with the exception of three residue points, 37 – 39, all residue points from 8 – 41 have a trend of increased PONDR score. The increase is 10 times at points around the mutation and dampens as moved further. This is evident in Figure 3.109b where this triple mutation tends to pull the NTD away from the x-axis and towards the threshold of PONDR score. The double mutant F26YI27T also shows similar trend, however, the increase in score is not as high as the triple mutant. Moreover, residues 12 – 14 and 33 – 39 show a decrease in PONDR score for F26YI27T (Appendix E, Table 16).

The mutations V31G (3.13% decrease in hydrophobicity) and V31IL33I (1.6% increase in hydrophobicity) exhibit exactly opposite results where the former results in slight increase in PONDR score and the latter causes a slight decrease in PONDR score from residue 13 - 45 (Appendix E, Table 17). However, all four of these mutants have only one residue, M1, predicted to lie in the disordered region (Figure 3.109). The mutant F8YV31G shows the characteristics of both the respective single mutants, F8Y and V31G. Similar to F8Y, the double mutant shows decrease in PONDR score from residue 3 – 16 and similar to V31G, the double mutant shows an increase in PONDR score from residue 17 – 45.

Overall results show that the mutation in the proximal region of NTD resulted in increase in disordered region. This is due to the fact that, that region was already lying close to the PONDR threshold score of 0.5.

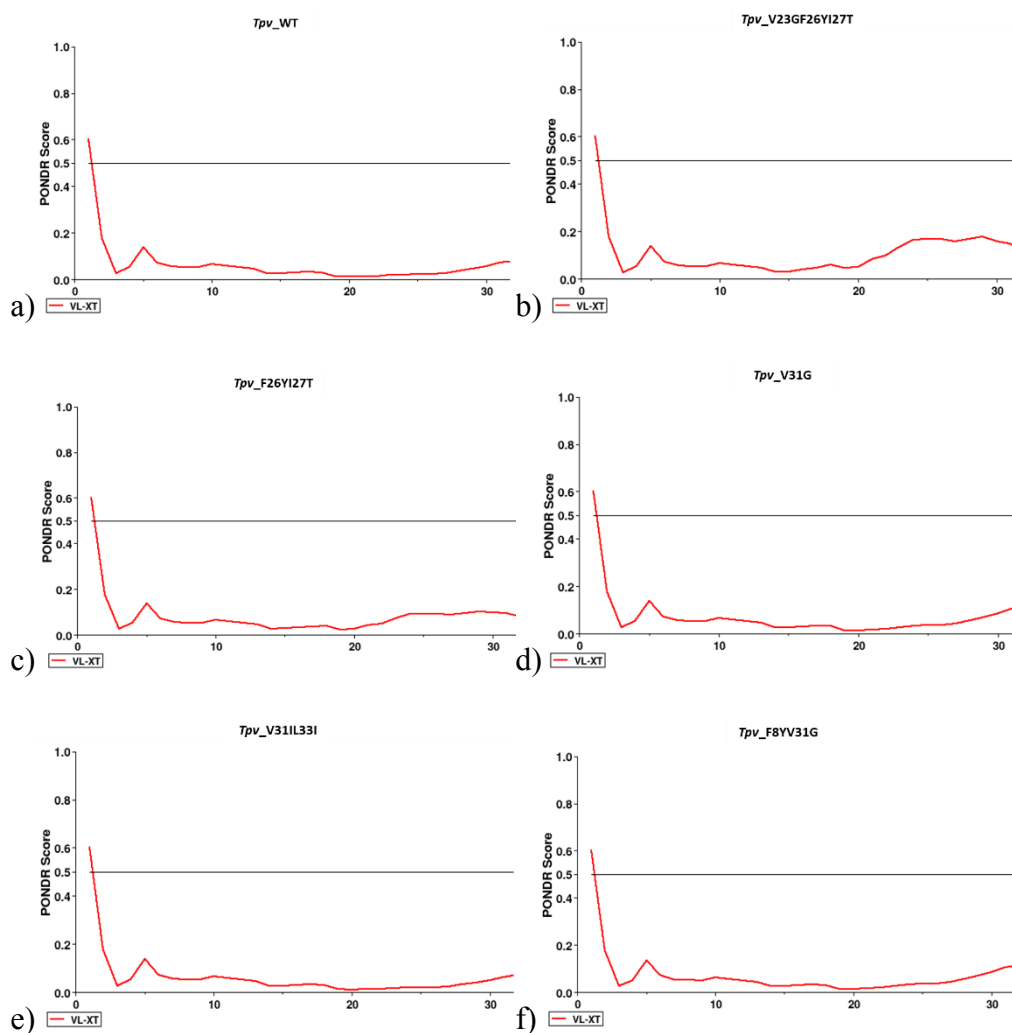


Figure 3.109 PONDR score plots of NTD of *Tpv* HSP 14.3 WT (a), V23GF26YI27T (b), F26YI27T (c), V31G (d), V31IL33I (e), and F8YV31G (f). The x-axis represents the residue number while the y-axis shows the PONDR score of each residue. The grey line represents the threshold, above which is the disordered region and below, the ordered region.

3.13.4.2 Analysis using PrDOS (Protein Disorder Prediction Server)

For confirmation and cross checking, another online tool, PrDOS was used for IDR analysis. The false positive (FP) rate, which is the percentage of residues falsely predicted to be disordered was set at 5% which is the default of the server. The output for *Tpv* sHSP 14.3 WT, as plot is shown in Figure 3.110. Each residue, along x-axis, is plotted versus its disorder probability score, along y-axis. Those that lie beyond

the threshold line of 0.5 are anticipated as disordered. Each residue has a specific disorder probability score which is also given in the form of table format, (Ishida and Kinoshita, 2007). When compared to PONDR, the output of both the online tools differs slightly. For the NTD, the PrDOS claims two disordered regions (residue 1 – 2 and residues 21 – 22), in contrast to PONDR (only residue 1 lies in disordered region). In ACD, where PONDR claims two disordered regions (61 – 68 and 73 – 78), PrDOS states presence of one disordered region comprising of 11 amino acids (72 – 82). In contrast to PONDR, the PrDOS predicts a disordered region in CTD of the protein from 118 – 124 amino acids. In addition to this, the two online servers differ in prediction about NTD. In PONDR, almost whole of the NTD lies away from the threshold of 0.5. However, as claimed by PrDOS, the initial 24 amino acids have a score of 0.4 or more, positioning them near the threshold of 0.5.

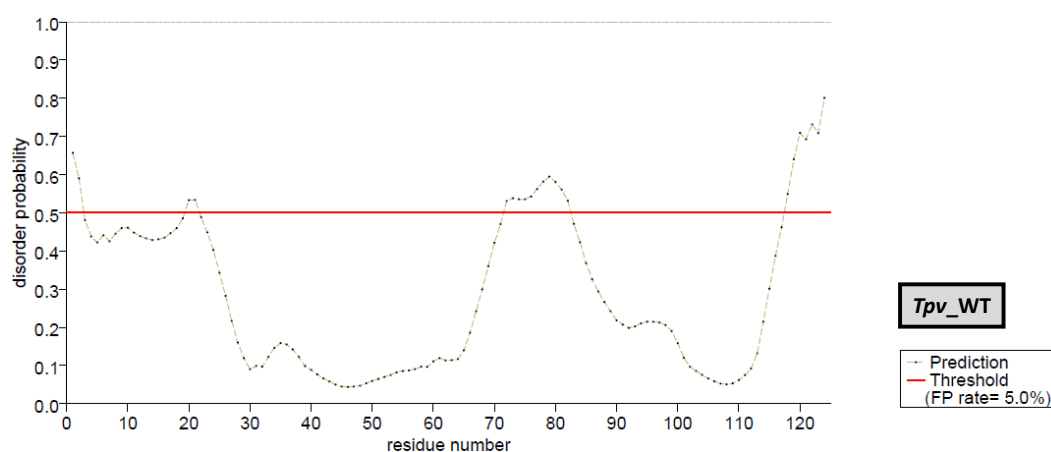


Figure 3.110 Graphical output from PrDOS of *Tpv* sHSP 14.3 protein. The x-axis represent each amino acid in the protein monomer. Disorder probability score of each amino acid is given on y-axis.

3.13.4.2.1 PrDOS output for NTD proximal mutant variants

Reducing the mean hydrophobicity by replacing a highly hydrophobic isoleucine with a neutral polar amino acid Threonine at position 5 resulted in a distinct increase (above 0.06) in probability score of two third of the NTD (initial 19 amino acids) and a moderate increase for the next 6 amino acids, shown in Table 19 (Appendix

E). This may lead to prediction of proximal 22 amino acids of the NTD (except 16th residue) as disordered, (Figure 3.111b). On the other hand, mutation of phenylalanine at position 8, by Tyrosine, caused slight increase in the score and probability of being disordered for initial 12 residues, (Figure 3.111c). However it resulted in decreasing the score and thus the probability of residues from 16 onwards almost throughout the NTD (until 29th residue), similar to PONDR output. The mutant I5T depicts increase in the disorder probability and F8Y caused some decrease in it, around the mutation region in the primary sequence. The double mutant, where both of the above single mutations are combined, I5TF8Y, exhibited a combined effect, *i.e.*, the increase in disorder probability is higher than F8Y, but lower than I5T, alone, from residues 6 – 28, as seen in Figure 3.111d. But for residues 1 – 6, the probability score is higher than both the single mutants. Since serine is one of the commonly found residue in the disordered region (Romero *et al.*, 2001), replacing a hydrophobic amino acid phenylalanine with serine, at position seven, in addition to the substitution of phenylalanine at position eight with a less hydrophobic tyrosine resulted in further elongation of the two disordered regions in the NTD, as compared to WT and F8Y single mutation variant, (Figure 3.111e). The first IDR consisted of initial 11 residues while the second IDR comprised of six residues as compared to two residues in WT and F8Y mutant protein. The increase in probability score caused by F7SF8Y double mutation is also higher than F8Y mutation alone. Combining all the three mutations I5T, F7S and F8Y in a single protein (I5TF7SF8Y), resulted in substantial increase in the disorder probability score of the NTD. This effect ultimately led to the combining of the two disordered regions of NTD in WT, which now appeared as a single long disordered region from residue 1 – 22, (Figure 3.111f), in the triple mutant. Although residues 23 – 25 also showed increase in the score, but did not pass the threshold of 0.5. This increase is the highest among all mutants studied and for some residues (13 – 15 and 17 – 21), match with the I5T single mutant. While considering the effect of mutations on whole protein, all the mutations showed a decrease of 0.01 – 0.05 in disorder probability score especially in the disordered region of the ACD and CTD, where the decrease was

most pronounced in the triple mutation protein, also shown by table (Appendix E, Table 19).

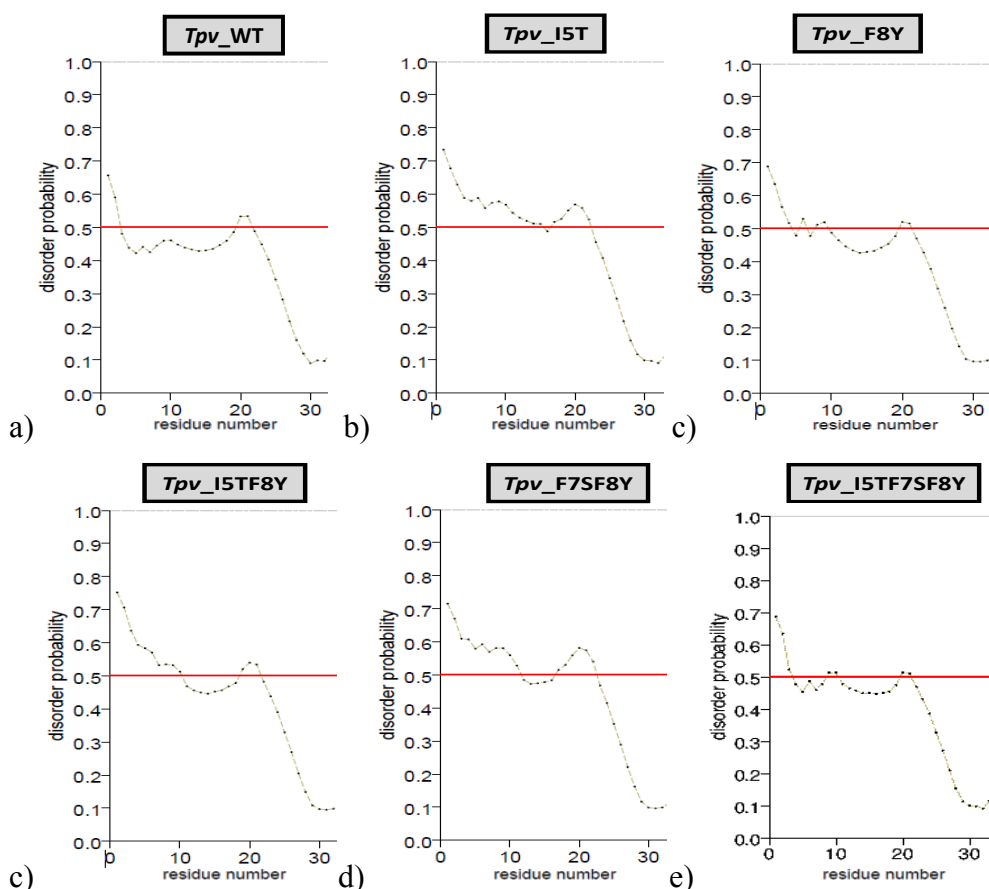


Figure 3.111 Disorder probability score plots of variants of *Tpv* sHSP 14.3; WT (a), I5T (b), F8Y (c), I5TF8Y (d), F7SF8Y (e), I5TF7SF8Y (f). The x-axis represents the residue number while the y-axis shows the disorder probability score of each residue. The red line represents the threshold, above which is the disordered region and below, the ordered region. The dotted line represent the disorder probability score of each residue.

3.13.4.2.2 PrDOS output for middle NTD mutant variants

The related mutations, in which the negatively charged residues are replaced with hydrophobic one, E11V and E22G, both do not result in drastic change in the disorder probability score of the proximal NTD. The results are similar to WT (Figure 3.112a and 3.112c); first three residues lie in the disordered region as compared to two in WT (E11V has residues 9 and 10 also in disordered region). This is also evident by the 3-D structures in Figure 3.88 and 3.90. However, the mutation

decreasing the hydrophobicity at position 12, M12T, resulted in a severely altered structure where prediction according to PrDOS is that the initial 14 residues are disordered (Figure 3.112b). However, 3-D structure of M12T (Figure 3.92) shows that residues from 1 – 7 are in a state of disordered coil. As a result, a single substitution at a highly conserved position 12, increased the length of the disordered region in the proximal NTD. Throughout the protein, there are alterations in the probability score (Appendix B, Table 20) for all the three mutants (also seen in 3-D structures) suggesting that a single point mutation might lead to such shift in the spatial arrangement of the variant protein.

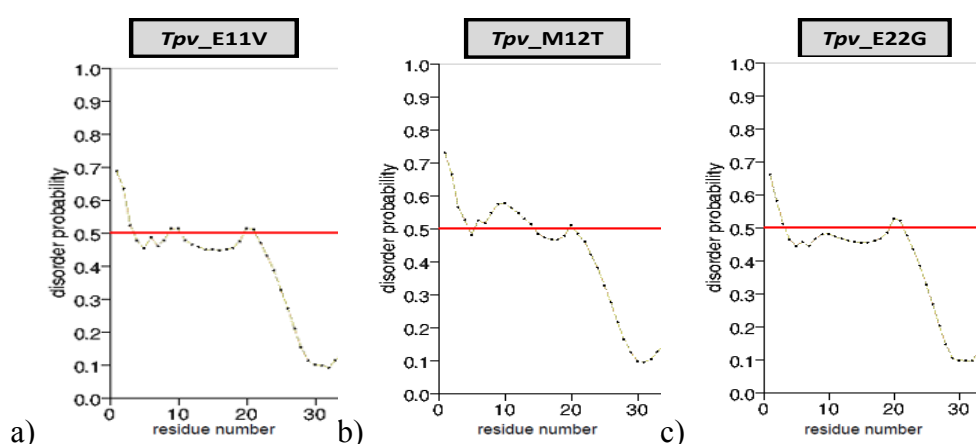


Figure 3.112 Disorder probability score plots of variants of *Tpv* sHSP 14.3; E11V (a), M12T (b) and E22G (c). The x-axis represents the residue number while the y-axis shows the disorder probability score of each residue. The red line represents the threshold, above which is the disordered region and below, the ordered region. The dotted line represent the disorder probability score of each residue.

3.13.4.2.3 PrDOS output for NTD distal mutant variants

In this group, considering the related mutants, F26YI27T seems to have rather severe effect (Figure 3.113c) on increasing the disordered region in the proximal NTD as compared to the related triple mutant, V23GF26YI27T (Figure 3.113b). The former mutant is predicted to have first 10 residues lying in the disordered region (except residue 7), while the latter has only three residues in this region, similar to WT, which has two. Not only this, the triple mutant also decreased the number of residues lying in the disordered region, in the latter NTD and ACD, by half (Appendix E, Table

21). The length of NTD predicted to be disordered is also increased by mutating Val31 to Gly in V31G (Figure 3.113d) and F8YV31G (Figure 3.113f), in the proximal NTD, where, both of the mutant variants have initial 9 residues in this range (Appendix E, Table 22). Moreover, the distal NTD in both these mutants show ordered behaviour, as opposed to WT. The 3-D structure of V31G (Figure 3.98) also coincides with this result. Similarly, the 3-D structure (Figure 3.100) as well as PONDR probability score (Figure 3.113e) reveals that the variant V31IL33I does not differ markedly from the WT, at the distal NTD region. However, this mutation increases the disordered region in the proximal NTD, according to PrDOS predictions.

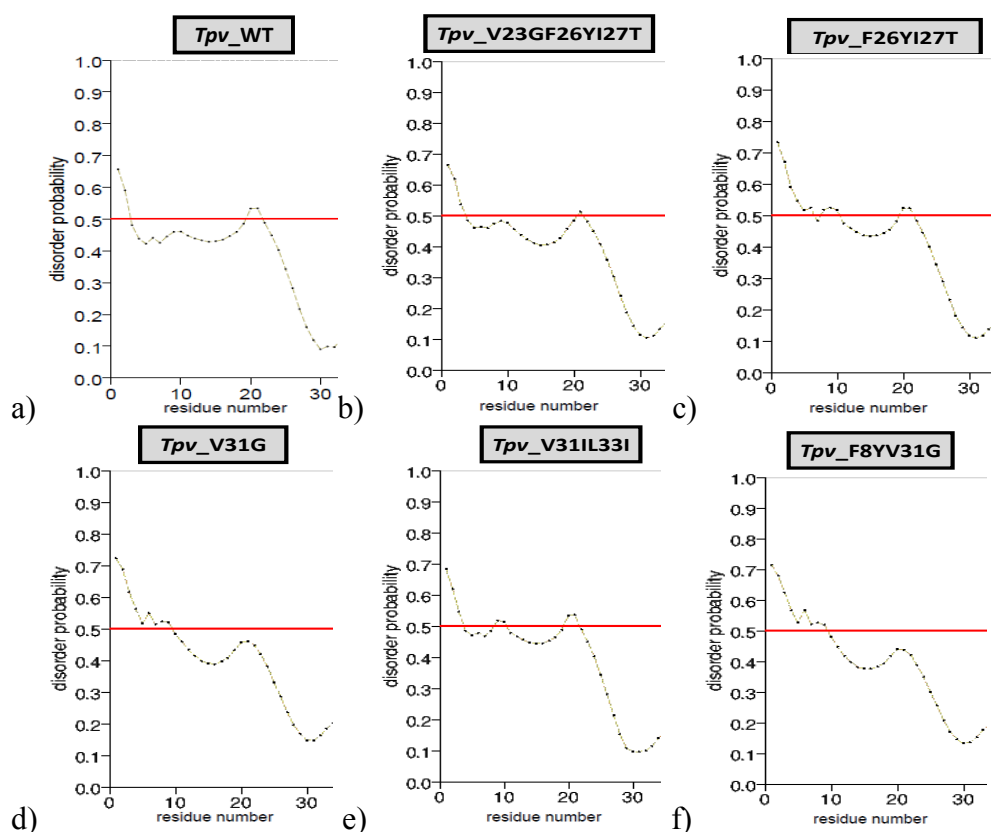


Figure 3.113 Disorder probability score plots of variants of Tpv sHSP 14.3; WT (a), V23GF26YI27T (b), F26YI27T (c), V31G (d), V31IL33I (e), and F8YV31G (f). The x-axis represents the residue number while the y-axis shows the disorder probability score of each residue. The red line represents the threshold, above which is the disordered region and below, the ordered region. The dotted line represents the disorder probability score of each residue.

3.13.5 Analysis of Structure and Surface Hydrophobicity of WT and Mutants by Chimera

The surface hydrophobicity of the dimer protein of *Tpv* sHSP 14.3 WT and those mutants which involve the decrease in hydrophobicity in the proximal region of NTD, was studied using Chimera 1.11.2 (Pettersen *et al.*, 2004). The surface hydrophobicity, as viewed from the top of the dimer, for *Tpv* HSP 14.3 WT is depicted in Figure 3.114. The black arrows point out the two NTDs in the dimers, and the orange red color in front of the black arrows demonstrate the presence of hydrophobic region of the proximal NTD where lies the beginning of helix and the three hydrophobic residues I5, F7 and F8, involved in a series of mutants produced, involving the NTD proximity. The yellow arrows indicate the CTD of both monomers in the dimer. The hydrophobicity exposed to the outer surface in a dimer can be essential for substrate recognition and binding. It is seen that the deep center of the dimer molecule is rich in the orange red color, indicating the burying of hydrophobic surfaces generally towards the inside of a globular protein, while hydrophilic/blue surfaces are facing towards the outer environment, expected to interact with water molecules around.

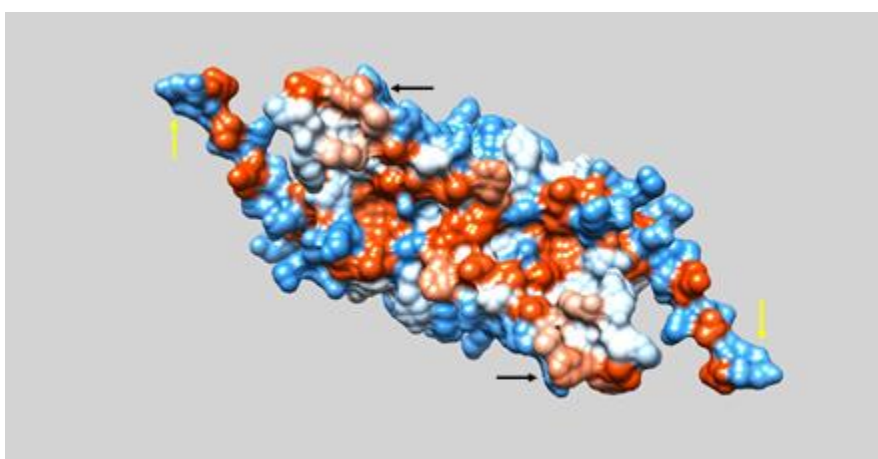


Figure 3.114 Three dimensional dimer structure of *Tpv* sHSP 14.3 WT showing the surface hydrophobicity. The dodger blue color indicates the hydrophilic surface, while orange red indicates the hydrophobic surface. Black and yellow arrows indicate NTD and CTD of the two monomers in a dimer.

Almost all mutations planned in this study show either a slight change or visible shift/bends, not only in the NTD but almost throughout the dimer structure, particularly in the terminals (NTD and CTD). This is clearly visible in the superimposed ribbon structures of each mutant, alongside the hydrophobicity surface figures. In the superimposed structures, the WT is colored as magenta while the mutant molecule has a blue color. Each mutant residue and its correspondence residue in WT are shown in stick model as well as labelled along with the monomer.

The mutation I5T, resulted in loss of dimer's surface hydrophobicity (in proximal NTD) (Figure 3.115b). It also caused a structural change where the whole dimer has shifted from the ends/terminals. The NTD shifts/turns towards outer side of the dimer and away from the center/axis or the monomer-monomer interface, as opposed to WT, where the NTD of the two monomers in a dimer, are facing towards each other. Additionally, the two CTD in the dimer of the mutant also face away from the dimer interface (Figure 3.115a), which is in contrast to the WT. The resultant dimer molecule appears shorter, and with exposed ends (NTD and CTD) and the two sides of dimer (ACD) more exposed, as compared to the WT.

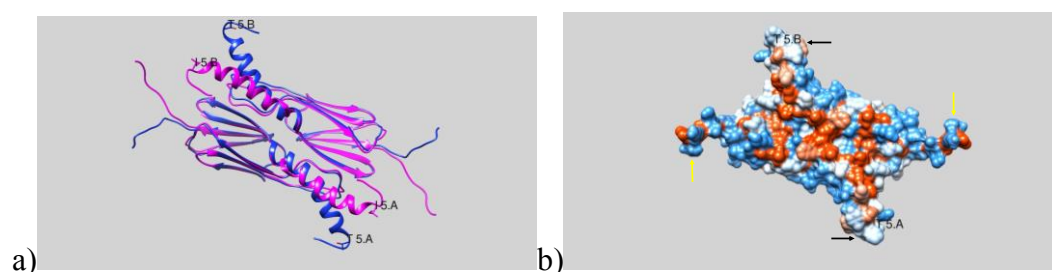


Figure 3.115 Effect of I5T mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of WT and I5T ribbon structures, (b) 3-D structure of I5T showing the surface hydrophobicity.

Another single mutation F8Y also resulted in decrease in the surface hydrophobicity at the position of mutation, (Figure 3.116b). The length of the dimer, as viewed from top, in the mutant F8Y remains similar to WT. However, in F8Y, the orientation of CTD in the 3-D structure is altered, similar to I5T mutant, *i.e.*, moving away from the center of the dimer. On the other hand, the NTD remains similar to that of WT, in terms of the number of helices and as well as its orientation. In both the single

mutations, I5T and F8Y, the CTD shows bending from residue Lys119 and Ile118, respectively.

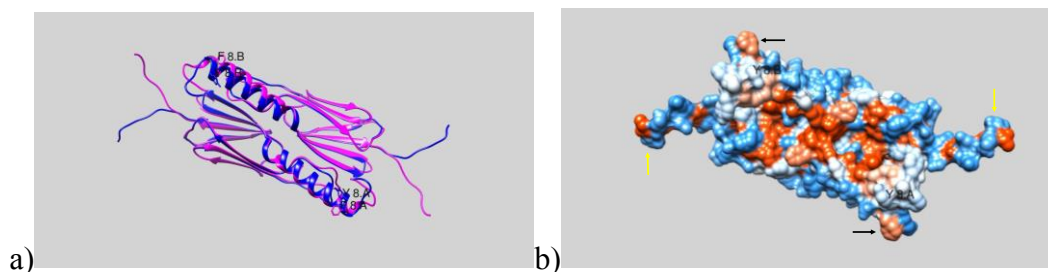


Figure 3.116 Effect of F8Y mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of WT and F8Y ribbon structures, (b) 3-D structure of F8Y showing the surface hydrophobicity.

Combining the two single mutations, I5T and F8Y, resulted in an obvious decrease/loss of surface hydrophobicity (Figure 3.117b) in the dimer at the two mutation points. However, interestingly, the arrangement of the monomers in 3-D structure of dimer of the mutant was very similar to WT. Although the variant protein showed a visible loss of surface hydrophobicity at mutation points, however, the overall shape and orientation of NTD and CTD were found to be very similar to WT, in 3-D structure, (Figure 3.117a). The distortion/bent present in the CTD in both the single mutants I5T and F8Y has completely disappeared on combining the two mutations in the single protein.

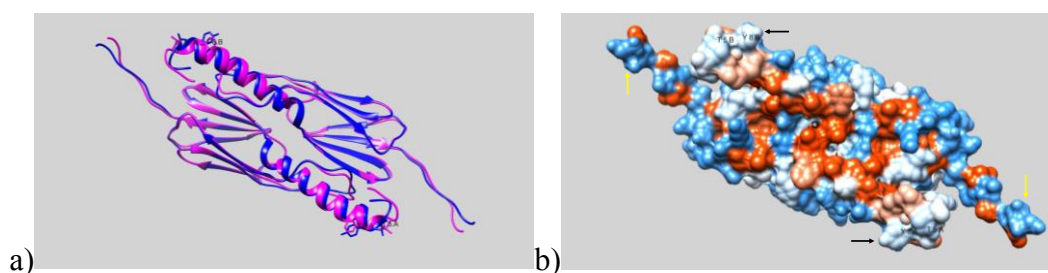


Figure 3.117 Effect of I5TF8Y mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of WT and variant ribbon structures, (b) 3-D structure of I5TF8Y showing the surface hydrophobicity.

Moreover, mutating the adjacent phenylalanine residues at position 7 and 8, resulted in loss of surface hydrophobicity, greater than double mutant I5TF8Y, (Figure 3.118b), however this loss is even more when overall hydrophobicity is decreased in

the proximal region (in triple mutation, I5TF7SF8Y), (Figure 3.119b). The double mutation F7SF8Y slightly effects the arrangement of CTD in the orientation (more upward and outward oriented in mutant) and slightly in ACD in the center of the dimer (Figure 3.118a). However, the simultaneous triple mutations, similar to I5T, effects the dimer structure to the maximum extent, seeming to making it compact, but the CTD arrangement is different from I5T and all other related variants protein, and is towards the dimer interface (Figure 3.119a).The NTD in both the variants, F7SF8Y and I5TF7SF8Y, exhibits an alteration where the residues 6 and 7 (Lys6 and Phe7), which were part of first helical turn in the WT, now become a coil like structure. This coil twists inward in the double mutant while in the triple mutant it bends further away from the dimer molecule.

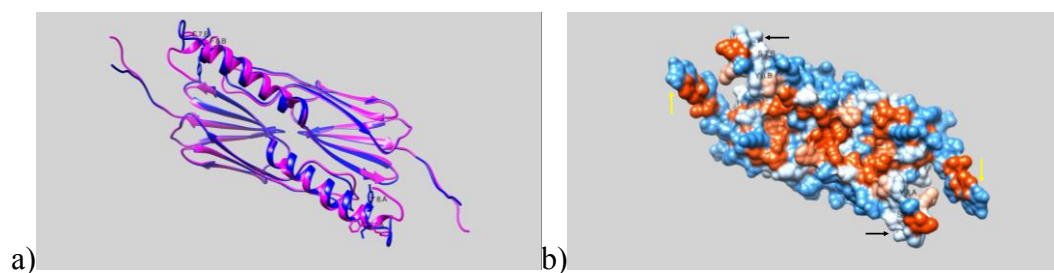


Figure 3.118 Effect of F7SF8Y mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of WT and variant ribbon structures, (b) 3-D structure of F7SF8Y showing the surface hydrophobicity.

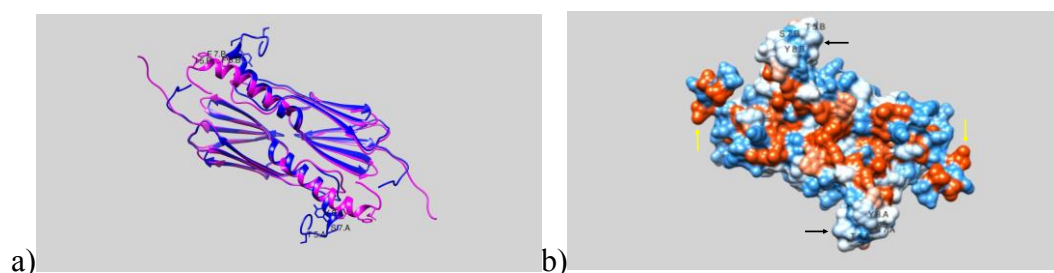


Figure 3.119 Effect of I5TF7SF8Y mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of WT and variant ribbon structures, (b) 3-D structure of I5TF7SF8Y showing the surface hydrophobicity.

A decrease in the hydrophobicity on the surface of the dimer, at the N-termini of the dimer molecule (Figure 3.120b) is also observed as a result of substituting hydrophobic residue Met12 with Thr. Not only this, the superimposition image

shows that this single mutation has resulted in converting the initial helix of the NTD to a unstructured segment, where Lys6 is now part of the coil, in contrast to WT (Figure 3.120a). This mutation also twists the NTD, pulling it away from the dimer center, similar to that seen in triple mutant I5F7SF8Y. The CTD of each monomer is also seen to bend away from the dimer interface, an effect observed in I5T and F8Y mutants. This might result in loosening of the dimer or in other words, a more open structure as compared to the WT, which in turn might be due to the loss of intramolecular hydrophobic bond between residues 8 and 12.

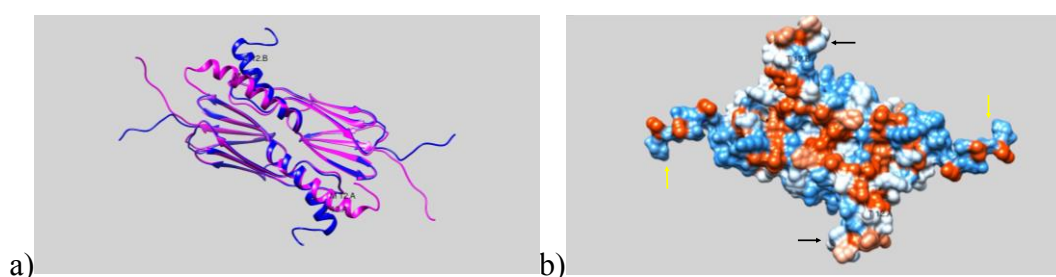


Figure 3.120 Effect of M12T mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of ribbon structures of WT and variant, (b) 3-D structure of M12T showing the surface hydrophobicity.

Increasing the hydrophobicity at position 11 (E11V), results in change in the orientation of CTD in three dimensional space. However, as described earlier, this mutation causes two-third of mainly terminal part of the NTD (bending initiates at Thr19) to bend further towards the dimer interface (Figure 3.121a). The CTD of this mutant variant dimer is pointed outwards, and although the overall size of the dimer seems to be almost similar to WT dimer. The space between CTD and NTD from the side of the dimer is reduced, making it a rather compact one. The surface hydrophobicity of the dimer is seen to be increased as indicated by the orange red color at the point of mutation (Figure 3.121b).

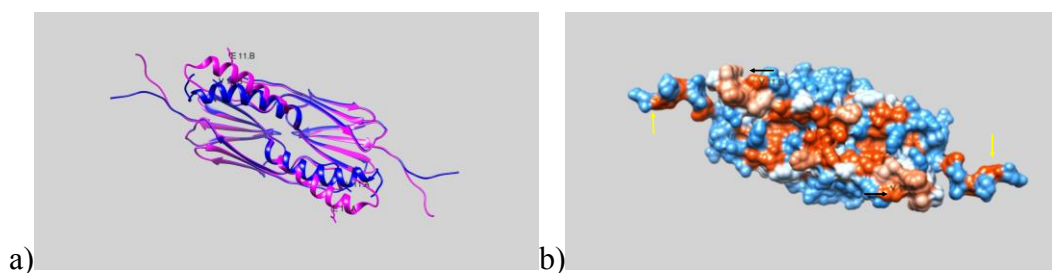


Figure 3.121 Effect of E11V mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of ribbon structures of WT and variant, (b) 3-D structure of E11V showing the surface hydrophobicity.

On the other hand, the substitution of glutamic acid at position 22, by a small and hydrophobic amino acid, Gly, results in almost an opposite action on the structure of NTD, as compared to E11V variant. The twist/shape of the coil consisting of initial five residues of the NTD remains similar to WT, however, the NTD bends far away from the axis/surface of dimer interface, as opposed to E11V mutant (Figure 3.122a). In addition to this, the CTD also undergoes structural changes and is seen to fold towards the dimer. This results in shortening of the size of the dimer in spatial dimensions. A slight increase in surface hydrophobicity (indicated by the orange color at the site of mutation) is also observed at the point of mutation (Figure 3.122b), which might be due to the fact that, the Gly residue is a short one with no side chain.

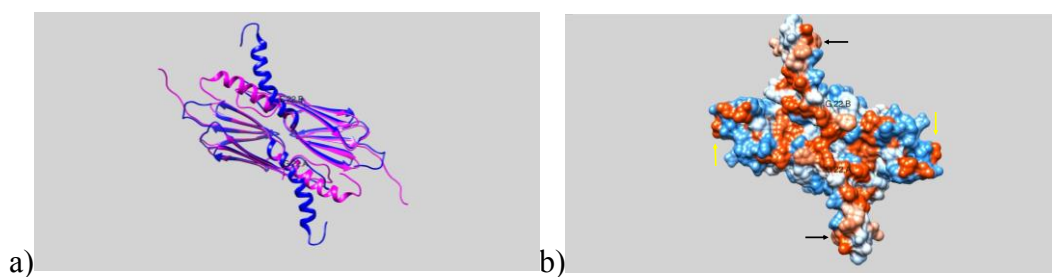


Figure 3.122 Effect of E22G mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of ribbon structures of WT and variant, (b) 3-D structure of E22G showing the surface hydrophobicity.

There was a visible decrease in the surface hydrophobicity in the central pocket of the dimer molecule which was contributed by Phe26 and Ile27 from both the monomers and was mutated in F26YI27T (Figure 3.123b). This hydrophobicity was further reduced in the triple mutant, V23GF26YI27T (Figure 3.124b), where the

center of the dimer was devoid of orange red tint and rich in hydrophilic colors. However, the spatial arrangement of NTD and CTD in both the related mutants have striking differences. The double mutant shows similar to WT structure and orientation (Figure 3.123a), the triple mutations simultaneously in the protein molecule lead to increase in the length of the disordered NTD. The helix in triple mutant starts from Thr9, while in the WT it starts from Lys6. In addition to this, the helical structure, as well as the CTD, bended away from the dimer center (Figure 3.124a).

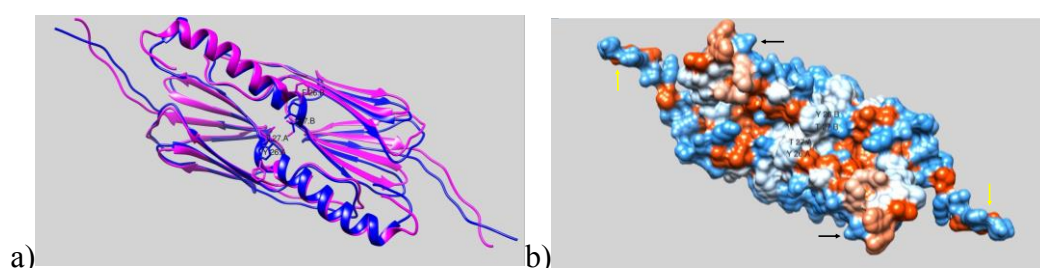


Figure 3.123 Effect of F26YI27T mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of ribbon structures of WT and variant, (b) 3-D structure of F26YI27T showing the surface hydrophobicity.

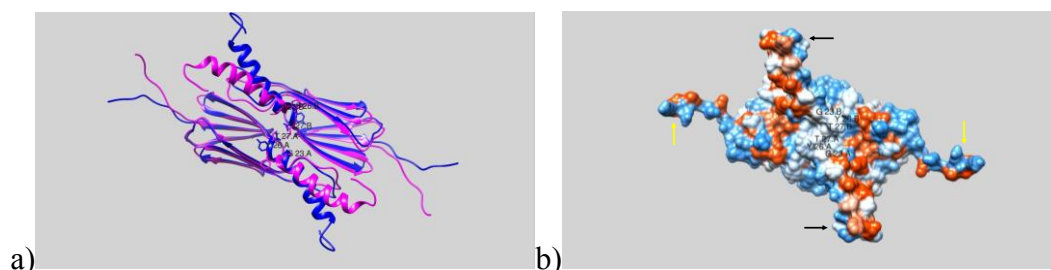


Figure 3.124 Effect of V23GF26YI27T mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of ribbon structures of WT and variant, (b) 3-D structure of V23GF26YI27T showing the surface hydrophobicity.

For the mutants V31G and V31IL33I, the surface hydrophobicity in the center of the dimer is decreased in the former (Figure3.125b), while visibly increased in the latter variant (Figure 3.126b) and the CTD arrangement in both the variants seem to take similar turns, *i.e.*, moving away from the center of the dimer. This bent in the CTD, occurs at the same residue (Lys119) in both the variants but is more pronounced in the double mutant (Figure 3.126a). However, the main difference between the two

variants is seen in the NTD where, substitution of Val31 with Gly resulted in twisting of the NTD from third helical turn, away from the dimer (Figure 3.125a). On the other hand increasing the hydrophobicity at position 31 and 33 simultaneously does not affect the NTD, which lays completely superimposed on the WT structure. This could be due to the loss of intramolecular hydrophobic interaction present between residues 31 and 44, in the WT and are not found in the V31G mutant. However, since the protection effect of both the mutants, against heat induced aggregation is similar to each other, this twist in the NTD of V31G might not affect its role in chaperone activity of the protein.

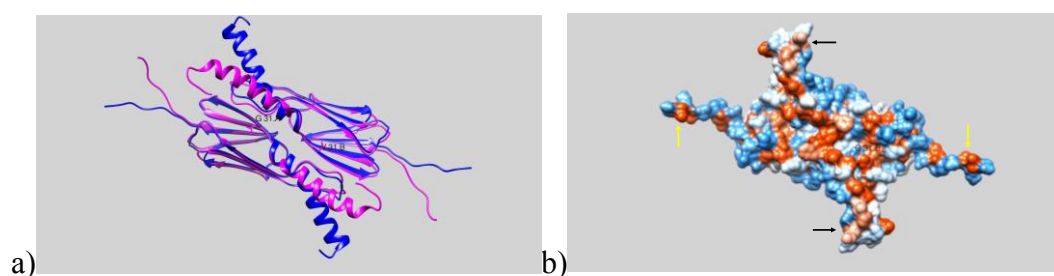


Figure 3.125 Effect of V31G mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of ribbon structures of WT and variant, (b) 3-D structure of V31G showing the surface hydrophobicity.

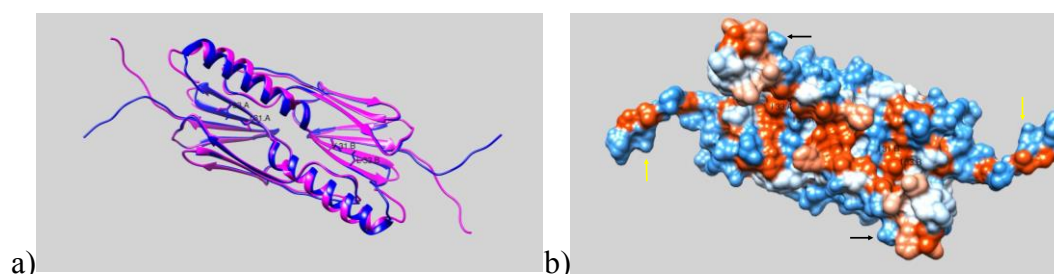


Figure 3.126 Effect of V31I/L33I mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of ribbon structures of WT and variant, (b) 3-D structure of V31I/L33I showing the surface hydrophobicity.

Combining a mutation from the proximal NTD and distal NTD, together in the same protein F8YV31G, resulted in a mutant dimer, that although lacked the surface hydrophobicity at the residue substituted points (Figure 3.127a), but the orientation defects of the NTD and CTD that occurred in the single mutants vanished in this double mutant (Figure 3.127b). The shape of the coil present in the beginning of

NTD, the length and spatial arrangement of the CTD and, therefore, the overall size of the dimer, all are similar to WT.

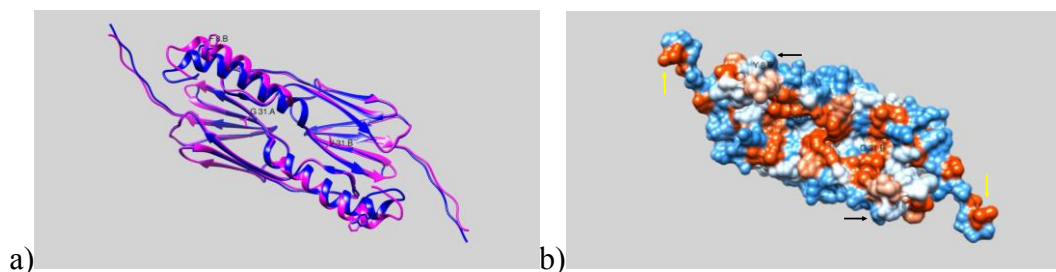


Figure 3.127 Effect of F8YV31G mutation on 3-D structure of *Tpv* sHSP 14.3. (a) Superimposition of ribbon structures of WT and variant, (b) 3-D structure of F8YV31G showing the surface hydrophobicity.

Overall, the least effect on the three dimensional arrangement of *Tpv* sHSP 14.3 is observed in the distant double mutation F8YV31G, where the size as well as the spatial arrangement of the whole protein molecule remains similar. Minor differences are seen in the proximal NTD and that might be due to the loss of intramolecular hydrophobic interaction present between residue 8 and 12 in the WT. The central part of the sHSP, the orientation of ACD and the parts of NTD and CTD close to ACD seem to be the least affected, if at all, by any of the mutations.

Since, the NTD and CTD are responsible for formation and stabilization of oligomers in many sHSPs (Heirbaut *et al.*, 2017), these changes in their orientation lead to alterations in dimer-dimer interaction which in turn will alter the oligomeric assembly of the mutants, as seen in the BN-PAGE results of some specific mutants. The mutations that most severely affected the surface hydrophobicity of the dimer are the triple mutations in the proximal and distal NTD, separately (I5TF7SF8Y and V23GF26YI27T). In addition to affecting the surface hydrophobicity, these two set of mutations also affected the proximal NTD the most, increasing the length of instructed NTD and its orientation in 3-D structure.

3.13.6 Thermodynamic Stability of Mutants

An estimation of thermodynamic stability enables to analyze the general trend about the $\Delta\Delta G$ values (Hawkes, Grutter and Schellman, 1984). The difference between the free energy of the WT and the variant, provides us the $\Delta\Delta G$ which helps in estimating/commenting on kinetic stability of the mutant protein. Two online tools/servers MUpro and I-Mutant2.0 were used to this end. There is a relationship between thermal stability and chaperone activity of sHSPs where decrease in thermal stability directly proportionate with a decrease in protection effect of a sHSP (Kazakov *et al.*, 2009). The $\Delta\Delta G$ numerical value can have a positive or negative sign, indicating increase or decrease in thermodynamic stability, respectively, following mutation. In MUpro, the confidence lies between -1 and 1, where a score less than 0 indicates a decrease in protein thermodynamic stability and *vice versa* (Cheng, Randall and Baldi, 2006). The reliability index in I-Mutant2.0 is between 0 -10 and indicates the efficiency of the output (Capriotti, Fariselli and Casadio, 2005).

The thermodynamic energy values for all the single mutants generated in this study are listed in Table 3.15. A general trend of decrease in thermodynamic stability was observed for all the mutants (except E11V in I-Mutant2.0) but the results from both the online tools are consistent only for some of the mutants. This could be due to difference in the algorithms and backbone structures used by these tools. The overall results indicate that decreasing the hydrophobicity in the any region of NTD leads to decrease in thermal stability although degree of decline varied greatly. According to MUpro the pronounced decrease in thermodynamic stability (values below -1.5) is seen in four of the six single mutant variants, namely, I5T, M12T, E22G and V31G. Considering the two mutant variants of the proximal NTD, I5T, the decrease in hydrophobicity nearest to the disordered region, shows higher decrease in thermodynamic stability. This mutant has shown drastic decrease in the ability to protect Citrate Synthase against heat induced inactivation. Thus, it indicates that, hydrophobicity in the proximal NTD of *Tpv* HSP, particularly near the disordered region might play an important role in recognition and/or binding with the substrate

as well as in thermodynamic stability. Among the two mutations, where charge is abolished and hydrophobicity is introduced, the variant E22G showed greater decrease in thermodynamic stability of the protein (according to MuPro, since I-Mutant reliability index for E11V is low). The maximum decrease in thermodynamic stability is exhibited when hydrophobicity is reduced at the junction of NTD and ACD, V31G mutation. This protein also was unstable at higher temperatures, as compared to WT, during heat treatment experiments (Figure 3.43).

Table 3.15 Thermodynamic values by I-Mutant2.0 and MUpro of single mutants involving the proximal part of NTD, studied.

	Results from I-Mutant2.0			Results from MUpro		
	$\Delta\Delta G$	Comment	Reliability Index	$\Delta\Delta G$	Comment	Confidence Score
I5T	-0.99	Decrease	6	-1.521	Decrease	-1
F8Y	-1.30	Decrease	8	-1.385	Decrease	-0.523
E11V	0.53	Decrease	0	-0.34	Decrease	-0.13
M12T	-0.59	Decrease	5	-1.82	Decrease	-1
E22G	-2.21	Decrease	8	-1.58	Decrease	-1
V31G	-5.11	Decrease	9	-2.09	Decrease	-1

3.14 *Tpv* sHSP 14.3 Mini-Peptides

In this study, two short peptides including the sequence of the NTD of *Tpv* sHSP 14.3 were designed, in order to investigate for their chaperone activity. A long peptide, comprising the whole NTD, (32 amino acids) was labelled as peptide1. Another peptide, a rather shorter one, comprising of the initial 12 amino acids of the NTD was labelled as peptide2 (see Figure 3.76). The chaperone activity of the peptides was investigated by their ability to protect the aggregation of substrate enzyme alcohol dehydrogenase from yeast (yADH), induced by heat and was defined as the percentage of protection against aggregation. In all the experiments, the scattering of light by yADH, without mini chaperones (control) at 43°C was taken as maximum and the protection thus was considered 0%. Aggregation was studied by light scattering of the aggregated particles with Multiskan Go by Thermo Scientific, using 96-well plate.

The aggregation of ADH was studied in the absence and presence of 0.1% - 0.91% acetonitrile concentrations. Our results showed that within this range, the organic solvent did not increase or decrease the aggregation of ADH. The experiments for each sample was performed in triplicate. The mean and standard deviation were calculated, therein.

Once the effect of the solvent on the substrate aggregation was confirmed, the ability of both the peptides to suppress the aggregation of yADH was studied. In all the reactions, the substrate (yADH) concentration was kept constant at 233 µg/ml, while the peptide varied between 10 µg/ml and 65 µg/ml. Overall result shows that, both the peptides exhibited protection against temperature induced aggregation of temperature sensitive enzyme, ADH from yeast.

The concentration of the peptide and the protection provided against aggregation were positively correlated, *i.e.*, the protection effect of both peptides increased with the increase in their concentration. The short peptide (peptide2) protected the yADH more effectively as compared to the NTD-peptide (peptide1), at concentrations 20 µg/ml and above (Figure 3.129 – 3.132).

Around 20% of the substrate was protected against aggregation at concentration of 10 µg/ml of peptide (Figure 3.128). On doubling the concentration of the peptide, the protection was increased, but not exactly doubled (Figure 3.129). However at concentrations of 20 µg/ml and above, the shorter peptide (peptide2) proved better protector against aggregation than the larger peptide (peptide1).

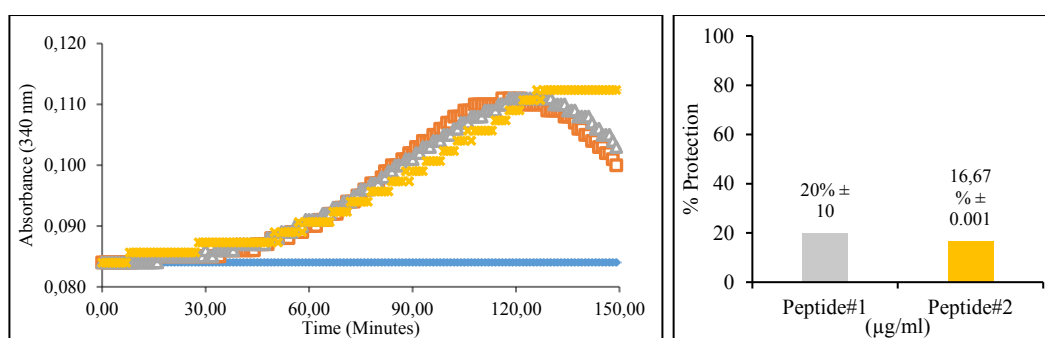


Figure 3.128 Light scattering plot (—♦—) Blank, (—□—) Control, (—▲—) Peptide1 and (—✦—) Peptide2 and bar chart (representing the slopes) of the percentage protection of yADH against heat induced aggregation in the presence of 10 µg/ml of peptide1 and peptide2, at 43°C. Percent protection is relative to the protection of the control, *i.e.*, in the absence of mini peptide (0%). Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

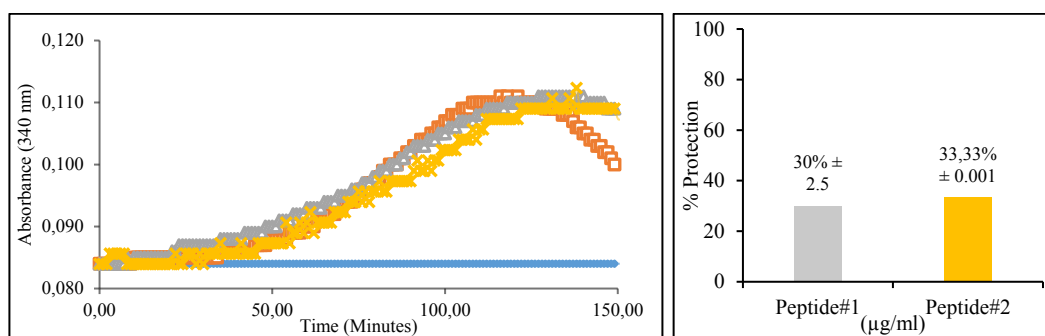


Figure 3.129 Light scattering plot (—♦—) Blank, (—□—) Control, (—▲—) Peptide1 and (—✦—) Peptide2 and bar chart (representing the slopes) of the percentage protection of yADH against heat induced aggregation in the presence of 20 µg/ml of peptide1 and peptide2, at 43°C. Percent protection is relative to the protection of the control, *i.e.*, in the absence of mini peptide (0%). Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

Around half of the given substrate was protected by peptide2 at substrate:chaperone weight ratios around 8:1. However, at similar milligrams concentrations, the whole NTD peptide protected only 40% of the substrate (Figure 3.130). In terms of molar ratios, this protection of 50% provided by the peptide2 was possible at substrate:chaperone molar ratios of 1:20. Further increasing the concentrations of chaperones to 40 µg/ml of the peptides, the percentage protection result was exactly double of what was achieved at 20 µg/ml of the peptides (Figure 3.131).

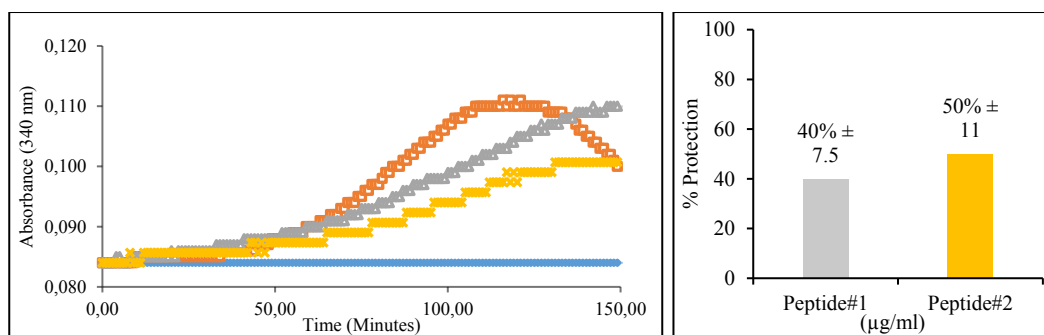


Figure 3.130 Light scattering plot (—♦—) Blank, (—□—) Control, (—▲—) Peptide1 and (—■—) Peptide2 and bar chart (representing the slopes) of the percentage protection of yADH against heat induced aggregation in the presence of 30 µg/ml of peptide1 and peptide2, at 43°C. Percent protection is relative to the protection of the control, *i.e.*, in the absence of mini peptide (0%). Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

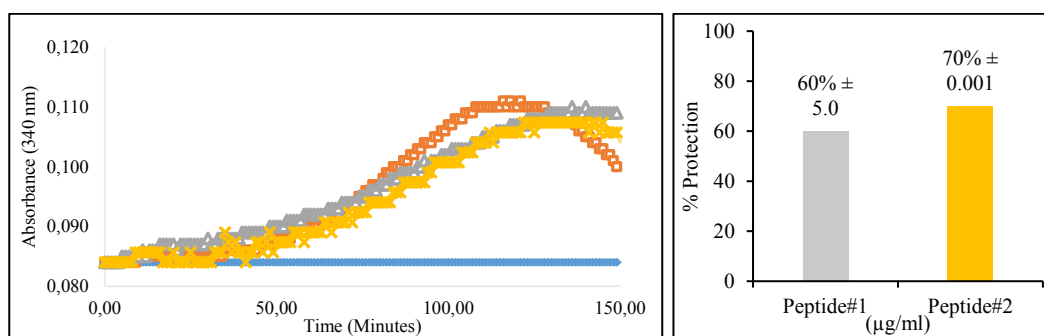


Figure 3.131 Light scattering plot (—♦—) Blank, (—□—) Control, (—▲—) Peptide1 and (—■—) Peptide2 and bar chart (representing the slopes) of the percentage protection of yADH against heat induced aggregation in the presence of 40 µg/ml of peptide1 and peptide2, at 43°C. Percent protection is relative to the protection of the control, *i.e.*, in the absence of mini peptide (0%). Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

The protection of 70% or above of the substrate against the heat induced aggregation was achieved, when the chaperone peptide concentration was increased to 5:1 substrate:chaperone weight ratio (Figure 3.132). However, increasing the concentration of the peptides to six times of what we started with (4:1 substrate:chaperone weight ratio; 50 µg/ml), protected nearly all substrate enzyme from aggregating at sub-optimal temperature. The small peptide showed complete protection of the ADH enzyme (Figure 3.133). These results confirm the positive

correlation between the concentration of the peptide and the percentage of protective effect on aggregating substrate.

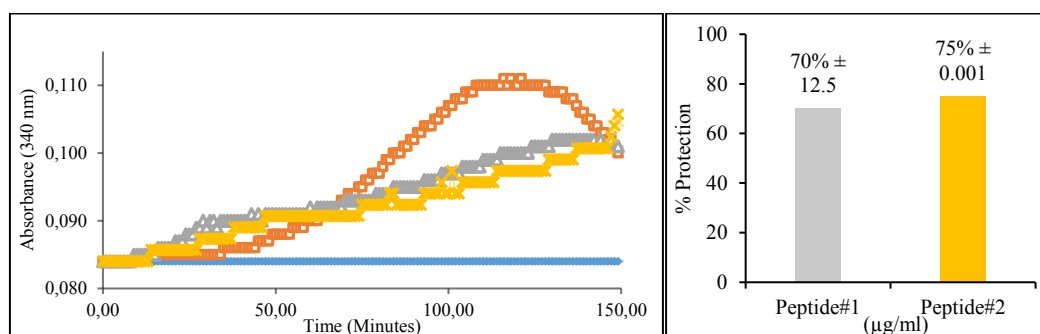


Figure 3.132 Light scattering plot (—◆—) Blank, (—□—) Control, (—▲—) Peptide1 and (—★—) Peptide2 and bar chart (representing the slopes) of the percentage protection of yADH against heat induced aggregation in the presence of 50 µg/ml of peptide1 and peptide2, at 43°C. Percent protection is relative to the protection of the control, *i.e.*, in the absence of mini peptide (0%). Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

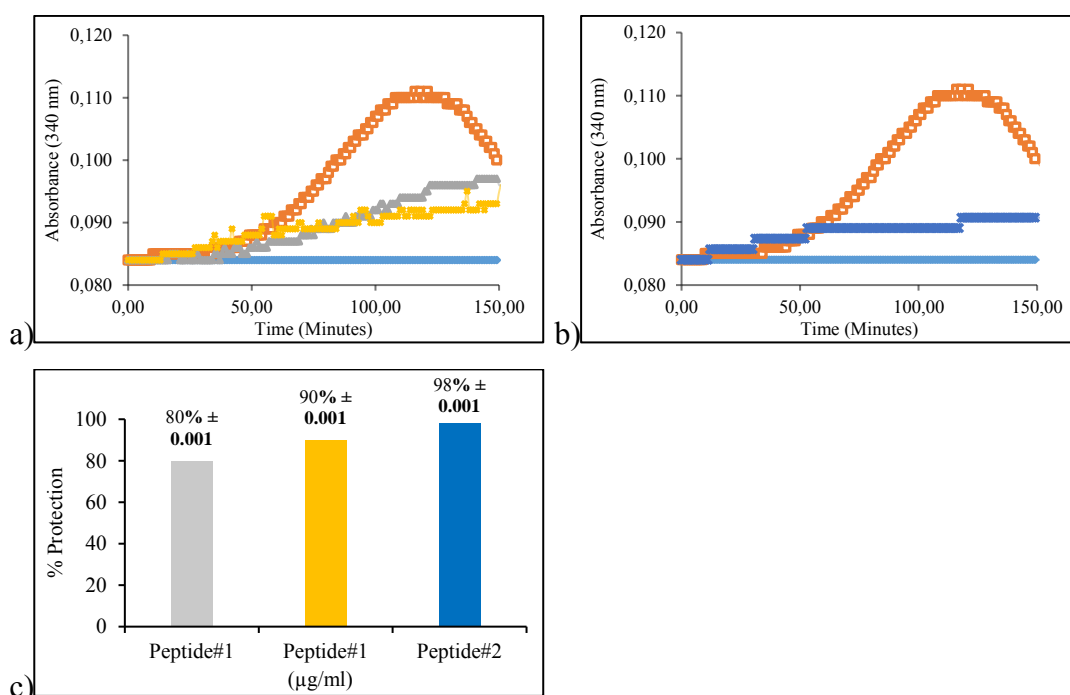


Figure 3.133 Light scattering plot (—◆—) Blank, (—□—) Control, (—▲—), Peptide1 (55 µg/ml) (—★—) Peptide1 (60 µg/ml), (—■—) Peptide 2 (65 µg/ml) and bar chart (representing the slopes) of the percentage protection of yADH against heat induced aggregation in the presence of peptide1 and peptide2, at 43°C. Percent protection is relative to the protection of the control, *i.e.*, in the absence of mini peptide (0%). Each data point represents the mean of three independent trials and is shown as the mean ± standard deviation (SD).

CHAPTER 4

DISCUSSION

At any time, cells face different stresses. The most common ones are when the eye protein is exposed to UV-light, and when other body cells are exposed to oxidative stress. sHSPs play an important role as a first line of defense, to protect the cells from the deleterious effects of these stresses by making the cell more resistant to the stressors such as reactive oxygen species and divalent ions, iron and mercury (Janowska *et al.*, 2019). Natural mutations in the HSPBs in human have been a cause of several disease (Boncoraglio, Minoia and Carra, 2012). Increase in their production is seen in aging cells (De Los Reyes and Casas-Tintó, 2022). In addition to this, sHSPs have been shown to serve as promising therapeutic agents for the treatment and cure of retinal neurodegenerative diseases and other neuropathies (Phadte, Sluzala and Fort, 2021).

In terms of length and sequence, the most diverse segment in a sHSP is its NTD. Besides a few conserved fractions or residues in related domains/species, the remaining primary sequence of any NTD shows huge diversification. The NTD contributes to the structure of the sHSP by stabilizing the oligomers and to the function of the sHSPs, by recognition of the partially unfolded proteins which are the substrates for sHSP. Structurally, some part of the NTD attains a helical configuration while other parts are present in a disordered state (Boelens, 2020). A recent study shows that the deletion of an initial part of sHSP NTD (14 residues), of *Synechococcus* sp. sHSP, did show a decrease in the chaperone effect of the respective sHSP but the effect imparted was not so drastic. However, complete deletion of the NTD (along with CTD) resulted in the complete loss of the protection effect of sHSP (Biswas *et al.*, 2021).

It has long been speculated (Raman and Rao, 1994) that the sHSPs recognize its substrate, the partially unfolded protein, by virtue of its hydrophobic residues. This is seen in a naturally occurring mutation P20L, in human HSPB6, where, increasing hydrophobicity by replacing the aromatic proline, towards the proximal NTD, with aliphatic and highly hydrophobic leucine, takes place. The resultant mutant protein displayed higher protection effect to the substrate with increase in the NTD hydrophobicity, while an increase in aggregation at 70°C (Shatov and Gusev, 2020). Nonetheless, the oligomeric nature of the protein is not affected by this mutation. However, not always, an increase in hydrophobicity is followed by an increase in chaperone function. For example, introduction of a positive charge in place of hydrophobic, proline, P20R of the HSPB6, is a naturally occurring mutation leading to posterior polar cataract, which is inherited as autosomal dominant trait (Xia *et al.*, 2014). This mutation resulted in alteration in the secondary structure, which is followed by increase propensity of the protein to form fibril like structure and drastic decrease in protecting unfolding proteins (Ghahramani *et al.*, 2020). On the other hand, abolishing a highly conserved charge and replacing it with polarity, R7S, a naturally occurring mutation in human HSPB3, displays phenotypically as idiopathic peripheral neuropathy (Kolb *et al.*, 2010). Not only this, a single mutation in α A-crystallin sHSP, in the positive charged residue, R21, with glutamine results in cataract formation, however, in-vitro, the chaperone activity with ADH and insulin was observed to be increased by virtue of increase in hydrophobic surface, with no effect on structure and oligomerization of the protein (Phadte, Santhoshkumar and Sharma, 2018). While another study involving the positively charged arginine in the highly conserved SRLFDQFFG sequence of HSPB8 of humans, showed that introducing hydrophobicity at this position led to formation of larger oligomers and also affected the chaperone activity (Shatov, Sluchanko and Gusev, 2021). This conserved motif also plays a role in formation of heterooligomers among the human sHsps (Shatov, Strelkov and Gusev, 2020). Single substitution mutations in the NTD of HSPB1, resulting in reversing or introduction of charges, has resulted in variety of neuropathy diseases (Nefedova *et al.*, 2015). In addition to this, the positively

charged lysine in the NTD is expected to play a role in stabilization of oligomers by crosslinking between subunits in an oligomer (Lambert *et al.*, 2011). As mentioned above, the highly conserved positively charged residue, arginine, has been widely studied, there are not many studies present on negatively charged residues in the NTD, especially glutamic acid. Therefore, the two highly conserved glutamate at position 11 and 22, in *Tpv* sHSP 14.3, were studied in this thesis.

Following sub-cloning of the gene of *Tpv* sHSP 14.3, TVN0775, into pET23a(+), the expression of WT was analyzed in both BL21(DE3) and T7-Express competent cells, where the former expression system exhibited high protein concentration in the cell lysate. The T7-Express competent cells also manifested protein expression, however, the expressed protein level in the respective cell lysate was slight, as seen by the density of the band on SDS-PAGE. The same results were also observed for all the variants whose cell lysates were initially prepared in T7-Express competent cells. The expressed protein of *Tpv* sHSP 14.3 was found to be stable at temperatures of 60°C - 80°C. Thermostability of the proteins from *Thermoplasma volcanium* in general is within this temperature range since the source organism is a moderately thermophilic one. Other studies also have shown that sHSPs from thermophilic or hyper thermophilic archaea were stable enough when subjected to high temperatures (75°C or 80°C). Therefore, this step has been employed by them in the first stage for the purification of the protein (Kim *et al.*, 1998, 2003; Usui, 2004; Saji *et al.*, 2008). In this study, chaperone activity of the purified protein *Tpv* sHSP 14.3 WT was analyzed by estimating the protection capabilities of the protein against thermal aggregation of a temperature sensitive mesophilic enzyme, citrate synthase (optimal temperature 35°C), at two different molar ratios; 1:7 and 1:35 substrate:chaperone ratios. Studies have established this fact that sHSPs have the ability to protect citrate synthase enzyme from aggregating due to unfolding, when exposed to higher than its optimal temperature (35°C). Different ratios have been reported to be sufficient for protecting citrate synthase at higher temperature depending on the origin of the sHSPs. For example, a sHSP from *Rana catesbeiana*, hsp30 efficiently protected the CS from aggregation (up to 93%) at merely a 1:1 molar ratios and showed complete

protection at higher sHSP concentrations (Kaldis, Atkinson and Heikkila, 2004). Similar results were also obtained with sHSP from *C.elegans*, HSP16-2, which prevented aggregation of citrate synthase at 1:1 and higher sHSP concentrations (Leroux *et al.*, 1997). However, Hsp16.3 from *M.tuberculosis* tends to protect citrate synthase above 90% (almost completely) at 1:9 substrate:sHSP ratio (Chang *et al.*, 1996), while CRYAA and CRYAB at 1:0.6 substrate:sHSP ratios were adequate to completely protect the substrate from aggregating, when exposed to 43°C (Santhoshkumar and Sharma, 2001). In this study, the *Tpv* sHSP 14.3, protected the pig heart citrate synthase enzyme against thermal aggregation effectively (~70% protection) at 1:7 substrate:sHSP molar ratios. At the higher concentration of sHSP (1:35 substrate:sHSP molar ratio) the protection rate was increased to 80%.

We also analyzed the ability of this *Tpv* sHSP 14.3 WT to assist the citrate synthase in restoring its activity after exposing it to high temperature (47°C for 10 minutes) in the presence of chaperone. Without the chaperone, citrate synthase activity was almost completely lost within 10 minutes of high temperature (47°C) exposure. At temperatures above 35°C, citrate synthase aggregates irreversibly and thus leads to loss of activity. However, the WT of *T.volcanium* sHSP, when present in excess quantities (1:1700 substrate:sHSP molar ratio) or (0.44 µg/ml CS and 214 µg/ml chaperone, *i.e.*, 1:486 substrate:sHSP weight ratio) restored the activity of the enzyme almost 81% of positive control. Similar results are seen for chaperones like Hsp18.1 and Hsp17.7 from Pea and sHSP from *M.tuberculosis* (Lee, Pokala and Vierling, 1995; Chang *et al.*, 1996) and also was previously presented by our lab, that *Tpv* sHSP 14.3 protein in excess concentration (400 µg/ml) was able to protect the CS activity completely (Kocabiyik and Aygar, 2012). This protective action of sHSP exhibited concentration dependent nature and, therefore, the shielding effect of the chaperone against aggregation was reduced by 30% and 11% as compared to positive control and 1:1700, respectively, when the sHSP concentration was reduced by half, *i.e.*, 1:850 substrate:sHSP molar ratio.

The structural characterization of sHSP under study in this research was done by resolving its oligomeric structures and their distribution on a Blue-Native PAGE gel. From the structures determined to-date, one of the most notable features of sHSPs is their organization as large oligomeric structures, where the outside diameters range from 100 Å to 180 Å and the oligomers are in the form of hollow ball like structures, but the number of subunits in the oligomers vary: *M. jannaschii* Hsp16.5 and yeast Hsp26 complexes have 24 subunits, whereas *M.tuberculosis* Hsp16.3 has 12. Hsp16.9 from wheat, meanwhile, is a barrel-shaped structure assembled from two hexameric double disks with a total of 12 subunits (Kim, Kim and Kim, 1998; Van Montfort *et al.*, 2001; Kennaway *et al.*, 2005). In contrast to these defined oligomeric structures, the quaternary structure of α -crystallin is variable, *i.e.*, it forms polydisperse oligomeric assemblies with up to 50 subunits per complex (Haley, Horwitz and Stewart, 1998). The 24-mer, observed for the archaea *M.jannaschii*, was observed also for *Tpv* sHSP 14.3 WT at room temperature, in this study. We know that where the ACD is responsible for dimer formation, the partially resolved NTR seem to be important in stabilizing the oligomers, via hydrophobic contacts, especially in eukaryotes, like wheat Hsp16.9 (Van Montfort *et al.*, 2001). Therefore, it is expected that change in hydrophobicity will result in altered oligomeric states of the variants. Oligomerization is likely to depend on the successive union of subunits driven by the interaction of oligomerization determinants, leading to increasingly larger assembly units such as dimers, trimers or tetramers (Koteiche and Mchaourab, 2002). On the other hand, oligomerization is thought to be required for substrate binding and chaperone function, although oligomer roles vary from one sHSP to another (Avilov, Aleksandrova and Demchenko, 2005). For efficient recognition and binding to the substrate proteins, the majority of sHSPs shift the equilibrium of the oligomeric assembly from an inactive state which comprise of a high fraction of large oligomers to an ensemble weighted toward smaller species representing the active state (Haslbeck *et al.*, 2005) (Haslbeck *et al.*, 2015), which is necessary to reveal the hydrophobic NTR, previously hidden in the oligomer. Therefore, presence of unfolded substrate or presence of conditions that trigger unfolding of substrate

proteins, will shift the sHSP equilibrium towards active state (Benesch *et al.*, 2008). Studies have shown that heat stress is the general trigger for activating sHSPs (Haslbeck *et al.*, 2015). However, the temperature range for activation varies with physiological temperature of organisms, for example for yeast Hsp26 it is 20°C to 43°C, but for thermophilic *M.jannaschii*, it ranges from 60°C to 95°C (Haslbeck *et al.*, 1999, 2008). The oligomeric states of *Tpv* sHSP 14.3 WT at room temperature and after heating at 70°C when resolved through BN-PAGE gel, showed the presence of the characteristic 24-mer oligomer in addition to other smaller and larger oligomers. A band that corresponds to 12-mer is present in both, the non-heated and heated, samples, however another distinct band present adjacent to 12-mer band is seen to disappear upon heating. This may be due to conversion of the oligomer of that size to smaller species. Such dissociation of larger oligomers to dimers is previously revealed in studies, for example, the 24-mer oligomer of HSP26 dissociated into dimers at temperature 20°C and above, and this dissociation increased with increase in temperature (Skouri-Panet *et al.*, 2006). Also at any time and temperature, presence of a number of species in an equilibrium is also recently shown in HspB1 of Chinese Hamster (Kurokawa *et al.*, 2021). Other distinct species present on the solution at any time are dimer and 60-mer oligomers. The 24-mer and 60-mer species of sHSPs are reported in most of the studied sHSPs. For example, a recently studied sHSP from a virus, cyanophage is reported to possess oligomers with 24, 48 and 60 subunits (Biswas *et al.*, 2021).

The 3-D structure of the monomer and dimer of *Tpv* sHSP 14.3 was generated using Easy Modeller 4.0 and UCSF Chimera, 1.11.2 respectively. The structural elements are a combination of α -helices (NTD), β -sheets (ACD), coils and unstructured segments (CTD). Similar structure of sHSP has been reported previously by Liu, J. Y. Chen, *et al.*, (2015) for the sHSP of *Sulfolobus solfataricus* (SsHsp14.1). In the dimeric structure of SsHsp14.1, the NTDs are facing away from each other in a dimer and the β_6 of each monomer overlaps the β -sheets of the other monomer to form interactions with β_2 . This type of structural organization forms the basis of dimer formation typical for non-metazoans sHSP (Stamler *et al.*, 2005). In the 3-D model

of *Tpv* sHSP 14.3, orientation of the NTDs are also away from each other in a dimer and does not show any intermolecular interaction in the initial part of the NTD. In fact the first monomer-monomer interaction is seen in *Tpv* sHSP 14.3 at position 27, where the isoleucine of both the monomer form hydrophobic interaction with a distance of 3.98 Å. It has been reported that, in Hsp16.3 of *M.tuberculosis*, the hydrophobic surfaces remain buried or are not exposed (probably interacting with other hydrophobic residues) in the inactive state and get exposed on exposure to a stressor (chemical or temperature) (Yang *et al.*, 2008). Therefore, the hydrophobic interaction between Ile27 of the two monomers, might possibly be important during subunit exchange in *Tpv* sHSP 14.3, to recognize and interact with the substrate proteins. Further along the NTD of the *Tpv* sHSP 14.3, valine at position 31 is also involved in forming intermolecular hydrogen bond between monomers with glutamine at position 80. Previously, it is reported for the sHSPs, Ta16.9 and Ps18.1, they interact with substrate in the dimeric state and, therefore, the intermolecular interactions stabilizing the dimer are considered important for chaperone activity (Santhanagopalan *et al.*, 2018). Therefore, the intermolecular interaction between Val31 of one monomer and Gln80 of the other, in *Tpv* sHSP 14.3 dimer, might be important for maintaining intact dimer structure, before and/or during its interaction with substrate. As expected from its primary structure, the NTD of the *Tpv* sHSP 14.3 is highly hydrophobic in nature. Considering Gly, Ser, Thr, Cys, Asn, Pro and Gln as neutral aminoacids, Glu and Asp as acidic, Arg, Lys and His as basic, while all the rest as hydrophobic amino acids, the NTD of *Tpv* sHSP 14.3 is composed of 46.88% hydrophobic amino acids in nature (Figure 3.67). The rest of the pie is shared between neutral, basic and acidic having 37.5%, 9.38% and 6.25% share, respectively. This hydrophobicity has been the main concern in this research, along with two acidic residues, to understand the significance of each targeted residue in the structure and function of the sHSP. 3-D structure analysis of human HSPB6 monomer revealed that the NTD adjacent to ACD arranges itself in the tertiary structure such that it overlays the β 3 and interacts with it (Weeks *et al.*, 2014). Similar structural arrangement is observed for *Tpv* sHSP 14.3 monomer that is

possibly established by interaction of valine 31, the last hydrophobic residue in the NTD of *Tpv* sHSP 14.3, with alanine 44 (residue lies in β 3). It was previously reported that equivalent residue of HSPB6, V64, makes interactions with the hydrophobic pocket (residues V92 and V132) of the ACD of adjacent dimer (Weeks *et al.*, 2014).

Lens protein gets exposed to UV stress and is prone to form fibrils, however, it is protected from aggregation by, other proteins called crystallins. The crystallins are, therefore, present in abundance in the eye, and they themselves have high propensity to form β -amyloid fibrils, under specific conditions. Small heat shock proteins possess structurally very similar ACD of eye crystallins. For this reason they are also susceptible to form fibril like structure that result in aggregates. Generally the proximal and distal part of the ACD are prone to form fibrils (Carver *et al.*, 2017). When viewing the overall picture of the protein *Tpv* sHSP 14.3 WT, similar to these reports, ACD, particularly the part of ACD lying close to the NTD is more inclined to forming β -amyloid fibrils (see Figure 3.91), as revealed by Rosetta Energy study using online tools. This is probably balanced by the rest of the protein which show less incline towards forming such fibrils. This has been reported by Abeln *et al.*, (2008) who studied the aggregation propensity of central region (rich in hydrophobic peptides) with or without flanking disordered regions. They report that the presence of highly aggregating region in the middle of unstructured polypeptides is advantageous for decreasing the propensity of aggregation (Abeln and Frenkel, 2008). In case of WT *Tpv* sHSP 14.3, considering the NTD alone, there is a perfect balance, where out of the 21 energy bands present, slightly above half (12) does not cross the Rosetta Energy limit, while the rest half (9) show the propensity to form fibrils. Therefore, the equilibrium point is shifted slightly towards less propensity of the NTD to form fibrils and, may play an important role in maintaining the soluble state of the protein. Thus any disturbance in this balance, as seen later, by amino acid exchanges, may or may not shift the equilibrium towards increase fibril propensity thus affecting the structure and in turn the function of the sHSP.

The dimer of a sHSP is not only the building block of oligomers, but most likely this is also the state that the sHSP interacts with the partially unfolded substrate. It is hypothesized that the sHSP, especially, the NTD and CTD section thus may undergo transition from ordered to disordered state, during dissociation of oligomers, to interact with the substrate (Bardwell and Jakob, 2012; Kovacs and Tompa, 2012). Therefore, the ordered-disordered balance of the NTD is tried to be analyzed in this state for a better outlook on the mechanism of substrate recognition by the sHSP. It has been numerous times proved that NTD is intrinsically disordered and it is thought that the NTD of both monomer in a dimer exhibit an equilibrium between an open (more disordered) and closed state (more ordered/helical) (Klevit, 2020), where the hydrophobic residues present on the NTD become available as a large hydrophobic surface for interaction with a diverse range of substrates (Jaya, Garcia and Vierling, 2009; Patel, Vierling and Tama, 2014; Baughman *et al.*, 2020). Recently, there has been provided some proof that the ACDs in a dimer also unfold partially, attaining a disordered state, during dissociation of monomers and, therefore, contribute to substrate recognition, binding or stabilization after binding (Alderson *et al.*, 2020). Not only sHSPs, but also many other intrinsically disordered proteins play a role in neurodegenerative diseases, where folding in unfamiliar state of these disordered structures of proteins results in aggregates that might trigger its onset or progress of the disease (Ayyadevara *et al.*, 2021). Overall results from PONDR and PrDOS, predicting the IDR in the *Tpv* sHSP 14.3 show that the proximal NTD is anticipated to be in a disordered state. Since both methods use primary sequence of the monomer as template, they predict intrinsic ordered/disordered state. Similar predictions have been reported for HSPBs of human, where either a segment (of HSPB4), or segments (of HSPB1, HSPB2, HSPB3, HSPB5, HSPB6 and HSPB10) or whole NTD (of HSPB7, HSPB8 and HSPB9) are foreseen as being disordered (Webster *et al.*, 2019). Combining the predictions of the two software together, the ACD of *Tpv* sHSP 14.3 was predicted to be intrinsically disordered towards the middle (overlapping residues of both software, 72 – 78). Comparably, besides HSPB2 and HSPB3, all the other known human HSPBs are also predicted to have a disordered region, towards the

middle of the ACD. Although the PONDR result does not predict any other disordered region in *Tpv* sHSP 14.3, the PrDOS result predicts 118 – 124 residues to be disordered. Parallel results have been shown for all human HSPBs where either the whole CTD or a part of it is anticipated to have high score to be expected in disordered state (Webster *et al.*, 2019).

It can be summarized that, WT of *Tpv* sHSP 14.3 is a sHSP chaperone that has the capacity to protect the substrate protein (pig heart citrate synthase) at sub-optimal temperature from aggregating by forming stable and soluble complexes with it. After cooling the solution, by withdrawal of the sHSP, CS can refold in such a way as to retain its enzymatic activity to almost complete extent. This could be possible due to the perfect balance of Rosetta Energy of all the hexa-peptides of the *Tpv* sHSP 14.3 polypeptide, so that the overall propensity to form fibril like structure is low. In addition, the equilibrium of the oligomer distribution of the chaperone between smaller species (the active state) and larger ones (the storage form) should be optimum.

4.1 Group 1: Decreasing hydrophobicity in the proximal NTD

This group involves all the mutants which involved the substitution in the initial 10 residues of the NTD of *Tpv* sHSP 14.3, (I5T, F8Y, I5TF8Y, F7SF8Y and I5TF7SF8Y). All the variants bunched in this group displayed high protein concentration in BL21(DE3) *E.coli* expression system, similar to WT, in the cell free extract. In addition to this, all the afore-mentioned variants' proteins demonstrated high level of thermal stability when exposed to temperatures as high as 80°C. This character of the proteins proved helpful and thus it was employed, similar to WT, as the initial step for the purification of these proteins from the rest of the host *E.coli* proteins that were soluble in the lysis buffer. As compared to others, the growth rate of the I5TF8Y and F7SF8Y mutant variants were slow. This retardation in *E.coli* growth could be possibly explained in a way that the mutant sHSP may have a high affinity to a single or set of proteins (that are critical for bacterial growth) and form

stable complexes with them, thus interfering or causing hindrance in the normal growth pace.

All the five mutants grouped here exhibited the property to guard the substrate (CS) from aggregating due to unfolding when it was exposed to a continuous high temperature. The single mutant F8Y, and the double mutants I5TF8Y and F7SF8Y exhibited concentration independent behavior and the extent of their protection was 69%, 60% and 40% at both molar ratios (1:7 and 1:35 substrate:sHSP), respectively. However, for I5T and I5TF7SF8Y, an increase in activity was observed with increase in the chaperone concentration, similar to WT, thus showing concentration dependent behavior. The protection effect of I5T was 29% less than that of WT at lower concentration and when the concentration was increased by five times, the effect was 20% less than that of WT. In other words, I5T showed 40% and 60% protection at 1:7 and 1:35 CS:sHSP molar ratios, respectively. However, the ability of this mutant to protect the activity of substrate from heat, was concentration independent and almost half of that of WT. The results from IDR shows that single mutation decreasing hydrophobicity at position 5 results in increase in the disordered region of the monomer. This can be explained by the carbon H-bond between isoleucine and threonine at positions 5 and 3 respectively, present in the WT. This bond disappears in the mutant and a new bond of the same nature appears between threonine and phenylalanine at position 5 and 8, respectively. As a result of shift in the intramolecular hydrogen bonding, the proximal segment of NTD may become more disordered. This might lead to structural re-arrangement in such a way, that the 24-mer oligomeric specie, characteristic of archaeal sHSPs and also seen in *Tpv* sHSP 14.3 WT, is absent in this mutation, as seen in its BN-PAGE results. Although, similar to WT, the 60-mer oligomer presence is witnessed. Besides, the two oligomers, (36-mer and 16-mer) are found in I5T mutant but they are absent in WT protein. This oligomeric behavior would have been made possible due to structural changes at dimer level, (Figure 3.2). The dimer has attained an open structure with the spatial arrangement of NTD and also CTD away from the dimer core, (Figure 4.1). Thus the decrease in the chaperone activity of the mutant I5T could be due to

decreased surface hydrophobicity of the dimer as well as, structural re-arrangement in the proximal NTD that might lead to a decrease in substrate recognition, and/or a decrease in stability of substrate-sHSP complex or both.

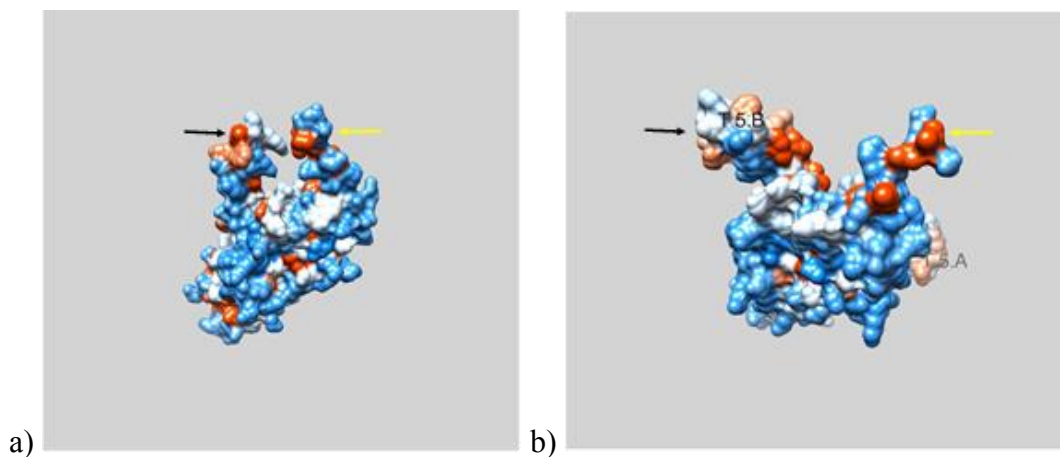


Figure 4.1 Side view of the dimer of WT (a) and I5T mutant variant (b). Black arrows point towards NTD, yellow arrows towards CTD.

The other single mutant F8Y, in addition to the larger species and 36-mer oligomer like I5T, also showed the presence of 24-mer oligomer. This could be the reason that this mutant demonstrated chaperone activity somewhat similar to the WT only at lower chaperone concentration in terms of protecting the CS from aggregation. F8Y preserved CS from getting aggregated at higher temperature and the protective effect was 69% at both lower and higher molar ratios. Similar to WT, this mutant chaperone had a concentration dependent effect of chaperoning. By maintaining the CS in a soluble state at high temperature, the mutant F8Y, also enabled the substrate to retrieve its native folded state. The remaining CS activity was 38% and 66% at low and high molar ratios, respectively (Figure 4.2 and 4.3). The effectiveness was same as WT (69%) at low molar ratio but less than WT at higher molar ratio, (81% of WT at 1:1700). Examination of the interactions gained and lost due to this mutation, at molecular level reveals a stunning picture. The summary of which is that, in the WT, phenylalanine at position 8 was in contact with residues glutamic acid (an electrostatic interaction and a hydrogen bond) and methionine (a hydrophobic interaction and a hydrogen bond) at position 11 and 12, respectively. When

substituted by tyrosine, the position 8 becomes closer to the proximal NTD and forms new interactions with methionine (an electrostatic interaction and a hydrogen bond) and isoleucine at position 1 and 5, respectively. Thus, as expected, the intrinsic disorder (at NTD terminal) in the F8Y mutant is very similar to that of WT (for residues 1 and 2 and), rather a decrease in PONDR score is observed from residues 3 – 22 as compared to WT, thus pushing them further away from the threshold of being declared disordered. Looking at the 3-D structure obtained by Chimera, besides the obvious decrease in surface hydrophobicity, the NTD seems to be less disturbed than the CTD in this mutant. The Rosetta Energy values of this mutant, also, do not show significant difference from the WT. However, the decrease in ability of this mutant to protect the activity of substrate enzyme (as compared to WT) can be explained by the fact that mutations causing a decrease in disordered state of IDR can result in decrease of protein activity (Vacic and Iakoucheva, 2012; Vacic *et al.*, 2012). Previous reports showed that when hydrophobicity is reduced by substituting phenylalanine with a less hydrophobic amino acid, alanine or a charged one, the sHSPs (particularly metazoans sHSPs) have lost their protection ability against heat or chemical induced aggregation of substrate (Plater, Goode and Crabbe, 1996; Horwitz *et al.*, 1998; Pasta *et al.*, 2003; Hanazono *et al.*, 2013b).

When the above single mutants are combined in a single protein, I5TF8Y, the result is an effect, which lies in the middle of that of both of the single mutants. The reduced effect of I5T mutation in terms of protection of CS from aggregation is compensated by F8Y mutation and therefore, the % protection value is more than I5T and less than F8Y. This mutant was slightly less effective than two point mutants in protecting the activity of the CS enzyme at 1:850 CS:sHSP molar ratio. However at 1:1700 molar ratio its protection efficiency of CS activity from high temperature was higher than I5T mutant and lower than F8Y mutant (Figure 4.3). This double mutant was about 10% and 20% less efficient than WT in preventing aggregation of the CS at lower (1:7) and higher (1:35) CS:chaperone molar ratios, respectively. The oligomeric species distribution by this mutant was almost like combination of the single mutants and showed the same effect, where it possessed the 36-mer similar to I5T and 24-

mer, similar to F8Y. The disordered region in the double mutant I5TF8Y is similar to I5T, according to PONDR software, but PrDOS results show the double mutant having disordered region of the mutant proximal NTD, whose extent is more than the F8Y and less than I5T alone. This can be explained by the interactions at molecular level, where the hydrophobic interaction of at position 8 are lost while electrostatic interactions are retained, in I5TF8Y. As explained earlier, the F8Y mutation causes residue 8 to make interactions with the residues of the proximal coil (residues 1 and 5), thus pulling the coil towards the helix and decreasing the PONDR score of the proximal region. When F8Y is combined with I5T (I5TF8Y) the residue 8 shows no interaction with residues 1 and 5 and, therefore, the length of the disordered region is greater than F8Y (similar to I5T). The 3-D spatial arrangement of the NTD and CTD in the dimer of the mutant I5TF8Y is very much similar to that of WT, like the values of Rosetta Energy of all the hexa-peptides around the mutation sites. Increased disorderedness may adversely affect oligomer stability and alter their distribution. Besides, obvious decrease in the hydrophobicity could be the reason of its decreased efficiency in protection of CS from aggregation and its activity at high temperature, relative to WT.

Of the three phenylalanine residues present in the NTD of *Tpv* sHSP 14.3, two are in its proximal part (at position 7 and 8) and lie adjacent to each other. The hydrophobicity of this part was drastically decreased by replacing one phenylalanine with serine. The resultant mutant F7SF8Y showed detectable changes in the quaternary structure. In contrast to WT and F8Y, no oligomeric species larger than 24-mer was observed for this mutant. However, similar to WT, it showed presence of 24-mer, 12-mer and dimer species while 12-mer is not found in F8Y. Besides the obvious decrease in the surface hydrophobicity of the dimer, the dimer of the mutant, as a whole is more compact than that of the WT and F8Y mutant (Figure 3.5). Increased compactness could be explained by the fact that phenylalanine at position 8 forms a single hydrophobic interaction with methionine 12. The mutant, in addition to this interaction, also forms another hydrophobic interaction with lysine at position 6, thus pulling the proximal NTD towards the center. This interaction and absence

of electrostatic interactions were not found in F8Y mutant. The protection of substrate against aggregation of this mutant was (40%) less than WT, while only 26 – 27% of the enzyme activity was protected at low and high substrate:sHSP molar ratio (Figure 3.54). Moreover, efficiency for protecting the activity and aggregation was concentration independent. These results may suggest that it could be possible that this mutant formed stable complex with the substrate, which could be long lasting and/or the chaperone might not release the enzyme at all, or the chaperone might have bound to the enzyme in such a way so as to block the active site of the enzyme. A reverse mutation, serine-to-phenylalanine was performed for human sHSP, HSPB6 that functions as cardio-protectant in humans. The results showed that, switching polarity with hydrophobicity at position 10 in NTD, S10F, resulted in increase in the chaperone activity of the protein (Shatov and Gusev, 2020). The reverse in our case that decreased chaperone activity. So the two results are consistent. The Zipper DB results show that the energy values for the hexa-peptides ${}^5\text{IKFFTN}_{10}$ and ${}^7\text{FFTNE}_{12}$ increased significantly, while that for the latter peptide, passed the threshold of -23Kcal. Since now 10 out of 21 hexa-peptides have propensity to form fibrils, the equilibrium of fibril forming tendency of the protein, as a whole, is shifted towards right. This might be the reason for low protection capacity of the activity of the enzyme. Although the PONDR results indicate the disorder region of the F7SF8Y mutant to be short and similar to F8Y, results from PrDOS indicate a drastic effect of this double mutation on the proximal NTD coil. Residues from 1 – 11 and 17 – 22 are predicted to be disordered as compared to WT, which has only 1 – 2 and 20 – 21 residues of NTD in disordered region. This result is supported by superimposition results from chimera, where it is observed that the length of the coil in the proximal NTD, of F7SF8Y, is increased to residue 7, while the coil is restricted to residue 5 in WT. It has been shown that substitution, F → S, in ordered region which resulted in conversion of the ordered region to disordered one, altered the protein activity and this accounts for 3.6% of all of the studied order to disorder mutation diseases (Vacic and Iakoucheva, 2012).

The last mutant in this group was an attempt to combine three single mutations in a protein, *i.e.*, I5TF7SF8Y. As far as intrinsic disordered regions are concerned, the results were according to expectations. According to PONDR, the initial seven residues were found to be disordered, as compared to two or one residues being disordered in other mutants. On the other hand, PrDOS affirmed the residues up to 22 passing the threshold of 0.5 and thus established as disordered. There are reports on mutations in the IDRs that result in a structural change, transforming the disordered region into ordered one, and the resulting protein losing its function almost completely. For example, a single mutation at position 166 from Serine to Phenylalanine in Troponin I, cardiac muscle, resulted in decreasing the PONDR score at that specific position by 0.2. A same point mutation in another protein Tumor suppressor p53 at position 46 (S to F) resulted in decrease in PONDR score by 0.29 (Vacic and Iakoucheva, 2012). Phenylalanine at position 7, when replaced by serine, a complete reverse result is seen in our study. Furthermore, the triple mutant I5TF7SF8Y and single mutant I5T, (which show substantial changes in ordered to disordered region transformation), have shown drastic decrease in chaperone activity as compared to WT. Studies have shown that a decrease of 0.05 PONDR score caused by a single mutation can result in loss of protein function leading to a disease. The interactions at molecular level of this triple mutant are similar to WT, while the Rosetta Energy values are similar to the double mutant F7SF8Y. While the chaperone functional results and structure analysis results from Chimera and PrDOS show this triple mutant to have striking similarities to the single mutant I5T. The dimer in the 3-D spatial arrangement obtained from Chimera shows the dimer to be as more of sphere or square, different than the longitudinal/cylindrical WT dimer but the most similar to I5T. Moreover, the BN-PAGE results of this mutant are exactly same to the I5T mutant, except for 16mer which is only in I5T. Not only in structure, the efficiency of protection of activity of CS, of the triple mutant is similar to I5T especially at 1:1700 molar ratio (Figure 4.3). The thermodynamic study of all the mutants in this group also support the general behaviour of the mutants. The mutant with low thermodynamic stability I5T, F7SF8Y and I5TF7SF8Y, displayed lower

chaperone activity as compared to others, in terms of protecting CS against inactivation.

Many studies predict that phenylalanine in NTD may have noticeable substrate interaction which is lost after any mutation in the residue. This was mainly attributed to the observations that phenylalanine in NTD proximal part interact by hydrophobic interactions with phenylalanine in the ACD (Van Montfort *et al.*, 2001). Since oligomer or higher order species formation is necessary to stabilize the chaperone-substrate complex, such a mutant where both phenylalanine are replaced will show less protection effect. In a study with *Schizosaccharomyces pombe* sHSP, substitution of the phenylalanine residues at positions 6, 7 and 9 of the proximal NTD resulted in loss of oligomer formation and consequently decrease in chaperone activity (Hanazono *et al.*, 2013b). In another study, in human HSPB6, substitution of Q31 to Leucine led to increase in chaperone activity, while F33 to Alanine replacement resulted in decrease in chaperone activity (Heirbaut *et al.*, 2014). In another study, sHSP from *Drosophila melanogaster*, DmHsp27, substituting phenylalanine at position 29 and 31 with a more hydrophilic tyrosine or less hydrophobic alanine, had similar or decreased effects on protection of CS aggregation as compared to WT (Moutaoufik *et al.*, 2017). Highly conserved, NTD phenylalanine at position 27 of α B-crystallin of Humans and mouse, is of pronounced significance and a single mutation at this position abolishes the chaperone activity, either partially or completely for various protein substrates. This single mutation in human α B-crystallin, at elevated temperatures, results in not only the complete loss of protection effect, but the sHSP itself aggregates and precipitate (Horwitz *et al.*, 1998). Similar to our F7SF8Y results, the replacement of phenylalanine F24 as well as F27 by a charged residue arginine in mouse α B-crystallin significantly obstructs its chaperone activity. So much so, even replacing the phenylalanine F27 by another, comparatively less hydrophobic, alanine residue also was not sufficient to protect chemically and thermally denatured protein substrate (Plater, Goode and Crabbe, 1996). This suggests that F24 in mouse α B-crystallin and F27 in both mouse and human α B-crystallin is not only important for

interaction with the substrate, these residues also have critical role in the stability of α B-crystallin at temperatures above 60°C. The two adjacent phenylalanine residues in the NTD of *Pisum sativum* PsHsp18.1 when mutated, separately, to decrease the hydrophobicity and substrate interaction, oligomer sizes of the mutant variants were comparable to their WT. Similar result was observed in the F8Y mutant. Moreover, the chaperone activity of the mutant, PsHsp18.1, was also similar to its WT (Jaya, Garcia and Vierling, 2009). We also observed that F8Y mutant variant's chaperone activity (protection against heat induced aggregation of CS at 1:850 substrate:sHSP molar ratio) was similar to *Tpv* sHSP 14.3 WT. In an archaeal sHSP, SsHsp14.1, residues of the proximal NTD, *i.e.*, isoleucine and leucine at position 9 and 11, participated in hydrophobic interaction with same residues of an adjacent dimer (Liu, J. Y. Chen, *et al.*, 2015). On the other hand, mutations in Phenylalanine 6 and 7 of *S.pombe* resulted in formation of smaller oligomers, while, mutations of phenylalanine 29 and 33 of HSPB6 led to impairment of hetero-oligomerization. This result indicated that those hydrophobic residues are involved in intermolecular interactions beyond dimer (tetramers and higher order oligomers) formation (Heirbaut *et al.*, 2017). Similar functional role can be attributed to the phenylalanine residues in the proximal NTD of the *Tpv* sHSP 14.3.

A highly conserved phenylalanine rich motif in the NTD of HSPBs is predicted to have a helical structure (Heirbaut *et al.*, 2014). This motif plays a role in maintaining the structural and functional integrity of the sHSPs. Deletion of this motif resulted in altered secondary and tertiary structure in the HSPB4 and HSPB5, and the resulting mutants showed decrease in oligomeric size and alteration in chaperone activity (Pasta *et al.*, 2003). This motif, especially its conserved phenylalanine residues are found to be essential for subunit exchange and association of HSPB1/HSPB6 during hetero-oligomerization (Heirbaut *et al.*, 2017). In addition to the hydrophobic residues, the arginine of the motif also effects oligomerization in HSPBs and substituting it with alanine leads to altered oligomers as compared to the respective WTs (Shatov *et al.*, 2018).

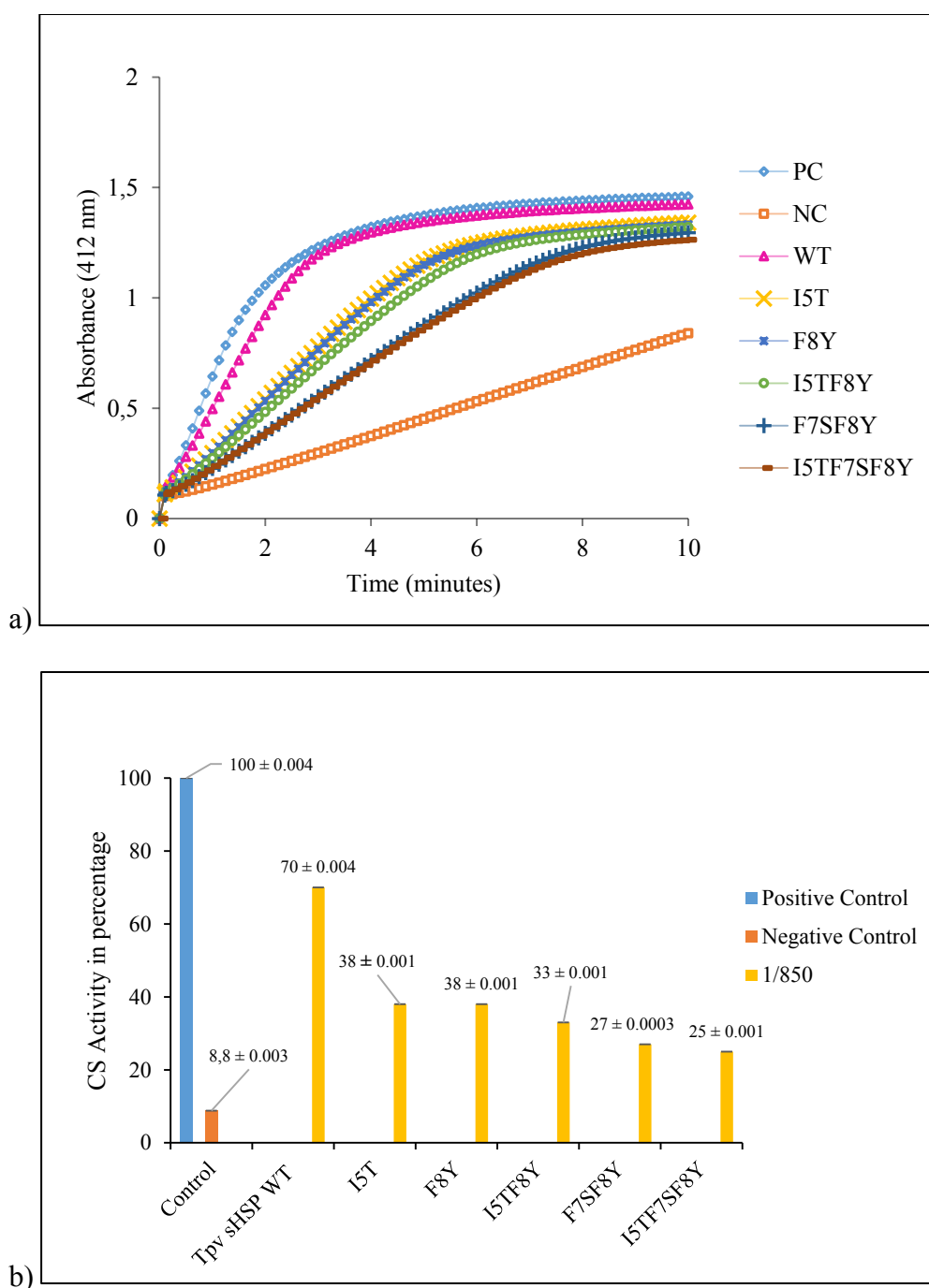


Figure 4.2 Enzyme activity plot (a) and bar chart (b) (representing the slopes) of the percentage protection of citrate synthase (0.44 $\mu\text{g/ml}$) against heat induced aggregation in the presence of 107 $\mu\text{g/ml}$ of *Tpv* sHSP 14.3 WT and various mutants (1:850 CS:chaperone molar ratio), at 47°C. Error bars show the \pm SD of three independent experiments.

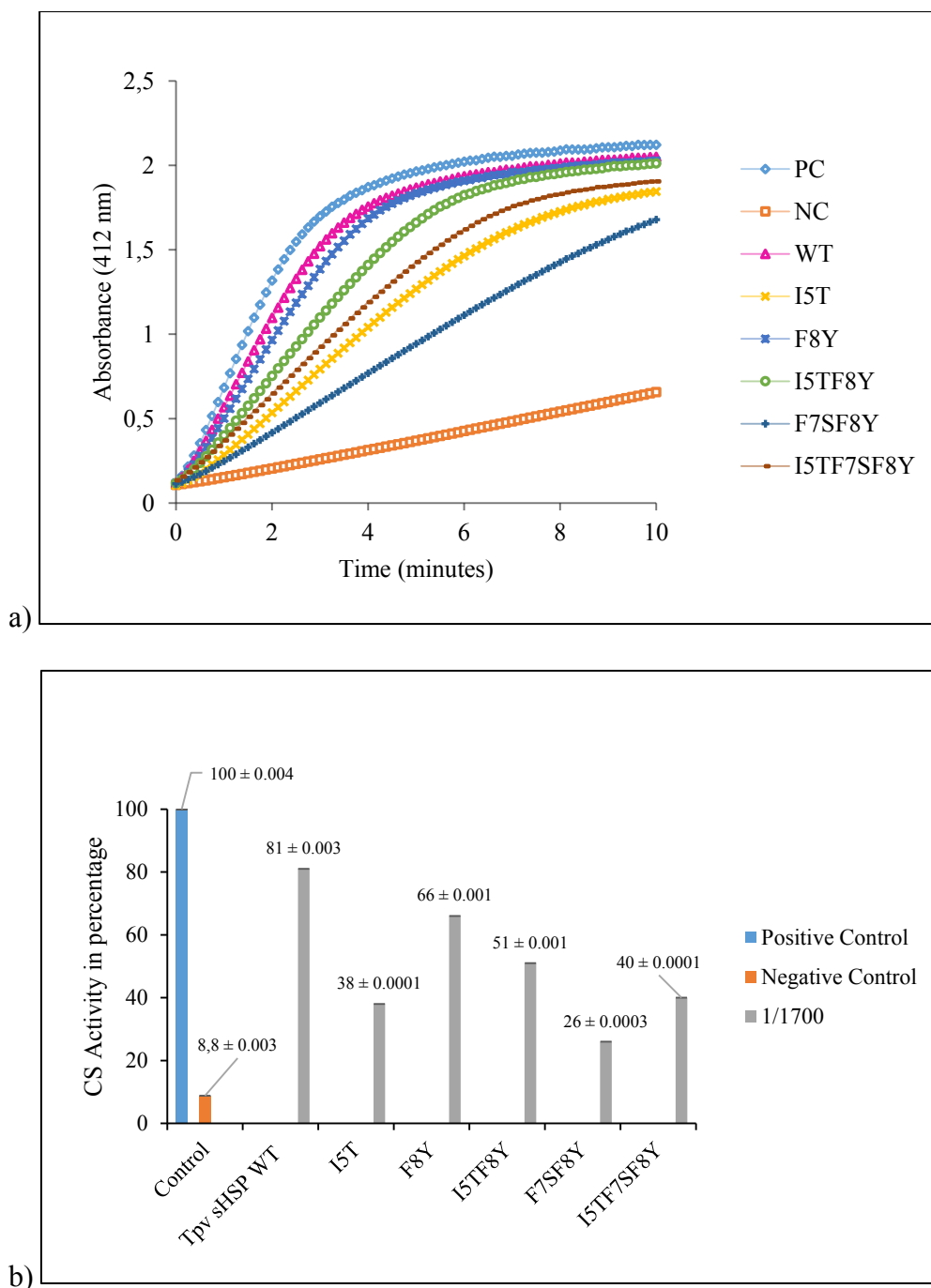


Figure 4.3 Enzyme activity plot (a) and bar chart (b) (representing the slopes) of the percentage protection of citrate synthase (0.44 $\mu\text{g/ml}$) against heat induced aggregation in the presence of 214 $\mu\text{g/mL}$ of *Tpv* sHSP 14.3 WT and various mutants, at 47°C. Error bars show the \pm SD of three independent experiments.

Except the group of the mutants in which the substitution takes place in the initial 10 residues of the NTD (considered above in Group 1), all other mutant variants (Group

2 – 5) displayed comparatively less amount of the protein in the cell extract. One reason could be that, at higher temperature the sHSP tends to protect the partially folded proteins of the host cell, *i.e.*, *E.coli*, but in doing so, it may be irreversibly bound to the substrate and end up with formation of insoluble complexes that could be removed during centrifugation.

4.2 Group 2: Negatively charged residues in the middle domain of NTD (E11V and E22G)

The glutamic acid at 11th position (E11) is highly conserved in all the three domains, namely, archaea, eukaryotes and prokaryotes. However, E22 appears not to be conserved in distinct domains prokaryotes and eukaryotes, but the charge as well as the residue is highly conserved at this position in archaea domain. The glutamic acid at 11th position also shows high conservation, when compared by MSA with sHSPs of eukaryotes studied in literature. The substitutions, E11V and E22G were decided to investigate the importance of negatively charged residues in the NTD of *Tpv* sHSP 14.3.

The expression and preparation of cell lysate of both the mutant variants, involving substitution of charged residue with hydrophobic or neutral amino acid, E11V and E22G, was initially tried in T7-Express competent cells. However, expressed protein level was not sufficient as deduced by SDS-PAGE. Then, expression of these mutant proteins were studied using BL21(DE3) competent cells. Based on the density of the band corresponding to the proteins in SDS-PAGE showed that the expressed protein was not in substantial amount in this system too. On the other hand, the disappearance of the same protein bands in the SDS-PAGE lanes, after heat treatment at 60°C and above indicated the decrease in thermal stability of the mutant protein, as compared to that of WT. For this reason we could not obtain enough protein to perform functional analyses with these mutants.

When the 3-D structure models of the mutants were analyzed for gain or loss of any inter and intra molecular interactions owing to substitution, it was found that the electrostatic interactions that glutamate at position 11 formed with phenylalanine at position 8 in each monomer of the WT, were completely lost in the variant, while the number of hydrogen bonds remained same. It was previously shown that along with hydrophobic interactions, electrostatic interactions are required for the stabilization of not only the sHSP dimers (Wen *et al.*, 2010), but also the substrate-sHSP complex after recognition (Mao, Ke and Chang, 2001; Morris *et al.*, 2008; Nandi *et al.*, 2015). Besides, there is a possibility that these intra molecular electrostatic interaction would also be important for maintaining the structure of the respective NTD in a stable helix. Losing of these interactions might have resulted in a change of the confirmation of NTD in a way as to increase the exposure of hydrophobic surface, thus assisting the strong binding of the subunits in oligomers leading to large tightly bound oligomers or aggregates. This can explain the reason of low protein concentration in the soluble cell-free extract, making it explicit that although the protein would have been expressed in large quantities, it would be insoluble owing to large oligomer and aggregates and thus gets removed during centrifugation process. On the other hand, the glutamate at position 22 does not form any intramolecular electrostatic interactions, with in a monomer or intermolecular ones, within the dimer. However this cannot exclude the possibility of this residue forming intermolecular electrostatic interactions within an oligomer, for its stabilization. Therefore, its loss by mutation may result in unstable protein oligomeric structure. Electrostatic interaction impart significant impact on the structure and function of any sHSP. So much so, in the presence of NaCl, which weakens the electrostatic interactions in proteins, reduction in subunit exchange efficiency, the oligomeric size and its stability and the chaperone function, has been reported for Hsp18 of *M.leprae* (Nandi *et al.*, 2015). In our study, the six hydrogen bonds of E22 with neighboring residues in a monomer, five were retained when this residue was substituted with glycine. There is loss of a single carbon hydrogen bond with S25. There could be two reasons for low quantity of the E22G mutant protein

in the cell lysate: one being similar to E11V, reduced oligomeric stability, the second being that smaller oligomeric structures increases the binding of these sHSPs to other cellular proteins causing precipitation of the large complexes. Both the mutants displayed low chaperone activity when analyzed, to protect CS against thermal aggregation, at two molar ratios 1:7 and 1:35 substrate:sHSP ratio. The protecting ability of E11V was 38% less than that of WT (around 2 folds less than WT), with a single residue mutation E11V and this behavior was concentration independent, with increasing the chaperone five times also resulting in analogous protection effect as lower concentrations. The chaperone activity of the variant E22G was even lower than E11V at lower molar ratio, where the protection was 48% lower (3.3 folds less than WT) than that of the corresponding WT molar concentrations. However, this variant showed concentration dependent activity. There are reports similar to our results of chaperone activity and supporting the speculations about these mutant variants forming larger oligomers. These studies showed that reversing the charge or introducing hydrophobicity at highly conserved glutamic acid at position 41 in HSPB1 of human is a natural mutation leading to various diseases like distal hereditary motor neuropathy (Houlden *et al.*, 2008). They suggested that formation of the larger sizes of oligomers which are unable to dissociate, can lead to decrease in chaperone activity (Shashidharamurthy *et al.*, 2005; Nefedova *et al.*, 2015). The E41 of the HSPB1, corresponds to E11 in our protein, according to MSA result. Also E11 of *Tpv* sHSP 14.3 corresponds to aspartic acid in highly conserved motifs, SRLFDQFFG, SXVFD and WD/EPF, in NTD of human sHSPs, TaHsp16.9 and HSPB1 of eukaryotes respectively. It was suggested that this negatively charged aspartate residue has role in stabilization of oligomers by forming inter and intra molecular bonds with hydrophobic and charged residues of the ACD. Disruption or deletion of D/E or a set of residues including this negatively charged amino acid leads to dissociation of dimers, and therefore, its alteration leads to decrease in chaperone effects, by effecting the equilibrium of dissociation of oligomers when under stress (Thériault *et al.*, 2004; Jaya, Garcia and Vierling, 2009). Depending on these results, we can expect that, it is possible that the highly conserved E11 and E22

in the NTD of *Tpv* sHSP14.3 might be involved in similar inter and/or intra molecular interactions with one or more of the arginines in the ACD (similar to E in the WEPF motif), that may stabilize the oligomers.

The Rosetta Energy results of both the mutants further supports our speculations that these mutations increased the aggregation propensity of the protein leading to the removal of most of it by centrifugation. The substitution E11V shifts the equilibrium towards the fibril formation, *i.e.*, more than half (12 out of 21) of the total energy bands showed the propensity to form fibrils, a clear indication of the resultant protein prone to form large aggregates. On the other hand, E22G, although shows only 8 energy bands (as compared to 9 in WT), passing the threshold of -23Kcal, the energy of two bands present towards the center of NTD increased where increase in one is quite significant (23.6 → 27.2). This indicates that this region which was already rich in fibril forming peptide sequence now shows further increase in it. Since the change in Rosetta Energy is drastic, this could explain the low chaperone effect of the variant at low and high concentrations.

According to the analysis of the 3-D model structures, the two mutations, where negatively charged amino acids are replaced with hydrophobic amino acids, have contrary effects on the NTD of the respective sHSP mutant proteins. The E11V mutation tends to pull the NTD towards the dimer interface while E22G mutation bends the NTD away from the dimer interface. However, in both the mutant protein, this bending of the NTD occurs at residue Thr19. The loss of intramolecular bonds (electrostatic bond in E11V and hydrogen bond in E22G) might be responsible for this bend in the NTD of the mutant proteins. Moreover, the presence of glycine (no methyl group) may have imparted the freedom of rotation to the NTD (Foley and Parra, 2004). In both the mutant proteins, the surface hydrophilicity of the dimer is decreased, upon mutation. Otherwise, the side chain of Glu11 and Glu22 are exposed to the surface, attracting water molecules, and thus might play a role in the solubility of the dimer, in WT. According to the PONDR and PrDOS analysis, the difference between WT and mutant proteins regarding the length of the disordered region in the proximal NTD is not significant. However, slight changes in the score, with respect

to the WT, are found, which is also depicted in the superimposed structures. Previously it has been reported that abolishing a negative charge by replacing it with a hydrophobic residue ($D \rightarrow G$ or $D \rightarrow V$), in the disordered regions, decreased the PONDR score of the residue by 0.07 and 0.11, respectively and consequently had deleterious effect on the function of the protein (superoxide dismutase) leading to disease (sclerosis) (Vacic and Iakoucheva, 2012). Both the residues, E11 and E22 in *Tpv* sHSP 14.3 lie in the predicted ordered region, still replacing them with Val and Gly, respectively produced a difference of 0.05 and 0.00015 in PONDR score, respectively. The thermodynamic stability of the mutant E22G is less than E11V, which could be the reason of its low chaperone activity at 1:7 substrate:chaperone ratio for protecting CS against heat induced aggregation.

4.3 Group 3: Hydrophobic residue in the middle domain of NTD (M12T)

Information about the structure and function of the unstructured NTD of the sHSPs, in general, is not sufficient, as compared to the ACD, due to certain structural limitations. Nevertheless, most of the studies of NTD present, involve the hydrophobic residues, especially phenylalanine, or positively charged residues like arginine. However, in far related (to the archaea and eukaryotes) species, plants, methionine in the NTD of their sHSP is also extensively explored for its role in the structure as well as the chaperone activity of the respective protein. sHSPs from the chloroplasts of angiosperms show a conserved methionine rich region (Waters and Vierling, 1999) which imparts characteristic structural features to the respective sHSPs, for example, the rotation of the two hexameric discs in the dodecamer of HSP21 of *A.thaliana* (Lambert *et al.*, 2011). Since, it has been established that the presence of Hsp16.3 is inevitable for the replication of *Mycobacterium tuberculosis* in the host macrophages, it is speculated that methionine on the sHSPs play a role in neutralizing the reactive oxygen species by reversible oxidation (Nakamoto and Vigh, 2007; Sudnitsyna, Sluchanko and Gusev, 2015). A total of four methionine are present in the primary structure of *Tpv* sHSP 14.3, and two of them lie in the NTD;

one which initiates the protein translation and the second one is present at position 12. The hydrophobicity at position 12 was decreased by mutating it to a polar, threonine. This substitution, M12T, resulted in decrease in the *Tpv* sHSP 14.3 protein concentration in cell lysates of both T7 Express and BL21(DE3) competent cells. There could be two reasons for this: either the resulting protein was toxic for the cell. Else, the mutated protein was more prone to aggregation and removed by centrifugation. The latter is supported by the results of Rosetta Energy, where this single mutation results in shifting the equilibrium towards increased propensity to form fibrils. As compared to 9 energy bands with Rosetta Energy value > -23 Kcal in WT, in the mutant 11 out of the total 21 energy bands in NTD of *Tpv* sHSP 14.3, that crossed the threshold of -23 Kcal. This may lead to disturbance of the balance within the protein that is essential for its solubilization. Therefore, the M12T mutant variant is expected to form β -amyloid like structures and aggregate. Our speculation is supported by a previous study where oxidation of methionine in apolipoprotein (hampering methionine function) hindered the protein's natural process of oligomerization and instead the protein became prone to aggregate by forming cross- β -amyloid like structures (Wong *et al.*, 2010). In another study the sulfoxidation of the conserved methionine resulted in formation of large oligomers, which was explained to be due to exposure of hydrophobic surfaces of sHSPs, which then interacted among themselves leading to increase in molecular mass of oligomers. (Härndahl *et al.*, 2001). Our mutant M12T also was unstable at temperature $\geq 60^{\circ}\text{C}$. The thermodynamic stability calculated by MuPro, of the mutant also shows that the ΔG of the mutant is higher than WT, thus the mutant is thermodynamically unstable. Similarly, methionine oxidation in apolipoprotein resulted in decrease in its thermal stability by 10°C (Wong *et al.*, 2010). Also, the M68T mutant of CRYAB was reported to exhibit decreased thermal stability and aggregated when subjected to higher temperature, 62°C for 15 minutes. On contrary, increasing the hydrophobicity at same position led to increase in thermal stability of the same mutant, as compared to the WT (Shroff *et al.*, 2001).

In addition to this, the mutant's, M12T, ability to protect CS from aggregation at higher temperatures, decreased (by 27%) as compared to WT, at lower chaperone concentration (1:7 substrate:sHSP ratio). Although the M68 lies in the initial ACD of the sHSP CRYAB, still, its replacement by T provided the results in parallel to ours. The M68T mutant of CRYAB was also reported to have decreased chaperone activity, where the extent of decrease was different for different substrates. On the contrary, increasing hydrophobic nature at this position by replacing it with leucine or valine increased the chaperone effect (Shroff *et al.*, 2001). However, when the concentration of our mutant protein, M12T, was increased to five times, (1:35 substrate:sHSP ratio), interestingly the light scattering increased more than the control, indicating that the mutant protein itself might be aggregated along with substrate protein, CS, although the chaperone itself does not show any self aggregation at experiment temperature. Similar results were reported previously where sulfoxidation of the conserved methionine in the NTD of Hsp21 of *A.thaliana* resulted in loss of chaperone activity, where the oxidized Hsp21 not only failed to protect the CS against heat induced aggregation, it itself also co-precipitated with it, as deduced by increase in light scattering (Härndahl *et al.*, 2001). Co-precipitation of the chaperone along with the partially unfolded client has also been reported for native sHSPs, for example, Hsp22, a mitochondrial sHSP of Pea (Avelange-Macherel *et al.*, 2020) Conversely, in another study, where the highly conserved methionine residues were substituted by leucine residues, the chaperone activity against thermal aggregation of CS and chemical aggregation of insulin, of *A.thaliana* Hsp21 remained unaffected (Gustavsson *et al.*, 2001). In addition, it was reported that the oxidation of methionine induces structural changes in the oligomer of Hsp21.

The decrease in the M12T mutant thermal stability can be attributed to the loss of intramolecular hydrophobic interactions. In the WT dimer, this residue (Met12) forms intramolecular hydrophobic interaction with phenylalanine at position 8, one in each monomer. However, the number of intramolecular H-bonds remains unchanged by this mutation. As previously reported, hydrophobic interactions, in addition to H-bonds, play an important role in enabling the protein to maintain its 3-

D structure in native state, at elevated temperatures and are critical in maintaining the stability of thermophilic proteins (Scheraga, Nemethy and Steinberg, 1962; Carver and Lindner, 1998; Moon *et al.*, 2019).

Substitution mutation, M12T also results in alteration of the 3-D structure of the *Tpv* sHSP 14.3 monomer/dimer. The surface hydrophobicity of the dimer is decreased at the point of mutation. According to PONDR score of M12T, only Met1 is found to be disordered. However, the 3-D structure generated by homology modelling predicts the first seven residues to be present in a coil like structure, as compared to only five residues in the WT. The results from PrDos have predicted an increase in score of the first 18 residues of NTD, pushing 14 of them above the threshold line and in to disordered state (while the WT has first two residues in disordered state). This result is supported by the 3-D model structure in a way that a bent, starting from the residue Asn18, in NTD of the M12T mutant is noticed (away from the dimer interface). These structural changes might have been the reason of decrease in thermodynamic and thermal stability of the protein as well as its decreased chaperone activity.

As a result, we can expect that methionine at position 12 in *Tpv* sHSP 14.3 might be involved in substrate recognition and/or binding, and also might play a role in maintaining the protein structural and thermodynamic stability.

4.4 Group 4: Mutations in Distal domain of NTD (V23GF26YI27T, F26YI27T, V31G and V31IL33I)

The segment of NTD lying adjacent to the ACD is reported to play a crucial role in the secondary and tertiary structure of the sHSP (Jehle *et al.*, 2010, 2011). However, it does not affect the oligomerization and the chaperone activity of the sHSP. Mutations, in HSPB5, involving either changing the charge or introduction of hydrophobicity at the positions R56 and/or S59, resulted in changes in the secondary and tertiary structures of the individual mutant sHSP. However, the oligomeric size

of all the mutants remained same. Moreover, the chaperone activity of the mutants also remained comparable to the respective WT (Jovcevski *et al.*, 2018). Similar results have been achieved in this study where a set of four mutants were designed, involving substitution mutation in the distal part of NTD. Three of them were aimed at decreasing the hydrophobicity of the segment, namely V23GF26YI27T, F26YI27T and V31G, while one was planned to increase the hydrophobic nature around the junction of NTD and ACD, *i.e.*, V31IL33I. Of these four, the related mutants (V23GF26YI27T/F26YI27T and V31G/V31IL33I) displayed similar protection against temperature induced aggregation of CS. The two, V31G and V31IL33I, protected the CS against thermal aggregation nearly as efficiently as the WT, *Tpv* sHSP 14.3, at 1:35 substrate:sHSP molar ratios (13% less than WT). At lower chaperone concentrations, 1:7 substrate:sHSP molar ratios, the protective effect of the mutants slightly increased than that of WT. Similar result is seen previously where substitution of Val67, a position just at the junction of NTD and ACD of HSPB6 in human, to glycine (V67G), results in chaperone activity similar to the corresponding WT. The mutant V67G is also reported to have similar oligomeric profile, like its WT (Weeks *et al.*, 2014). On the other hand, the mutants, V23GF26YI27T and F26YI27T, display similar chaperone activity at both 1:7 and 1:35 molar ratios. However, their chaperone activity is around 25% higher than WT at 1:7 and around 10% higher than that of WT at 1:35 substrate:sHSP molar ratios. The 3-D structure predicted by homology modelling shows increase in the length of disordered region for V23GF26YI27T and a bend in the NTD of V31G mutant. The prediction results from PrDOS points towards increase in the disordered region of all the four mutants in this group. Considering the fact that disordered region plays a role in interaction with the substrate (Dunker *et al.*, 2001; Oldfield *et al.*, 2019), there may be a possibility that this would be the reason of slight increase in activity of above mentioned mutants as compared to WT at 1:7 CS:chaperone molar ratio. Although such an increase in the length of disordered region at the proximal NTD is also noticed in group1 mutants that is accompanied with decrease in the chaperone

activity, the reason for the later could be that the mutation in the region itself may have diminished any benefit effect, of the increasing disordered region.

It has been previously reported for the sHSPs of *B.japonicum*, a portion of NTD, lying close to ACD, is responsible in dimer formation and stabilization, in addition to the ACD itself, while complete intact NTD is inevitable for oligomer formation, and chaperone activity (Sun and MacRae, 2005). Another study shows that, the NTD segment adjacent to the ACD is involved in tetramer formation (Weeks *et al.*, 2014). Similar structural arrangements are witnessed in our *Tpv* sHSP 14.3 WT structure, where the first inter-monomer interaction (hydrophobic) occurs between isoleucine27 of both monomers. Residues 1 – 26 of the NTD do not show any monomer-monomer interactions. The loss of this interaction might result in loosening of the dimer at this end (Figure 4.4). This hydrophobic interaction is lost when isoleucine is substituted with a hydrophilic, threonine, in both the mutants, V23GF26YI27T and F26YI27T. However, the former/triple mutant might compensate this, by forming a hydrogen bond between the two threonine at position 27, but the double mutant, F26YI27T fails to do so (Appendix B, Table 10). However, the length possessed by hydrophobic interaction is not the same as H-bond. Also, since hydrophobic interactions are stronger than H-bonding, the monomer in the mutant dimer (V23GF26YI27T) might be loosely bound at this position and therefore, might explain the reason of V23GF26YI27T mutant having similar chaperone activity to F26YI27T mutant, which does not possess any interaction between residue 27 of each monomer in a dimer. Nevertheless, F26YI27T makes a new interaction (hydrophobic) with valine 23 of the same monomer. This interaction is absent in the triple mutant, as valine is substituted at position 23 by glycine. A new intramolecular hydrophobic interaction might result in shift of the last turn of the helix further towards the other helices.

The mutant V31G possessed similar intermolecular H-bonds like the WT (between V31 and glutamine 80), while in V31IL33I, in addition to retaining the intermolecular H-bonds of WT, four more intermolecular H-bonds were formed. Retention of intermolecular H-bonds by V31G and V31IL33I similar to WT,

explains the similar chaperone activity of these two mutants to the WT. Our speculations are also supported by the results obtained from Zipper DB software. The Rosetta Energy changes in the surrounding hexa peptides of V31G and V31IL33I variant proteins are not drastic. Although the Rosetta Energy values are increased, no single peptide energy value crosses the threshold of -23Kcal and the overall result of the NTD remains same before and after the mutation. However, the increase in protective effect of F26YI27T and V23GF26YI27T can be explained by loosening of the dimer. This could lead to easy subunit exchange, an increase in hydrophobicity available to the substrate and thus chaperone activity.

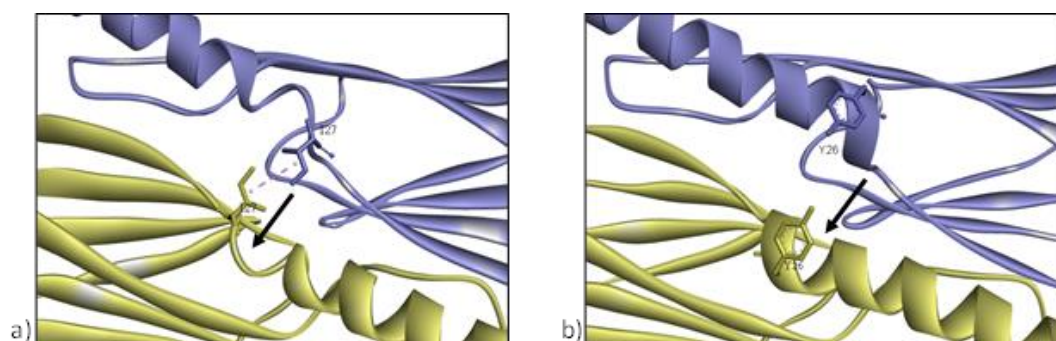


Figure 4.4 3-D structure of WT (a) and F26YI27T (b). Black arrow showing widening of the gap between the helix

On the other hand, both the related mutants, F26YI27T and V23GF26YI27T cause an overall decrease in the number of hexa-peptides passing the threshold energy of -23 Kcal/mol. Although an individual increase, to the same extent, in Rosetta Energy of peptide ${}_{23}\text{VSSFIY}_{28}$ is seen in both the mutants, the number of peptides that pass this threshold of -23 Kcal/mol is decreased to 8 and 5, in the double and triple mutant, respectively, as compare to 9 in the NTD of the WT. The ability of a protein to form fibril like structure is function of the overall balance of all the hexa peptides in a protein. The decrease in propensity to form fibrils in the NTD, might further enhance the fluidity/solubility of the protein, and might result in an increase ability of the protein to protect aggregating substrates at sub-optimal temperature. In literature, similar results are found where mutations increasing hydrophobicity, M68V and M68I, in CRYAB of mammals, also showed increase in protective effect against

chemically and heat induced denaturation of insulin, ADH and β _L-crystallin (Shroff *et al.*, 2001). Therefore, it is quite expected that an increase in hydrophobicity for V31IL33I mutant, increases its chaperone effect, even if it is slight, as compared to WT.

As reported previously the mutations at the far end of NTD resulted in decreased thermal stability of the proteins (Jovcevski *et al.*, 2018). Similar result was observed in our study, where all the four mutants grouped here showed decreased stability at temperatures 70°C and above, but were able to withstand the temperature up till 65°C (V31G and V31IL33I). All the four mutants in this group exhibited subtle amount of protein in the cell lysate when expressed using BL21(DE3) competent cell expression system. However, similar to the WT, V31G also did not show any appreciable expression in the T7-Express competent cells.

Studies have shown that when phenylalanine, at distal end of NTD, is substituted with a less hydrophobic residue, the protective effect is generally similar or comparable to that of WT. For example alanine was the new amino acid in place of phenylalanine at position 17 and 20 in *S.pombe* sHSP, and they had chaperone effect similar to their WT (Hanazono *et al.*, 2013b). In a study, substituting phenylalanine around the middle of NTD, F16 and F19 of PsHsp18.1, with Bpa, led to similar chaperone activity as the WT. However, the mutation at position 41 and 48 in resulted in decrease in chaperone activity towards certain substrates (Jaya, Garcia and Vierling, 2009). Therefore, it can be assessed that, not all phenylalanines in the NTD of a sHSP are involved in substrate recognition. Some might play a role in stabilizing of the dimers and oligomers prior to and following substrate-sHSP complex formation.

4.5 Group 5 F8YV31G

It has seen previously, in group 1, that as compared to single mutation, when two mutations are combined in a single protein, the resultant protein could have the

additional effect of both the mutations. An example of such combination in this study is the variant containing double substitution, where both the substitutions are far apart, F8YV31G. The mutation protein band was completely lost when subjected to 70°C, although the protein was stable at 65°C and below. This behavior, at high temperature, was in contrast to that of the single mutant F8Y, but similar to the single mutant V31G, which also is lost possibly due to aggregation when the cell extract is exposed to 70°C. This may imply that the valine at position 31 (a position close to ACD) may be involved in imparting stability to the sHSP at higher temperatures. When we consider the inter molecular and intramolecular interactions of amino acids at position 8 and 31 in both the WT and double mutant, the mutant loses the intramolecular hydrophobic interaction at position 8. Also, similar to single mutant V31G, the double mutant F8YV31G loses its intramolecular hydrophobic interactions with alanine at position 44 but possess the intermolecular hydrogen bonds with the same residue as in WT, *i.e.*, Glutamine 80 (although the atoms involved are different). On the other hand, the single mutation F8Y results in intramolecular hydrophobic and electrostatic interactions with new residues, but introduction of another distal mutation V31G along with F8Y, results in loss of hydrophobic interaction at position 8 (in the mutant) while the electrostatic interactions are retained, similar to WT. Also, the intramolecular hydrogen bonds, at position 8, in the double mutant F8YV31G are exactly similar to WT, whereas in single mutant F8Y, one of the two hydrogen bonds in WT, is replaced with a new residue. Studies have shown that the NTD of the sHSPs are involved in forming intermolecular interactions which is responsible for formation and stabilization of oligomers and, therefore, play a role in subunit exchange dynamics too (Delbecq, Rosenbaum and Klevit, 2015). The overall summary of interactions is that the double mutant is more similar to V31G in losing intramolecular hydrophobic interactions and retaining all other bonds, while F8Y has managed to form new bonds with new residues after losing the intramolecular hydrophobic, electrostatic and H-bonds.

When we consider the chaperone activity of this mutant, it is highly effective in protecting citrate synthase from heat aggregation, when exposed to the sub-optimal

temperature. This effect of F8YV31G is similar to that of WT (at 1:35 CS:sHSP molar ratio), which is less than both F8Y (by 11%) and V31G (by 14%). However, at lower molar ratios, (at 1:7 CS:sHSP molar ratio), the effect of double mutant F8YV31G is similar to V31G but more than F8Y at similar ratios. This may be the result of disturbance of the oligomer stability by loss of hydrophobic interactions. Thereby, small oligomeric and dimeric forms with high binding affinities may dominate. Similarly, the double substitution mutant, F8YV31G, showed protection of the activity of CS, comparable to that of WT, when present in excess. At 1:1700 CS:sHSP molar ratio, the activity of CS is protected by F8YV31G as effectively as WT. However, the single mutant F8Y, even at such high concentrations (1:1700 CS:sHSP molar ratio) had 40% less protection than the double mutant F8YV31G. This suggests that the distant mutation of V31G, helps to overcome the slight adverse effects of the decreasing of hydrophobicity in the proximal part of NTD. Interestingly, a similar result was reported in a previous study, where the mutation in the proximal (HSPB6-GPG) and distal (HSPB6V67G) NTD, separately, exhibited increased and decreased chaperone activity as compared to the WT. However, when both mutations were combined in a single protein (GPG-V67G), the effect was similar to WT, which means, less than the former mutant and more than the latter mutant (Weeks *et al.*, 2014).

The oligomer distribution on BN-PAGE shows that, F8YV31G mutant, in addition to possessing the 24-mer oligomer, like WT and F8Y, possessed oligomer of 60-mer, similar to WT but smaller than that of F8Y (76-mer). Another striking difference is the presence of a dodecamer in F8YV31G mutant, which is also present in WT, but absent in F8Y mutant. This profile of oligomers in F8YV31G mutant is same as WT but different than F8Y. A possible reason could be the difference in the intramolecular interactions (which are similar to WT and V31G mutant but different from F8Y mutant). This oligomer size distribution of the double mutant can explain the similarity of its chaperone activity to that of the WT but different than the F8Y mutant protein. The Rosetta Energy output of F8YV31G is a combination of both F8Y and V31G single mutants and also is similar to WT. Moreover, the

superimposition of WT and F8YV31G mutant 3-D models reveals the structural similarity between mutant protein and WT (Figure 3.14). Although the PrDOS results predict an increase in the length of the disordered region at the proximal NTD, the results from PONDR, like homology modelling predictions, suggest striking similarities in the overall structure of WT and F8YV31G. This might be accounting for the similar chaperone behavior of F8YV31G and WT sHSPs.

4.6 Mini-peptides

To date, a number of mini chaperones have been synthesized and exhibit great potential as therapeutic agents (Phadte, Sluzala and Fort, 2021). Most of the mini chaperones studied, are rich in charged residues and have comparatively fewer hydrophobic residues (Sreekumar *et al.*, 2013; Raju *et al.*, 2014). However, there are limited reports on identification of sequences as mini chaperones beyond the borders of ACD. Our study is the first time where the NTD alone of an archaeal sHSP is being explored for its ability to protect a substrate from aggregating when exposed to sub optimal temperatures. There are reports only about HSPB1 of human, where its whole NTD or parts of it showed protection of several substrates against thermal aggregation (Gliniewicz *et al.*, 2019; Kho *et al.*, 2021). However, the protection varied for different substrates and whole NTD showed minimal protection for CS substrate. For example, protection of a substrate (MDH) against aggregation was better than others, and almost accounted to complete protection (Gliniewicz *et al.*, 2019). However, high protection of CS by comparatively shorter peptides (4 – 10 amino acids), designed from the NTD of the HSPB1, which was named as macro-chaperone, was shown (Kho *et al.*, 2021). These studies, combined with our results suggest that NTD of sHSPs, including *Tpv* sHSP 14.3 can perform effective chaperone activity by interacting with the partially folded substrates.

Our results also showed a positive correlation between the peptide concentration and the percentage of protection (Figure 4.5 and 4.6). Moreover, both peptides were able to protect the substrate above 90% when the substrate:chaperone weight ratio was

4:1 or higher. In the presence of 65 $\mu\text{g/mL}$ of the peptide2, the protection of 233 $\mu\text{g/mL}$ of yADH was almost complete.

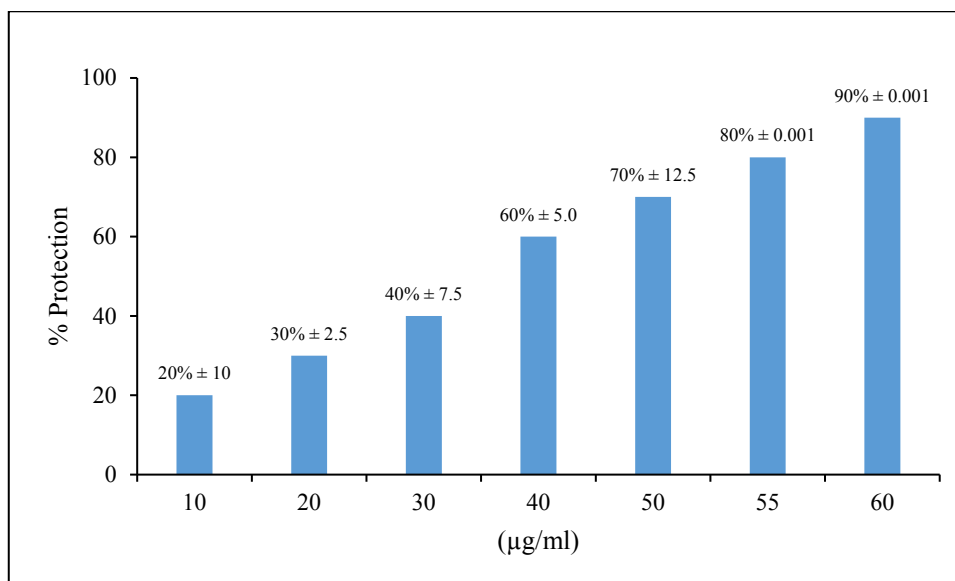


Figure 4.5 Bar chart depicting the percentage protection of yADH in the presence of peptide1 at different concentrations at 43°C. Percent protection is relative to the protection of the control, *i.e.*, in the absence of mini peptide (0%). Each data point represents the mean of three independent trials and is shown as the mean \pm standard deviation (SD).

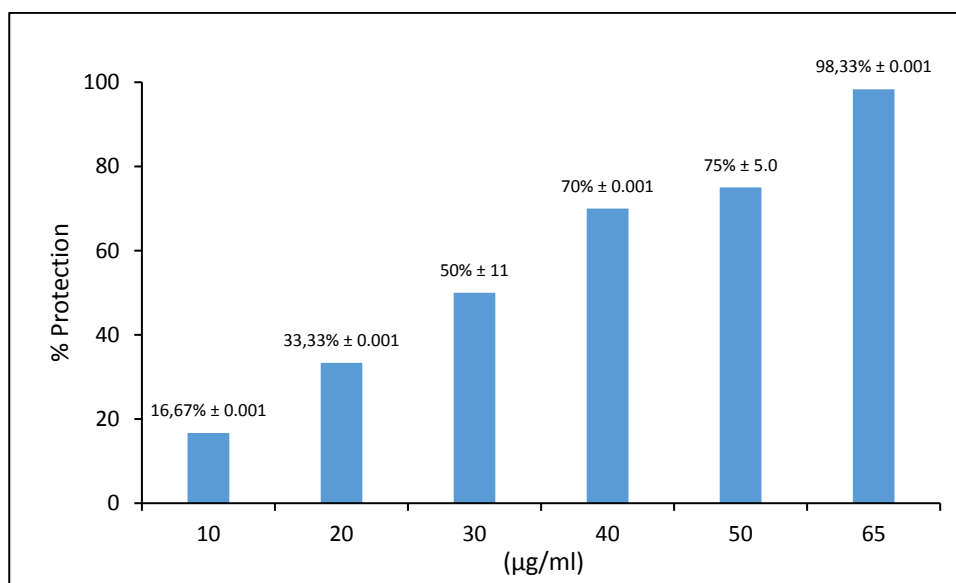


Figure 4.6 Bar chart depicting the percentage protection of yADH in the presence of peptide2 at different concentrations at 43°C. Percent protection is relative to the protection of the control, *i.e.*, in the absence of mini peptide (0%). Each data point represents the mean of three independent trials and is shown as the mean \pm standard deviation (SD).

CHAPTER 5

CONCLUSION

1. The NTD of sHSP of *Thermoplasma volcanium* is rich in hydrophobic residues at the proximal and distal end, while in the middle part, it has a higher ratio of residues with positive or negative charge. The monomer model 3-D structure we generated using homology modelling predicts the NTD to be mostly composed of a helical structure with disordered coils in the proximal and distal part. The last two residues of the NTD forms the beginning portion of β 2 sheet which is followed by the ACD section. In the 3-D model structure of the dimer, the two NTD face away from each other. The two monomers are stacked in a way that the β 6 of one monomer lies upon and interacts with the β 2 of the other monomer. In a dimer, only interaction between monomers, where NTD is involved in, is a hydrophobic interaction between Ile27 of both monomers.
2. The *Tpv* sHSP 14.3 WT is an efficient chaperone. When present in high concentrations, it does not only protect nearly 80% (as compared to control) of the substrate against heat induced aggregation, but also helps the substrate to retain its activity after removal of heat stress. The chaperone is stable at temperatures as high as 80°C and exists as the characteristic 24-mer oligomer, a trademark of most of the archaeal sHSPs, in addition to dimer dodecamer and 60-mer.
3. The primary sequence of NTD of the *Tpv* sHSP 14.3 WT exists in an equilibrium where its middle portion is more prone to form fibril like structures while the two ends are less likely. This equilibrium maintains the solubility of not only the NTD, but of the whole protein as well. The homology modelling predicts residues 1 – 5 of NTD being intrinsically disordered. However, the online tools PONDR (residue 1) and PrDOS (residues 1 – 2 and 20 – 21) have slightly different predictions.

4. Except the mutations, I5TF8Y and F8YV31G, all the other mutations in the NTD, resulted in alteration in the 3-D spatial arrangement of CTD.
5. The single mutations, E11V and E22G resulted in prominent increase in the surface hydrophobicity of the dimer. Besides, V31IL33I also resulted in slight increase in surface hydrophobicity. However, all other mutants slightly (single mutations) or drastically (double and triple mutations) decreased the surface hydrophobicity.
6. In addition to the surface hydrophobicity, the mutations also altered the 3-D arrangement of NTD in a dimer. Among the set of mutations performed in the proximal NTD, I5T and I5TF7SF8Y cause the bending of NTD away from the dimer axis, from residues Val16 and Met12, respectively. Other mutants in this group show spatial arrangement of NTD similar to that of WT. All mutants, besides E11V, E22G, I5T and F8Y, V31G and F8YV31G, increased the length of the disordered coil present in the proximal NTD. Moreover, NTD of the two monomers in a dimer, tend to bend away from the dimer axis in I5T, E22G, V31G and V23GF26YI27T and towards the dimer axis in E11V mutant protein.
7. The decrease in the hydrophobicity at the proximal region of NTD results in decrease in the chaperone activity of *Tpv* sHSP 14.3. While decreased hydrophobicity, at the distal end of NTD results in increase in the protection effect of the *Tpv* sHSP against heat induced aggregation of substrate protein. However, introduction of hydrophilicity (replacement of a highly conserved hydrophobic residue by a hydrophilic one) in the middle of NTD enhanced the aggregation of the substrate protein. Moreover, when the highly conserved negative charges are abolished in the NTD by introducing hydrophobic amino acids at that positions, chaperone activity decreased 3-fold as compare to the WT.
8. None of the mutants of proximal NTD of *Tpv* sHSP 14.3 was able to protect the activity of substrate enzyme, after exposure to suboptimal temperature, as efficiently as the WT. However at high concentrations, the protection effect of F8YV31G mutant protein approached to that of WT.

9. Decreased hydrophobicity in the proximal region of the NTD results in alteration in the oligomeric distribution of the resultant mutants, indicating a possible role for hydrophobic residues in the proximal NTD, for formation and stabilization of oligomers. The characteristic 24-mer was possessed by all mutants except I5T and I5TF7SF8Y. All mutants showed presence of large species (52-mer, 60-mer or 76-mer) except F7SF8Y. The active subunit, hypothesized to interact with the substrate, the dimer, was found in each of the mutants, similar to the WT. The mutant possessing the oligomer distribution most similar to WT was F8YV31G, a combination of proximal and distal mutation in a single protein. Heat stress altered the oligomeric distribution of the chaperones by increasing the band density of some oligomer size and appearance of some new size oligomers, in addition to the previous ones.
10. The mutations performed in this study, also altered the propensity of the NTD towards fibril formation. In addition to M12T, the two mutations where negative charges were abolished (E11V and E22G) resulted in the maximum increase in fibril forming tendency of the NTD. The mutations at the distal NTD resulted in an overall decrease in Rosetta Energy values of the surrounding peptides. Similar results were obtained for the proximal NTD mutations, which led to decrease in overall fibril formation tendency of NTD. An exception to this was the two mutants F7SF8Y and I5TF7SF8Y. These two mutant variants resulted in increasing the inclination of NTD towards fibril formation.
11. The NTD of *Tpv* sHSP 14.3, as a synthetic peptide, alone also possessed chaperone activity and was able to protect substrate against heat induced aggregation. Moreover, a segment of NTD, the first twelve amino acids, as a short peptide, also possessed equivalent heat protection effect of the substrate, as compared to the whole NTD peptide.

REFERENCES

- Abeln, S. and Frenkel, D. (2008) 'Disordered flanks prevent peptide aggregation', *PLoS Computational Biology*, 4(12). doi: 10.1371/journal.pcbi.1000241.
- Alderson, T. R. *et al.* (2020) 'Conditional Disorder in Small Heat-shock Proteins', *Journal of Molecular Biology*. Elsevier Ltd, 432(9), pp. 3033–3049. doi: 10.1016/j.jmb.2020.02.003.
- Altamirano, M. M. *et al.* (1997) 'Refolding chromatography with immobilized mini-chaperones', *Proceedings of the National Academy of Sciences of the United States of America*, 94(8), pp. 3576–3578. doi: 10.1073/pnas.94.8.3576.
- Avelange-Macherel, M. H. *et al.* (2020) 'The mitochondrial small heat shock protein HSP22 from pea is a thermosoluble chaperone prone to co-precipitate with unfolding client proteins', *International Journal of Molecular Sciences*, 21(1), pp. 1–23. doi: 10.3390/ijms21010097.
- Avilov, S., Aleksandrova, N. and Demchenko, A. (2005) 'Quaternary Structure Of α -Crystallin Is Necessary For The Binding Of Unfolded Proteins: A Surface Plasmon Resonance Study', *Protein & Peptide Letters*, 11(1), pp. 41–48. doi: 10.2174/0929866043478437.
- Ayyadevara, S. *et al.* (2021) 'Intrinsically disordered proteins identified in the aggregate proteome serve as biomarkers of neurodegeneration', *Metabolic Brain Disease*. Springer US, (0123456789). doi: 10.1007/s11011-021-00791-8.
- Bakthisaran, R., Tangirala, R. and Rao, C. M. (2015) 'Small heat shock proteins: Role in cellular functions and pathology', *Biochimica et Biophysica Acta - Proteins and Proteomics*. Elsevier B.V., 1854(4), pp. 291–319. doi: 10.1016/j.bbapap.2014.12.019.
- Bardwell, J. C. A. and Jakob, U. (2012) 'Conditional disorder in chaperone action', *Trends in Biochemical Sciences*. Elsevier Ltd, 37(12), pp. 517–525. doi:

10.1016/j.tibs.2012.08.006.

Basha, E., O'Neill, H. and Vierling, E. (2012) 'Small heat shock proteins and α -crystallins: Dynamic proteins with flexible functions', *Trends in Biochemical Sciences*. Elsevier Ltd, 37(3), pp. 106–117. doi: 10.1016/j.tibs.2011.11.005.

Baughman, H. E. R. *et al.* (2020) 'Release of a disordered domain enhances HspB1 chaperone activity toward tau', *Proceedings of the National Academy of Sciences of the United States of America*, 117(6), pp. 2923–2929. doi: 10.1073/pnas.1915099117.

Benesch, J. L. P. *et al.* (2008) 'Small heat shock protein activity is regulated by variable oligomeric substructure', *Journal of Biological Chemistry*. © 2008 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 283(42), pp. 28513–28517. doi: 10.1074/jbc.M804729200.

Bepperling, A. *et al.* (2012) 'Alternative bacterial two-component small heat shock protein systems', *Proceedings of the National Academy of Sciences of the United States of America*, 109(50), pp. 20407–20412. doi: 10.1073/pnas.1209565109.

Bhandari, S. *et al.* (2019) 'Dodecameric structure of a small heat shock protein from *Mycobacterium marinum* M', *Proteins: Structure, Function and Bioinformatics*, 87(5), pp. 365–379. doi: 10.1002/prot.25657.

Bhattacharyya, J. *et al.* (2006) 'Mini- α B-crystallin: A functional element of α B-crystallin with chaperone-like activity', *Biochemistry*, 45(9), pp. 3069–3076. doi: 10.1021/bi0518141.

Bhattacharyya, J. and Sharma, K. K. (2001) 'Conformational specificity of mini- α A-crystallin as a molecular chaperone', *Journal of Peptide Research*, 57(5), pp. 428–434. doi: 10.1034/j.1399-3011.2001.00871.x.

Biswas, S. *et al.* (2021) 'Multiple nanocages of a cyanophage small heat shock protein with icosahedral and octahedral symmetries', *Scientific Reports*. Nature

- Publishing Group UK, 11(1), pp. 1–17. doi: 10.1038/s41598-021-00172-2.
- Boelens, W. C. (2020) ‘Structural aspects of the human small heat shock proteins related to their functional activities’, *Cell Stress and Chaperones*. Cell Stress and Chaperones, 25(4), pp. 581–591. doi: 10.1007/s12192-020-01093-1.
- Boncoraglio, A., Minoia, M. and Carra, S. (2012) ‘The family of mammalian small heat shock proteins (HSPBs): Implications in protein deposit diseases and motor neuropathies’, *International Journal of Biochemistry and Cell Biology*. Elsevier Ltd, 44(10), pp. 1657–1669. doi: 10.1016/j.biocel.2012.03.011.
- Bondino, H. G., Valle, E. M. and ten Have, A. (2012) ‘Evolution and functional diversification of the small heat shock protein/ α -crystallin family in higher plants’, *Planta*, 235(6), pp. 1299–1313. doi: 10.1007/s00425-011-1575-9.
- Bova, M. P. *et al.* (1997) ‘Subunit exchange of α A-crystallin’, *Journal of Biological Chemistry*. © 1997 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 272(47), pp. 29511–29517. doi: 10.1074/jbc.272.47.29511.
- Buchner, J. *et al.* (1991) ‘GroE Facilitates Refolding of Citrate Synthase by Suppressing Aggregation’, *Biochemistry*, 30(6), pp. 1586–1591. doi: 10.1021/bi00220a020.
- Capriotti, E., Fariselli, P. and Casadio, R. (2005) ‘I-Mutant2.0: Predicting stability changes upon mutation from the protein sequence or structure’, *Nucleic Acids Research*, 33(SUPPL. 2), pp. 306–310. doi: 10.1093/nar/gki375.
- Carra, S. *et al.* (2017) ‘The growing world of small heat shock proteins: from structure to functions’, *Cell Stress and Chaperones*. Cell Stress and Chaperones, 22(4), pp. 601–611. doi: 10.1007/s12192-017-0787-8.
- Carra, S. *et al.* (2019) ‘Small heat shock proteins: multifaceted proteins with important implications for life’, *Cell Stress and Chaperones*. Cell Stress and Chaperones, pp. 295–308. doi: 10.1007/s12192-019-00979-z.

Carver, J. A. *et al.* (2017) 'The functional roles of the unstructured N- and C-terminal regions in α B-crystallin and other mammalian small heat-shock proteins', *Cell Stress and Chaperones*, 22(4), pp. 627–638. doi: 10.1007/s12192-017-0789-6.

Carver, J. A. and Lindner, R. A. (1998) 'NMR spectroscopy of α -crystallin. Insights into the structure, interactions and chaperone action of small heat-shock proteins', *International Journal of Biological Macromolecules*, 22(3–4), pp. 197–209. doi: 10.1016/S0141-8130(98)00017-8.

Chang, Z. *et al.* (1996) 'Mycobacterium tuberculosis 16-kDa antigen (Hsp16.3) functions as an oligomeric structure in vitro to suppress thermal aggregation', *Journal of Biological Chemistry*. © 1996 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 271(12), pp. 7218–7223. doi: 10.1074/jbc.271.12.7218.

Cheng, J., Randall, A. and Baldi, P. (2006) 'Prediction of protein stability changes for single-site mutations using support vector machines', *Proteins: Structure, Function and Genetics*, 62(4), pp. 1125–1132. doi: 10.1002/prot.20810.

Chernova, L. S. *et al.* (2020) 'N- And C-terminal regions of the small heat shock protein IbpA from: *Acholeplasma laidlawii* competitively govern its oligomerization pattern and chaperone-like activity', *RSC Advances*. Royal Society of Chemistry, 10(14), pp. 8364–8376. doi: 10.1039/c9ra10172a.

Chung, C. T., Niemela, S. L. and Miller, R. H. (1989) 'One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution', *Proceedings of the National Academy of Sciences*, 86(7), pp. 2172 LP – 2175. doi: 10.1073/pnas.86.7.2172.

Clark, A. R. *et al.* (2018) 'Terminal Regions Confer Plasticity to the Tetrameric Assembly of Human HspB2 and HspB3', *Journal of Molecular Biology*. Elsevier

Ltd, 430(18), pp. 3297–3310. doi: 10.1016/j.jmb.2018.06.047.

Collier, M. P. and Benesch, J. L. P. (2020) ‘Small heat-shock proteins and their role in mechanical stress’, *Cell Stress and Chaperones*. *Cell Stress and Chaperones*, 25(4), pp. 601–613. doi: 10.1007/s12192-020-01095-z.

Delbecq, S. P., Rosenbaum, J. C. and Klevit, R. E. (2015) ‘A Mechanism of Subunit Recruitment in Human Small Heat Shock Protein Oligomers’, *Biochemistry*, 54(28), pp. 4276–4284. doi: 10.1021/acs.biochem.5b00490.

Dukay, B. *et al.* (2021) ‘Neuroinflammatory processes are augmented in mice overexpressing human heat-shock protein B1 following ethanol-induced brain injury’, *Journal of Neuroinflammation*, 18(1), pp. 1–24. doi: 10.1186/s12974-020-02070-2.

Dunker, A. K. *et al.* (2001) ‘Intrinsically disordered protein’, *Journal of Molecular Graphics and Modelling*, 19(1), pp. 26–59. doi: 10.1016/S1093-3263(00)00138-8.

Foley, J. and Parra, R. (2004) ‘An ab initio study of intermolecular interactions of glycine, alanine, and valine dipeptide-formaldehyde dimers’, *Journal of Undergraduate Chemistry ...*, (June), pp. 15–20. Available at: http://www.researchgate.net/publication/258209624_An_ab_initio_study_of_inter_molecular_interactions_of_glycine_alanine_and_valine_dipeptide-formaldehyde_dimers/file/e0b495273d9fef1062.pdf.

Freilich, R. *et al.* (2018) ‘Competing protein-protein interactions regulate binding of Hsp27 to its client protein tau’, *Nature Communications*, 9(1), pp. 1–11. doi: 10.1038/s41467-018-07012-4.

Fu, X. *et al.* (2005) ‘A dual role for the N-terminal region of Mycobacterium tuberculosis Hsp16.3 in self-oligomerization and binding denaturing substrate proteins’, *Journal of Biological Chemistry*, 280(8), pp. 6337–6348. doi: 10.1074/jbc.M406319200.

Ghahramani, M. *et al.* (2020) ‘The congenital cataract-causing mutations P20R and

A171T are associated with important changes in the amyloidogenic feature, structure and chaperone-like activity of human α B-crystallin', *Biopolymers*, 111(5), pp. 1–15. doi: 10.1002/bip.23350.

Gliniewicz, E. F. *et al.* (2019) 'Chaperone-like activity of the N-terminal region of a human small heat shock protein and chaperone-functionalized nanoparticles', *Proteins: Structure, Function and Bioinformatics*, 87(5), pp. 401–415. doi: 10.1002/prot.25662.

Gustavsson, N. *et al.* (2001) 'Substitution of conserved methionines by leucines in chloroplast small heat shock protein results in loss of redox-response but retained chaperone-like activity', *Protein Science*, 10(9), pp. 1785–1793. doi: 10.1110/ps.11301.

Haley, D. A., Horwitz, J. and Stewart, P. L. (1998) 'The small heat-shock protein, α B-Crystallin, has a variable quaternary structure', *Journal of Molecular Biology*, 277(1), pp. 27–35. doi: 10.1006/jmbi.1997.1611.

Hanazono, Y. *et al.* (2012) 'Structural studies on the oligomeric transition of a small heat shock protein, StHsp14.0', *Journal of Molecular Biology*. Elsevier Ltd, 422(1), pp. 100–108. doi: 10.1016/j.jmb.2012.05.017.

Hanazono, Y. *et al.* (2013a) 'Nonequivalence observed for the 16-meric structure of a small heat shock protein, SpHsp16.0, from *Schizosaccharomyces pombe*', *Structure*. Elsevier Ltd, 21(2), pp. 220–228. doi: 10.1016/j.str.2012.11.015.

Hanazono, Y. *et al.* (2013b) 'Nonequivalence observed for the 16-meric structure of a small heat shock protein, SpHsp16.0, from *Schizosaccharomyces pombe*', *Structure*. Elsevier Ltd, 21(2), pp. 220–228. doi: 10.1016/j.str.2012.11.015.

Härndahl, U. *et al.* (2001) 'The chaperone-like activity of a small heat shock protein is lost after sulfoxidation of conserved methionines in a surface-exposed amphipathic α -helix', *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, 1545(1–2), pp. 227–237. doi: 10.1016/S0167-4838(00)00280-6.

- Haslbeck, M. *et al.* (1999) ‘Hsp26 : a temperature-regulated chaperone’, 18(23), pp. 6744–6751.
- Haslbeck, M. *et al.* (2004) ‘A domain in the N-terminal part of Hsp26 is essential for chaperone function and oligomerization’, *Journal of Molecular Biology*, 343(2), pp. 445–455. doi: 10.1016/j.jmb.2004.08.048.
- Haslbeck, M. *et al.* (2005) ‘Some like it hot: The structure and function of small heat-shock proteins’, *Nature Structural and Molecular Biology*, 12(10), pp. 842–846. doi: 10.1038/nsmb993.
- Haslbeck, M. *et al.* (2008) ‘Structural Dynamics of Archaeal Small Heat Shock Proteins’, *Journal of Molecular Biology*, 378(2), pp. 362–374. doi: 10.1016/j.jmb.2008.01.095.
- Haslbeck, M. and Vierling, E. (2015) ‘A first line of stress defense: Small heat shock proteins and their function in protein homeostasis’, *Journal of Molecular Biology*. Elsevier Ltd, 427(7), pp. 1537–1548. doi: 10.1016/j.jmb.2015.02.002.
- Haslbeck, M., Weinkauff, S. and Buchner, J. (2019) ‘Small heat shock proteins: Simplicity meets complexity’, *Journal of Biological Chemistry*, 294(6), pp. 2121–2132. doi: 10.1074/jbc.REV118.002809.
- Hawkes, R., Grutter, M. G. and Schellman, J. (1984) ‘Thermodynamic stability and point mutations of bacteriophage T4 lysozyme’, *Journal of Molecular Biology*, 175(2), pp. 195–212. doi: 10.1016/0022-2836(84)90474-1.
- Heirbaut, M. *et al.* (2014) ‘Dissecting the functional role of the n-terminal domain of the human small heat shock protein HSPB6’, *PLoS ONE*, 9(8). doi: 10.1371/journal.pone.0105892.
- Heirbaut, M. *et al.* (2017) ‘Specific sequences in the N-terminal domain of human small heat-shock protein HSPB6 dictate preferential hetero-oligomerization with the orthologue HSPB1’, *Journal of Biological Chemistry*. © 2017 ASBMB. Currently published by Elsevier Inc; originally published by American Society for

Biochemistry and Molecular Biology, 292(24), pp. 9944–9957. doi: 10.1074/jbc.M116.773515.

Hilario, E. *et al.* (2011) ‘Crystal structures of xanthomonas small heat shock protein provide a structural basis for an active molecular chaperone oligomer’, *Journal of Molecular Biology*. Elsevier Ltd, 408(1), pp. 74–86. doi: 10.1016/j.jmb.2011.02.004.

Hopp, T. P. and Woods, K. R. (1981) ‘Prediction of protein antigenic determinants from amino acid sequences’, *Proceedings of the National Academy of Sciences of the United States of America*, 78(6 I), pp. 3824–3828. doi: 10.1073/pnas.78.6.3824.

Horwitz, J. (1992) ‘ α -Crystallin can function as a molecular chaperone’, *Proceedings of the National Academy of Sciences of the United States of America*, 89(21), pp. 10449–10453. doi: 10.1073/pnas.89.21.10449.

Horwitz, J. *et al.* (1998) ‘Mutation of α B-crystallin: Effects on chaperone-like activity’, *International Journal of Biological Macromolecules*, 22(3–4), pp. 263–269. doi: 10.1016/S0141-8130(98)00024-5.

Houlden, H. *et al.* (2008) ‘Mutations in the HSP27 (HSPB1) gene cause dominant, recessive, and sporadic distal HMN/CMT type 2’, *Neurology*, 71(21), pp. 1660–1668. doi: 10.1212/01.wnl.0000319696.14225.67.

Ishida, T. and Kinoshita, K. (2007) ‘PrDOS: Prediction of disordered protein regions from amino acid sequence’, *Nucleic Acids Research*, 35(SUPPL.2), pp. 460–464. doi: 10.1093/nar/gkm363.

Janowska, M. K. *et al.* (2019) ‘Mechanisms of small heat shock proteins’, *Cold Spring Harbor Perspectives in Biology*, 11(10). doi: 10.1101/cshperspect.a034025.

Jaroszyński, A. *et al.* (2021) ‘Heat Shock Protein 27 Is an Emerging Predictor of Contrast-Induced Acute Kidney Injury on Patients Subjected to Percutaneous Coronary Interventions’, *Cells*, 10(3), pp. 1–8. doi: 10.3390/cells10030684.

Jaya, N., Garcia, V. and Vierling, E. (2009) ‘Substrate binding site flexibility of the

small heat shock protein molecular chaperones', *Proceedings of the National Academy of Sciences of the United States of America*, 106(37), pp. 15604–15609. doi: 10.1073/pnas.0902177106.

Jehle, S. *et al.* (2010) 'Solid-state NMR and SAXS studies provide a structural basis for the activation of α B-crystallin oligomers', *Nature Structural and Molecular Biology*. Nature Publishing Group, 17(9), pp. 1037–1042. doi: 10.1038/nsmb.1891.

Jehle, S. *et al.* (2011) 'N-terminal domain of α B-crystallin provides a conformational switch for multimerization and structural heterogeneity', *Proceedings of the National Academy of Sciences of the United States of America*, 108(16), pp. 6409–6414. doi: 10.1073/pnas.1014656108.

Jovcevski, B. *et al.* (2015) 'Phosphomimics destabilize Hsp27 oligomeric assemblies and enhance chaperone activity', *Chemistry and Biology*. Elsevier Ltd, 22(2), pp. 186–195. doi: 10.1016/j.chembiol.2015.01.001.

Jovcevski, B. *et al.* (2018) 'The influence of the N-terminal region proximal to the core domain on the assembly and chaperone activity of α B-crystallin', *Cell Stress and Chaperones*. Cell Stress and Chaperones, 23(5), pp. 827–836. doi: 10.1007/s12192-018-0889-y.

Kaldis, A., Atkinson, B. G. and Heikkila, J. J. (2004) 'Molecular chaperone function of the *Rana catesbeiana* small heat shock protein, hsp30', *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, 139(2), pp. 175–182. doi: 10.1016/j.cbpb.2004.08.006.

Kappé, G. *et al.* (2004) 'Tsp36, a tapeworm small heat-shock protein with a duplicated α -crystallin domain, forms dimers and tetramers with good chaperone-like activity', *Proteins: Structure, Function and Genetics*, 57(1), pp. 109–117. doi: 10.1002/prot.20220.

Kawashima, T. *et al.* (2000) 'Archaeal adaptation to higher temperatures revealed by genomic sequence of *Thermoplasma volcanium*', *Proceedings of the National*

Academy of Sciences of the United States of America, 97(26), pp. 14257–14262.
doi: 10.1073/pnas.97.26.14257.

Kazakov, A. S. *et al.* (2009) ‘Thermally induced structural changes of intrinsically disordered small heat shock protein Hsp22’, *Biophysical Chemistry*. Elsevier B.V., 145(2–3), pp. 79–85. doi: 10.1016/j.bpc.2009.09.003.

Kennaway, C. K. *et al.* (2005) ‘Dodecameric structure of the small heat shock protein Acr1 from *Mycobacterium tuberculosis*’, *Journal of Biological Chemistry*. © 2005 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 280(39), pp. 33419–33425. doi: 10.1074/jbc.M504263200.

Kho, J. *et al.* (2021) ‘De Novo Design, Synthesis, and Mechanistic Evaluation of Short Peptides That Mimic Heat Shock Protein 27 Activity’, *ACS Medicinal Chemistry Letters*, 12(5), pp. 713–719. doi: 10.1021/acsmchemlett.0c00609.

Kim, K. K., Kim, R. and Kim, S. H. (1998) ‘Crystal structure of a small heat-shock protein’, *Nature*, 394(6693), pp. 595–599. doi: 10.1038/29106.

Kim, R. *et al.* (1998) ‘Small heat shock protein of *Methanococcus jannaschii*, a hyperthermophile’, *Proceedings of the National Academy of Sciences of the United States of America*, 95(16), pp. 9129–9133. doi: 10.1073/pnas.95.16.9129.

Kim, R. *et al.* (2003) ‘On the mechanism of chaperone activity of the small heat-shock protein of *Methanococcus jannaschii*’, *Proceedings of the National Academy of Sciences of the United States of America*, 100(14), pp. 8151–8155. doi: 10.1073/pnas.1032940100.

Klevit, R. E. (2020) ‘Peeking from behind the veil of enigma: emerging insights on small heat shock protein structure and function’, *Cell Stress and Chaperones*. Cell Stress and Chaperones, 25(4), pp. 573–580. doi: 10.1007/s12192-020-01092-2.

Kocabiyik, S. and Aygar, S. (2012) ‘Improvement of protein stability and enzyme recovery under stress conditions by using a small HSP (tpv-HSP 14.3) from

- Thermoplasma volcanium', *Process Biochemistry*. Elsevier Ltd, 47(11), pp. 1676–1683. doi: 10.1016/j.procbio.2011.11.014.
- Kolb, S. J. *et al.* (2010) 'Mutant small heat shock protein B3 causes motor neuropathy: Utility of a candidate gene approach', *Neurology*, 74(6), pp. 502–506. doi: 10.1212/WNL.0b013e3181cef84a.
- Koteiche, H. A. and Mchaourab, H. S. (2002) 'The determinants of the oligomeric structure in Hsp16.5 are encoded in the α -crystallin domain', *FEBS Letters*, 519(1–3), pp. 16–22. doi: 10.1016/S0014-5793(02)02688-1.
- Kovacs, D. and Tompa, P. (2012) 'Diverse functional manifestations of intrinsic structural disorder in molecular chaperones', *Biochemical Society Transactions*, 40(5), pp. 963–968. doi: 10.1042/BST20120108.
- Kriehuber, T. *et al.* (2010) 'Independent evolution of the core domain and its flanking sequences in small heat shock proteins', *FASEB Journal*, 24(10), pp. 3633–3642. doi: 10.1096/fj.10-156992.
- Krishna Sharma, K., Kaur, H. and Kester, K. (1997) 'Identification of binding sites in bovine lens α -crystallin during chaperone action', *Investigative Ophthalmology and Visual Science*, 38(4), pp. 217–222.
- Kuhlman, B. and Baker, D. (2000) 'Native protein sequences are close to optimal for their structures', *Proceedings of the National Academy of Sciences of the United States of America*, 97(19), pp. 10383–10388. doi: 10.1073/pnas.97.19.10383.
- Kuntal, B. K., Aparoy, P. and Reddanna, P. (2010) 'EasyModeller: A graphical interface to MODELLER', *BMC Research Notes*, 3, pp. 1–5. doi: 10.1186/1756-0500-3-226.
- Kurokawa, N. *et al.* (2021) 'Oligomeric structural transition of hspb1 from chinese hamster', *International Journal of Molecular Sciences*, 22(19). doi: 10.3390/ijms221910797.

- Laksanalamai, P. and Robb, F. T. (2004) 'Small heat shock proteins from extremophiles: A review', *Extremophiles*, 8(1), pp. 1–11. doi: 10.1007/s00792-003-0362-3.
- Lambert, W. *et al.* (2011) 'Subunit arrangement in the dodecameric chloroplast small heat shock protein Hsp21', *Protein Science*, 20(2), pp. 291–301. doi: 10.1002/pro.560.
- Lee, G. J., Pokala, N. and Vierling, E. (1995) 'Structure and in vitro molecular chaperone activity of cytosolic small heat shock proteins from pea', *Journal of Biological Chemistry*, 270(18), pp. 10432–10438. doi: 10.1074/jbc.270.18.10432.
- Van Der Lee, R. *et al.* (2014) 'Classification of intrinsically disordered regions and proteins', *Chemical Reviews*, 114(13), pp. 6589–6631. doi: 10.1021/cr400525m.
- Leroux, M. R. *et al.* (1997) 'Structure-function studies on small heat shock protein oligomeric assembly and interaction with unfolded polypeptides', *Journal of Biological Chemistry*, 272(39), pp. 24646–24666. doi: 10.1074/jbc.272.39.24646.
- Liu, L., Chen, J. Y., *et al.* (2015) 'Active-State Structures of a Small Heat-Shock Protein Revealed a Molecular Switch for Chaperone Function', *Structure*. Elsevier Ltd, 23(11), pp. 2066–2075. doi: 10.1016/j.str.2015.08.015.
- Liu, L., Chen, J., *et al.* (2015) 'Crystal structure and function of an unusual dimeric Hsp20.1 provide insight into the thermal protection mechanism of small heat shock proteins', *Biochemical and Biophysical Research Communications*. Elsevier Ltd, 458(2), pp. 429–434. doi: 10.1016/j.bbrc.2015.01.134.
- De Los Reyes, T. and Casas-Tintó, S. (2022) 'Neural functions of small heat shock proteins', *Neural Regeneration Research*, 17(3), pp. 512–515. doi: 10.4103/1673-5374.320975.
- Maiti, M., Kono, M. and Chakrabarti, B. (1988) 'Heat-induced changes in the conformation of α - and β -crystalline: Unique thermal stability of α -crystallin', *FEBS Letters*, 236(1), pp. 109–114. doi: 10.1016/0014-5793(88)80295-3.

Mani, N. *et al.* (2016) 'Multiple oligomeric structures of a bacterial small heat shock protein', *Scientific Reports*. Nature Publishing Group, 6(March), pp. 1–12. doi: 10.1038/srep24019.

Mao, Q., Ke, D. and Chang, Z. (2001) 'Electrostatic interactions play a critical role in Mycobacterium tuberculosis Hsp16.3 binding of substrate proteins', *Biokhimiya*, 66(8), pp. 1111–1116.

McDonald, E. T. *et al.* (2012) 'Sequence, structure, and dynamic determinants of Hsp27 (HspB1) equilibrium dissociation are encoded by the N-terminal domain', *Biochemistry*, 51(6), pp. 1257–1268. doi: 10.1021/bi2017624.

McHaourab, H. S., Lin, Y. L. and Spiller, B. W. (2012) 'Crystal structure of an activated variant of small heat shock protein Hsp16.5', *Biochemistry*, 51(25), pp. 5105–5112. doi: 10.1021/bi300525x.

Mishra, S. *et al.* (2018) 'Engineering of a Polydisperse Small Heat-Shock Protein Reveals Conserved Motifs of Oligomer Plasticity', *Structure*. Elsevier Ltd., 26(8), pp. 1116-1126.e4. doi: 10.1016/j.str.2018.05.015.

Mogk, A. *et al.* (2003) 'Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation', *Molecular Microbiology*, 50(2), pp. 585–595. doi: 10.1046/j.1365-2958.2003.03710.x.

Mogk, A., Ruger-Herreros, C. and Bukau, B. (2019) 'Cellular functions and mechanisms of action of small heat shock proteins', *Annual Review of Microbiology*, 73, pp. 89–110. doi: 10.1146/annurev-micro-020518-115515.

Montfort, B. R. O. B. V. A. N., Slingsby, C. and Elizabeth, V. I. E. R. L. I. N. G. (2002) 'STRUCTURE AND FUNCTION OF THE SMALL HEAT SHOCK PROTEIN / a-CRYSTALLIN FAMILY OF MOLECULAR CHAPERONES a focus on the mammalian proteins and cell biology was prepared by ofsHsps in h u m a n disease and eye lens structure are summarized by sev- substrate in', *Advances*, 59.

- Van Montfort, R. L. M. *et al.* (2001) ‘Crystal structure and assembly of a eukaryotic small heat shock protein’, *Nature Structural Biology*, 8(12), pp. 1025–1030. doi: 10.1038/nsb722.
- Moon, S. *et al.* (2019) ‘Structural and mutational analyses of psychrophilic and mesophilic adenylate kinases highlight the role of hydrophobic interactions in protein thermal stability’, *Structural Dynamics*, 6(2). doi: 10.1063/1.5089707.
- Morris, A. M. *et al.* (2008) ‘Glutamic acid residues in the C-terminal extension of small heat shock protein 25 are critical for structural and functional integrity’, *FEBS Journal*, 275(23), pp. 5885–5898. doi: 10.1111/j.1742-4658.2008.06719.x.
- Morrison, L. E. *et al.* (2003) ‘Mimicking phosphorylation of α B-crystallin on serine-59 is necessary and sufficient to provide maximal protection of cardiac myocytes from apoptosis’, *Circulation Research*, 92(2), pp. 203–211. doi: 10.1161/01.RES.0000052989.83995.A5.
- Moutaoufik, M. T. *et al.* (2017) ‘Effect of N-terminal region of nuclear *Drosophila melanogaster* small heat shock protein DmHsp27 on function and quaternary structure’, *PLoS ONE*, 12(5), pp. 1–19. doi: 10.1371/journal.pone.0177821.
- Muranova, L. K. *et al.* (2015) ‘Characterization of mutants of human small heat shock protein HspB1 carrying replacements in the N-terminal domain and associated with hereditary motor neuron diseases’, *PLoS ONE*, 10(5), pp. 1–24. doi: 10.1371/journal.pone.0126248.
- Muranova, L. K. *et al.* (2021) ‘Quaternary structure and hetero- oligomerization of recombinant human small heat shock protein hspb7 (Cvhsp)’, *International Journal of Molecular Sciences*, 22(15). doi: 10.3390/ijms22157777.
- Nakamoto, H. and Vigh, L. (2007) ‘The small heat shock proteins and their clients’, *Cellular and Molecular Life Sciences*, 64(3), pp. 294–306. doi: 10.1007/s00018-006-6321-2.
- Nandi, S. K. *et al.* (2015) ‘Role of subunit exchange and electrostatic interactions

on the chaperone activity of *Mycobacterium leprae* HSP18', *PLoS ONE*, 10(6), pp. 1–27. doi: 10.1371/journal.pone.0129734.

Navarro- zaragoza, J. *et al.* (2021) 'Could small heat shock protein hsp27 be a first- line target for preventing protein aggregation in parkinson's disease?', *International Journal of Molecular Sciences*, 22(6), pp. 1–13. doi: 10.3390/ijms22063038.

Nefedova, V. V. *et al.* (2015) 'Small Heat Shock Proteins and Distal Hereditary Neuropathies', *Biochemistry (Moscow)*, 80(13), pp. 1734–1747. doi: 10.1134/S000629791513009X.

Nelson, R. *et al.* (2005) 'Structure of the cross- β spine of amyloid-like fibrils', *Nature*, 435(7043), pp. 773–778. doi: 10.1038/nature03680.

Oldfield, C. J. *et al.* (2019) *Introduction to intrinsically disordered proteins and regions, Intrinsically Disordered Proteins*. Elsevier Inc. doi: 10.1016/b978-0-12-816348-1.00001-6.

Pasta, S. Y. *et al.* (2003) 'Role of the conserved SRLFDQFFG region of α -crystallin, a small heat shock protein: Effect on oligomeric size, subunit exchange, and chaperone-like activity', *Journal of Biological Chemistry*. © 2003 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 278(51), pp. 51159–51166. doi: 10.1074/jbc.M307523200.

Patel, S., Vierling, E. and Tama, F. (2014) 'Replica exchange molecular dynamics simulations provide insight into substrate recognition by small heat shock proteins', *Biophysical Journal*. Biophysical Society, 106(12), pp. 2644–2655. doi: 10.1016/j.bpj.2014.04.048.

Pettersen, E. F. *et al.* (2004) 'UCSF Chimera - A visualization system for exploratory research and analysis', *Journal of Computational Chemistry*, 25(13), pp. 1605–1612. doi: 10.1002/jcc.20084.

Phadte, A. S., Santhoshkumar, P. and Sharma, K. K. (2018) ‘Characterization of an N-terminal mutant of α A-crystallin α A–R21Q associated with congenital cataract’, *Experimental Eye Research*. Elsevier, 174(May), pp. 185–195. doi: 10.1016/j.exer.2018.05.016.

Phadte, A. S., Sluzala, Z. B. and Fort, P. E. (2021) ‘Therapeutic potential of α -crystallins in retinal neurodegenerative diseases’, *Antioxidants*, 10(7), pp. 1–13. doi: 10.3390/antiox10071001.

Plater, M. L., Goode, D. and Crabbe, M. J. C. (1996) ‘Effects of site-directed mutations on the chaperone-like activity of α B-crystallin’, *Journal of Biological Chemistry*. © 1996 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 271(45), pp. 28558–28566. doi: 10.1074/jbc.271.45.28558.

Poulain, P., Gelly, J. C. and Flatters, D. (2010) ‘Detection and architecture of small heat shock protein Monomers’, *PLoS ONE*, 5(4). doi: 10.1371/journal.pone.0009990.

Raju, M. *et al.* (2014) ‘Erratum: Addition of α a-crystallin sequence 164-173 to a mini-chaperone DVFIFLDVKHFSPELDT alters the conformation but not the chaperone-like activity (Biochemistry (2014) 53:16 (2615-2623) DOI: 10.1021/bi4017268)’, *Biochemistry*, 53(21), p. 3521. doi: 10.1021/bi500563d.

Raju, M., Santhoshkumar, P. and Sharma, K. K. (2016) ‘Alpha-crystallin-derived peptides as therapeutic chaperones’, *Biochimica et Biophysica Acta - General Subjects*. Elsevier B.V., 1860(1), pp. 246–251. doi: 10.1016/j.bbagen.2015.06.010.

Raman, B. and Rao, C. M. (1994) ‘Chaperone-like activity and quaternary structure of α -crystallin’, *Journal of Biological Chemistry*, 269(44), pp. 27264–27268. doi: 10.1016/s0021-9258(18)46978-5.

Reddy, V. S. *et al.* (2021) ‘Impact of chronic hyperglycemia on Small Heat Shock Proteins in diabetic rat brain’, *Archives of Biochemistry and Biophysics*. Elsevier Inc., 701(July 2020), p. 108816. doi: 10.1016/j.abb.2021.108816.

- Reinle, K., Mogk, A. and Bukau, B. (2021) 'The Diverse Functions of Small Heat Shock Proteins in the Proteostasis Network', *Journal of Molecular Biology*. Elsevier Ltd, (xxxx), p. 167157. doi: 10.1016/j.jmb.2021.167157.
- Remington, S., Wiegand, G. and Huber, R. (1982) 'Crystallographic refinement and atomic models of two different forms of citrate synthase at 2.7 and 1.7 Å resolution', *Journal of Molecular Biology*, 158(1), pp. 111–152. doi: 10.1016/0022-2836(82)90452-1.
- Romero, P. *et al.* (2001) 'Sequence complexity of disordered protein', *Proteins: Structure, Function and Genetics*, 42(1), pp. 38–48. doi: 10.1002/1097-0134(20010101)42:1<38::AID-PROT50>3.0.CO;2-3.
- Rutsdottir, G. *et al.* (2017) 'Structural model of dodecameric heat-shock protein Hsp21: Flexible N-terminal arms interact with client proteins while C-terminal tails maintain the dodecamer and chaperone activity', *Journal of Biological Chemistry*, 292(19), pp. 8103–8121. doi: 10.1074/jbc.M116.766816.
- Saji, H. *et al.* (2008) 'Role of the IXI/V motif in oligomer assembly and function of StHsp14.0, a small heat shock protein from the acidothermophilic archaeon, *Sulfolobus tokodaii* strain 7', *Proteins: Structure, Function and Genetics*, 71(2), pp. 771–782. doi: 10.1002/prot.21762.
- Santhanagopalan, I. *et al.* (2018) 'It takes a dimer to tango: Oligomeric small heat shock proteins dissociate to capture substrate', *Journal of Biological Chemistry*. © 2018 Santhanagopalan *et al.*, 293(51), pp. 19511–19521. doi: 10.1074/jbc.RA118.005421.
- Santhoshkumar, P. and Sharma, K. K. (2001) 'Analysis of α -crystallin chaperone function using restriction enzymes and citrate synthase', *Molecular Vision*, 7(July), pp. 172–177.
- Santhoshkumar, P. and Sharma, K. K. (2004) 'Inhibition of amyloid fibrillogenesis and toxicity by a peptide chaperone', *Molecular and Cellular Biochemistry*, 267(1–2), pp. 147–155. doi: 10.1023/B:MCBI.0000049373.15558.b8.

- Sawaya, M. R. *et al.* (2007) 'Atomic structures of amyloid cross- β spines reveal varied steric zippers', *Nature*, 447(7143), pp. 453–457. doi: 10.1038/nature05695.
- SCHERAGA, H. A., NEMETHY, G. and STEINBERG, I. Z. (1962) 'The contribution of hydrophobic bonds to the thermal stability of protein conformations.', *The Journal of biological chemistry*, 237, pp. 2506–2508. doi: 10.1016/s0021-9258(19)73780-6.
- Segerer, A., Langworthy, T. A. and Stetter, K. O. (1988) 'Thermoplasma acidophilum and Thermoplasma volcanium sp. nov. from Solfatara Fields', *Systematic and Applied Microbiology*. Gustav Fischer Verlag, Stuttgart · New York, 10(2), pp. 161–171. doi: 10.1016/S0723-2020(88)80031-6.
- Selig, E. E. *et al.* (2020) 'N- and C-terminal regions of α B-crystallin and Hsp27 mediate inhibition of amyloid nucleation, fibril binding, and fibril disaggregation', *The Journal of biological chemistry*, 295(29), pp. 9838–9854. doi: 10.1074/jbc.RA120.012748.
- Sharma, K. K. *et al.* (1998) 'Interaction of 1,1'-bi(4-anilino)naphthalene-5,5'-disulfonic acid with α -crystallin', *Journal of Biological Chemistry*. © 1998 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 273(15), pp. 8965–8970. doi: 10.1074/jbc.273.15.8965.
- Sharma, K. K. *et al.* (2000) 'Synthesis and characterization of a peptide identified as a functional element in α A-crystallin', *Journal of Biological Chemistry*, 275(6), pp. 3767–3771. doi: 10.1074/jbc.275.6.3767.
- Shashidharamurthy, R. *et al.* (2005) 'Mechanism of chaperone function in small heat shock proteins: Dissociation of the HSP27 oligomer is required for recognition and binding of destabilized T4 lysozyme', *Journal of Biological Chemistry*. © 2005 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 280(7), pp. 5281–5289. doi: 10.1074/jbc.M407236200.

- Shatov, V. M. *et al.* (2018) ‘The role of the arginine in the conserved N-terminal domain RLFDQxFG motif of human small heat shock proteins HspB1, HspB4, HspB5, HspB6, and HspB8’, *International Journal of Molecular Sciences*, 19(7). doi: 10.3390/ijms19072112.
- Shatov, V. M. and Gusev, N. B. (2020) ‘Physico-chemical properties of two point mutants of small heat shock protein HspB6 (Hsp20) with abrogated cardioprotection’, *Biochimie*. Elsevier Ltd, 174, pp. 126–135. doi: 10.1016/j.biochi.2020.04.021.
- Shatov, V. M., Sluchanko, N. N. and Gusev, N. B. (2021) ‘Replacement of Arg in the conserved N-terminal RLFDQxFG motif affects physico-chemical properties and chaperone-like activity of human small heat shock protein HspB8 (Hsp22)’, *PLoS ONE*, 16(6 June), pp. 1–15. doi: 10.1371/journal.pone.0253432.
- Shatov, V. M., Strelkov, S. V. and Gusev, N. B. (2020) ‘The heterooligomerization of human small heat shock proteins is controlled by conserved motif located in the n-terminal domain’, *International Journal of Molecular Sciences*, 21(12), pp. 1–18. doi: 10.3390/ijms21124248.
- Shi, J. *et al.* (2006) ‘Cryoelectron microscopy and EPR analysis of engineered symmetric and polydisperse Hsp16.5 assemblies reveals determinants of polydispersity and substrate binding’, *Journal of Biological Chemistry*. © 2006 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 281(52), pp. 40420–40428. doi: 10.1074/jbc.M608322200.
- Shroff, N. P. *et al.* (2001) ‘Substituted hydrophobic and hydrophilic residues at methionine-68 influence the chaperone-like function of α B-crystallin’, *Molecular and Cellular Biochemistry*, 220(1–2), pp. 127–133. doi: 10.1023/A:1010834107809.
- Skouri-Panet, F. *et al.* (2006) ‘sHSPs under temperature and pressure: The opposite behaviour of lens alpha-crystallins and yeast HSP26’, *Biochimica et Biophysica*

Acta - Proteins and Proteomics, 1764(3), pp. 372–383. doi:
10.1016/j.bbapap.2005.12.011.

Specht, S. *et al.* (2011) ‘Hsp42 is required for sequestration of protein aggregates into deposition sites in *Saccharomyces cerevisiae*’, *Journal of Cell Biology*, 195(4), pp. 617–629. doi: 10.1083/jcb.201106037.

Spector, A. and Zorn, M. (1967) ‘Studies upon the sulfhydryl groups of calf lens alpha-crystallins.’, *Journal of Biological Chemistry*. © 1967 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 242(16), pp. 3594–3600. doi:
10.1016/s0021-9258(18)95850-3.

Sreekumar, P. G. *et al.* (2013) ‘Antiapoptotic properties of α -crystallin-derived peptide chaperones and characterization of their uptake transporters in human RPE cells’, *Investigative Ophthalmology and Visual Science*, 54(4), pp. 2787–2798. doi: 10.1167/iovs.12-11571.

Stamler, R. *et al.* (2005) ‘Wrapping the α -crystallin domain fold in a chaperone assembly’, *Journal of Molecular Biology*, 353(1), pp. 68–79. doi:
10.1016/j.jmb.2005.08.025.

Stengel, F. *et al.* (2012) ‘Dissecting heterogeneous molecular chaperone complexes using a mass spectrum deconvolution approach’, *Chemistry and Biology*. Elsevier Ltd, 19(5), pp. 599–607. doi: 10.1016/j.chembiol.2012.04.007.

Stromer, T. *et al.* (2004) ‘Analysis of the regulation of the molecular chaperone Hsp26 by temperature-induced dissociation: The N-terminal domain is important for oligomer assembly and the binding of unfolding proteins’, *Journal of Biological Chemistry*, 279(12), pp. 11222–11228. doi: 10.1074/jbc.M310149200.

Studer, S. *et al.* (2002) ‘A critical motif for oligomerization and chaperone activity of bacterial α -heat shock proteins’, *European Journal of Biochemistry*, 269(14), pp. 3578–3586. doi: 10.1046/j.1432-1033.2002.03049.x.

- Studier, F. W. and Moffatt, B. A. (1986) 'Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes', *Journal of Molecular Biology*, 189(1), pp. 113–130. doi: 10.1016/0022-2836(86)90385-2.
- V. Sudnitsyna, M. *et al.* (2012) 'The Role of Intrinsically Disordered Regions in the Structure and Functioning of Small Heat Shock Proteins', *Current Protein & Peptide Science*, 13(1), pp. 76–85. doi: 10.2174/138920312799277875.
- Sudnitsyna, M. V, Sluchanko, N. N. and Gusev, N. B. (2015) 'The Big Book on Small Heat Shock Proteins', 8, pp. 255–266. doi: 10.1007/978-3-319-16077-1.
- Sun, Y. and MacRae, T. H. (2005) 'Small heat shock proteins: Molecular structure and chaperone function', *Cellular and Molecular Life Sciences*, 62(21), pp. 2460–2476. doi: 10.1007/s00018-005-5190-4.
- Takeda, K. *et al.* (2011) 'Dimer structure and conformational variability in the N-terminal region of an archaeal small heat shock protein, StHsp14.0', *Journal of Structural Biology*. Elsevier Inc., 174(1), pp. 92–99. doi: 10.1016/j.jsb.2010.12.006.
- Thériault, J. R. *et al.* (2004) 'Essential role of the NH₂-terminal WD/EPF motif in the phosphorylation-activated protective function of mammalian Hsp27', *Journal of Biological Chemistry*, 279(22), pp. 23463–23471. doi: 10.1074/jbc.M402325200.
- Thompson, M. J. *et al.* (2006) 'The 3D profile method for identifying fibril-forming segments of proteins', *Proceedings of the National Academy of Sciences of the United States of America*, 103(11), pp. 4074–4078. doi: 10.1073/pnas.0511295103.
- Treweek, T. M. *et al.* (2015) 'Small heat-shock proteins: Important players in regulating cellular proteostasis', *Cellular and Molecular Life Sciences*, 72(3), pp. 429–451. doi: 10.1007/s00018-014-1754-5.
- Ungelenk, S. *et al.* (2016) 'Small heat shock proteins sequester misfolding proteins

- in near-native conformation for cellular protection and efficient refolding’, *Nature Communications*. Nature Publishing Group, 7, pp. 1–14. doi: 10.1038/ncomms13673.
- Usui, K. (2004) ‘Expression and biochemical characterization of two small heat shock proteins from the thermoacidophilic crenarchaeon *Sulfolobus tokodaii* strain 7’, *Protein Science*, 13(1), pp. 134–144. doi: 10.1110/ps.03264204.
- Usui, K. *et al.* (2004) ‘Role of the N-terminal region of the crenarchaeal sHsp, StHsp14.0, in thermal-induced disassembly of the complex and molecular chaperone activity’, *Biochemical and Biophysical Research Communications*, 315(1), pp. 113–118. doi: 10.1016/j.bbrc.2004.01.031.
- Vacic, V. *et al.* (2012) ‘Disease-Associated Mutations Disrupt Functionally Important Regions of Intrinsic Protein Disorder’, *PLoS Computational Biology*, 8(10). doi: 10.1371/journal.pcbi.1002709.
- Vacic, V. and Iakoucheva, L. M. (2012) ‘Disease mutations in disordered regions - Exception to the rule?’, *Molecular BioSystems*, 8(1), pp. 27–32. doi: 10.1039/c1mb05251a.
- Vos, M. J. *et al.* (2010) ‘HSPB7 is the most potent polyQ aggregation suppressor within the HSPB family of molecular chaperones’, *Human Molecular Genetics*, 19(23), pp. 4677–4693. doi: 10.1093/hmg/ddq398.
- Waters, E. R. and Vierling, E. (1999) ‘Chloroplast small heat shock proteins: Evidence for atypical evolution of an organelle-localized protein’, *Proceedings of the National Academy of Sciences of the United States of America*, 96(25), pp. 14394–14399. doi: 10.1073/pnas.96.25.14394.
- Webster, J. M. *et al.* (2019) ‘Small heat shock proteins, big impact on protein aggregation in neurodegenerative disease’, *Frontiers in Pharmacology*, 10(September), pp. 1–18. doi: 10.3389/fphar.2019.01047.
- Webster, J. M. *et al.* (2020) ‘Hsp22 with an n-terminal domain truncation mediates

- a reduction in tau protein levels', *International Journal of Molecular Sciences*, 21(15), pp. 1–17. doi: 10.3390/ijms21155442.
- Weeks, S. D. *et al.* (2014) 'Molecular structure and dynamics of the dimeric human small heat shock protein HSPB6', *Journal of Structural Biology*. Elsevier Inc., 185(3), pp. 342–354. doi: 10.1016/j.jsb.2013.12.009.
- Wen, Z. *et al.* (2010) 'Importance of a Potential Salt Bridge and Hydrophobic Core in the Function and Oligomerization of a Small Heat Shock Protein', *Protein & Peptide Letters*, 17(6), pp. 751–758. doi: 10.2174/092986610791190273.
- Wittig, I., Braun, H. P. and Schägger, H. (2006) 'Blue native PAGE', *Nature Protocols*, 1(1), pp. 418–428. doi: 10.1038/nprot.2006.62.
- Wong, Y. Q. *et al.* (2010) 'Methionine oxidation induces amyloid fibril formation by full-length apolipoprotein A-I', *Proceedings of the National Academy of Sciences of the United States of America*, 107(5), pp. 1977–1982. doi: 10.1073/pnas.0910136107.
- Wu, D. *et al.* (2019) 'The N terminus of the small heat shock protein HSPB7 drives its polyQ aggregation-suppressing activity', *Journal of Biological Chemistry*, 294(25), pp. 9985–9994. doi: 10.1074/jbc.RA118.007117.
- Xia, X. Y. *et al.* (2014) 'A novel P20R mutation in the alpha-B crystallin gene causes autosomal dominant congenital posterior polar cataracts in a Chinese family', *BMC Ophthalmology*, 14(1), pp. 1–7. doi: 10.1186/1471-2415-14-108.
- Xiong, J. *et al.* (2020) 'Small heat shock proteins in cancers: Functions and therapeutic potential for cancer therapy', *International Journal of Molecular Sciences*, 21(18), pp. 1–22. doi: 10.3390/ijms21186611.
- Yang, H. *et al.* (2008) 'The mycobacterium tuberculosis small heat shock protein Hsp16.3 exposes hydrophobic surfaces at mild conditions: Conformational flexibility and molecular chaperone activity', *Protein Science*, 8(1), pp. 174–179. doi: 10.1110/ps.8.1.174.

Yelton, A. P. *et al.* (2011) ‘A semi-quantitative, synteny-based method to improve functional predictions for hypothetical and poorly annotated bacterial and archaeal genes’, *PLoS Computational Biology*, 7(10), pp. 1–12. doi: 10.1371/journal.pcbi.1002230.

Young, L. Sen *et al.* (1999) ‘Molecular characterization of *Oryza sativa* 16.9 kDa heat shock protein’, *Biochemical Journal*, 344(1), pp. 31–38. doi: 10.1042/0264-6021:3440031.

Yu, C. *et al.* (2021) ‘Structural basis of substrate recognition and thermal protection by a small heat shock protein’, *Nature Communications*. Springer US, 12(1), pp. 1–11. doi: 10.1038/s41467-021-23338-y.

Zahn, R. *et al.* (1996) ‘Chaperone activity and structure of monomeric polypeptide binding domains of groEL’, *Proceedings of the National Academy of Sciences of the United States of America*, 93(26), pp. 15024–15029. doi: 10.1073/pnas.93.26.15024.

Zhang, Xiaoling *et al.* (2021) ‘The role of heat shock proteins in the regulation of fibrotic diseases’, *Biomedicine and Pharmacotherapy*. Elsevier Masson SAS, 135, p. 111067. doi: 10.1016/j.biopha.2020.111067.

APPENDICES

A. Buffers and Solutions

SDS staining solution:

Coomassie (Brilliant Blue R)....	1g
Methanol.....	450 mL
Acetic Acid.....	100 mL
Water.....	450 mL

SDS de-staining solution

Methanol.....	450 mL
Acetic Acid.....	100 mL
Water.....	450 mL

NZY⁺ Broth (1000 mL)

NZY Amine (Caesin Hydrolysate).....	10g
Yeast Extract.....	5g
Sodium Chloride.....	5g

pH adjusted to 7.5 by 5M and 1M NaOH. (sterilized by autoclave)

Reconstitution of NZY⁺ Broth (1000 mL)

MgCl₂ (1M).....12.5 mL (sterilized by autoclave)

MgSO₄ (1M).....12.5 mL (sterilized by autoclave)

Glucose (2M).....10 mL (sterilized by autoclave)

**Luria Bertani Agar for QuikChange II Site-Directed mutagenesis kit protocol
(1000 mL)**

Yeast Extract.....5g

Tryptone.....10g

Sodium Chloride.....10g

Agar.....20g (2%)

pH adjusted to 7.0 with 5M NaOH. (sterilized by autoclave)

Luria Bertani Agar for BL21(DE3) cells (1000 mL)

Yeast Extract.....5g

Tryptone.....10g

Sodium Chloride.....10g

Agar.....15 g (1.5%)

pH adjusted to 7.5 with 5M NaOH. (sterilized by autoclave)

Transformation and Storage Solution (TSS) (50 mL)

2.5 mL DMSO

2.5 mL of 1M MgCl₂·6H₂O (Final 20 mM Mg⁺² ions) (sterilized by autoclave)

35 mL Luria Bertani Broth, pH 7.5 (sterilized by autoclave)

5 g of PEG-8000 dissolved in 10 mL of distilled water (sterilized by autoclave)

TAE Buffer (50x) (1000 mL)

Tris Base.....242.0 g

Glacial acetic acid.....57.1 mL

EDTA (0.5M).....100 mL

Preparation of SDS Buffer (1000 ml)

Tris Base.....3.02g

Glycine.....14.4g

SDS.....1.0g

Sterilized by autoclave

Gel Buffer (50 ml)

Imidazole.....0.255g

6-Aminohexanoic Acid....9.84g

Distilled Water.....up to 50 ml

pH adjusted to 7.0 by 5N HCl

B. Preparation of Gels

SDS Gel Preparation:

Chemicals	Separating Gel (12%)	Stacking Gel
Bis-Acrylamide (0.8%)	2400 μL	330 μL
Acrylamide (30%)		
Distilled Water	1200 μL	870 μL
Tris-HCl (pH 8.9)	1200 μL (1.88M)	400 μL (0.625 M)
SDS (0.5%)	1200 μL	400 μL
Ammonium per sulfate	30 μL	10 μL
TEMED	5 μL	2 μL
Incubation Time	30 minutes	30 minutes (or until set)

SDS Gel Preparation (For Twin Gel):

Chemicals	Separating Gel (12%)	Stacking Gel
Bis-Acrylamide (0.8%)	4800 μL	660 μL
Acrylamide (30%)		
Distilled Water	2400 μL	1740 μL
Tris-HCl (pH 8.9)	2400 μL (1.88M)	800 μL (0.625 M)
SDS (0.5%)	2400 μL	800 μL
Ammonium per sulfate	60 μL	20 μL
TEMED	10 μL	4 μL
Incubation Time	30 minutes	30 minutes (or until set)

Blue-Native Gel Preparation:

Chemicals	4 %	13 %
Bis-Acrylamide (0.8%)	680 μ L	1760 μ L
Acrylamide (30%)		
Gel Buffer (3x)	1665 μ L	1332 μ L
Glycerol	-	0.8 g
Distilled Water	2600 μ L	255 μ L
Ammonium per sulfate (10%)	50 μ L	20 μ L
TEMED	5 μ L	3 μ L
Incubation Time	30 minutes	30 minutes (or until set)

C. Primer sequence

Primers for amplification of target gene for Cloning:

NdeI Site primer: CATATGtatacaccataaagttctttacg

BamHI site primer: GTATGCTTGATGTGATTGGGTGGGATCC

D. Figures for Chapter 3

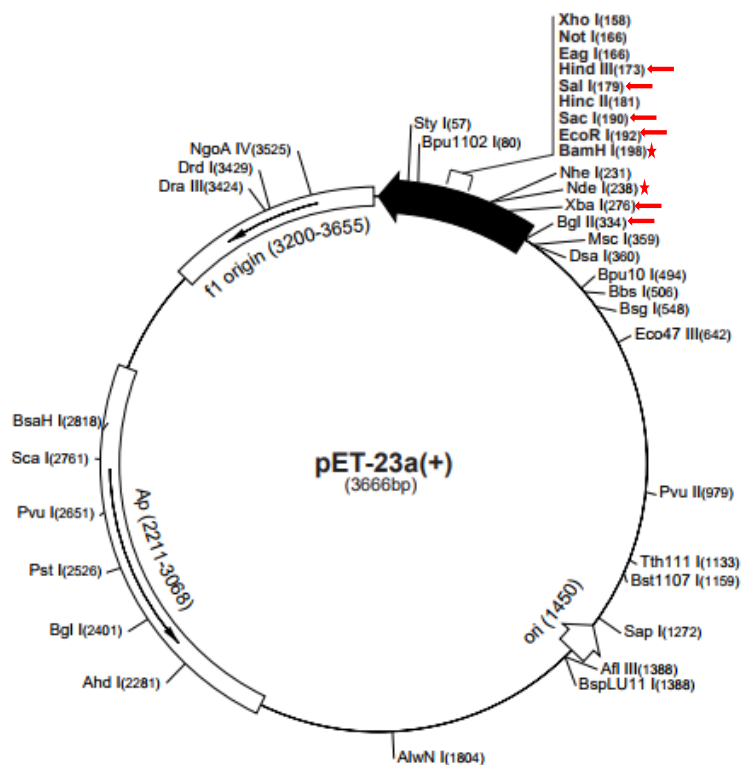


Figure 1: Vector Map of pET23a(+). The red star indicates the position where our gene (TVN0775) is inserted, *i.e.*, between NdeI (238 bp) and BamHI (198 bp). The red arrows indicate the restriction enzymes which were used for characterization of the gene (TVN0775). The enzymes used for characterization were HindIII (173), SalI (179), SacI (190), EcoRI (192), XbaI (276) and BglII (334)

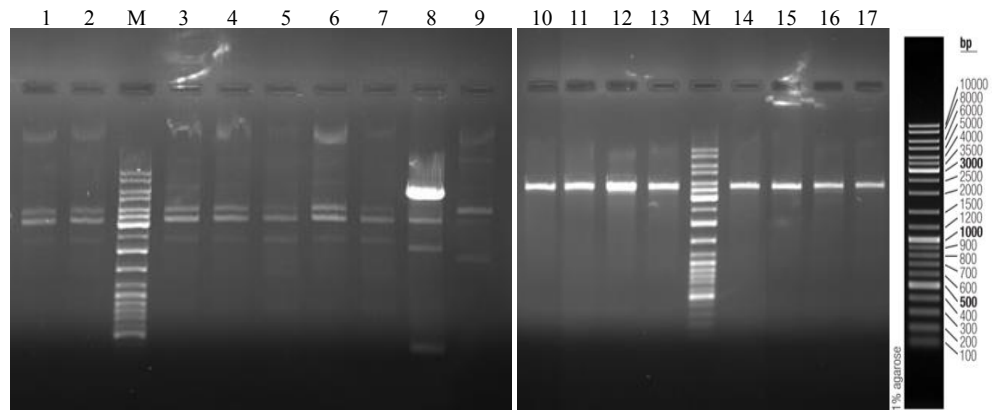


Figure 2: Isolated plasmid digested by restriction enzymes NdeI and BamHI. Lane1 – Lane 9 shows presence of more than one band, at coinciding position indicating problem with digestion which can be attributed to erroneous or no ligation of the gene. Lane 10 – Lane 17 also show single band of neatly digested pET23a(+) vector, however, there is no band smaller band, indication absence of insert in these colonies.

pET23_2T7 pET23a+gene pET23_2T7-terRC	----- TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCG -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- CAGCGTGACCCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTC -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- CTTCTCGCCACGTTTCGCCGGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTTAGG -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- GTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTC -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- ACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTT -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- CTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTC -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- TTTTGATTTATAAGGGATTTTGCCGATTTTCGGCTATTGGTTAAAAAATGAGCTGATTTA -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- ACAAAAATTTAACGCGAATTTAACAAAAATTAACGTTTACAATTTACGGTGGCACTTT -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- TCGGGGAAATGTGCGCGGAACCCCTATTTGTTATTTTCTAAATACATTCAAATATGTA -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- TCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGATAT -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- GAGTATTC AACATTTCCGTGTGCGCCTTATTCCTTTTTTTCGCGCATTTTCCTTCCTGT -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- TTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACG -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- AGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGA -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- AGAACGTTTTCCAATGATGACACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCG -----

Figure 3 continued

```
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
TATTGACGCCGGCAAGAGCAACTCGGTCGCCGCATACACTATTTCTCAGAATGACTTGGT  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
TGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATG  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
CAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCTGACAACGATCGG  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
AGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCCTTGA  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
TCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCC  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
TGCAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTCTAGCTTC  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
CCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTGCAGGACCACTTCTGCGCTC  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
GGCCCTCCCGCTGGCTGGTTTATGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCG  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
CGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACAC  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
GACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTC  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
ACTGATTAAGCATTGGTAACTGTGACACCAAGTTTACTCATATATACTTTAGATTGATTT  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
AAAACCTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGAC  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
CAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAA  
-----  
  
-----  
pET23_2T7  
pET23a+gene  
pET23_2T7-terRC  
-----  
AGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTGCAACAAAAAACC  
-----
```

Figure 3 continued

```
pET23_2T7 -----  
pET23a+gene AACTGGCTTCCAGCAGCGGTGGTTTGGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGT  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene AACTGGCTTCCAGCAGCGCGAGATACCAAATACTGCCTTCTAGTGTAGCCGTAGTTAGG  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene CCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACC  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene AGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTT  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene ACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGTTTCGTGCACACAGCCAGCTTGA  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene GCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCAGCT  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene TCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCG  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene CACGAGGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCTGTTCGGGTTTCGCCA  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene CCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAA  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene CGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTT  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene CTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGA  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene GCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGATATATGG  
pET23_2T7-terRC -----  
  
pET23_2T7 -----  
pET23a+gene TGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTCCGCTAT  
pET23_2T7-terRC -----
```

Figure 3 continued

pET23_2T7 pET23a+gene pET23_2T7-terRC	----- CGCTACGTGACTGGGTTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGGGCCCT -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- GACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- GCATGTGTCAGAGGTTTTACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCT -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- CATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTTCATCCGCTCCAGCTCGT -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- TGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGG -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- TTTTTTCCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGATTCTGTTCATGGGGGTAA -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- TGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCCC -----
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- GGTFACTGGAACGTTGTGAGGGTAAACAACCTGGCGGTATGGATGCGGCGGGACCAGAGAA -GTTAATTGAACGTTGTGAGGGTAAACAACCTGGCGGTATGGATGCG-CGGGACCAGAGAA
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- AAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTT-CCACAGGGT AAATCACTCAGGGTCATTGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGT
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- AGCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGC AGCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGC
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- GTTTCCAGACTTTACGAAACACGGAAACCGAAGACCATTTCATGTTGTTGCTCAGGTCGCA GTTTCCAGACTTTACGAAACACGGAAACCGAAGACCATTTCATGTTGTTGCTCAGGTCGCA
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- GACGTTTTGCAGCAGCAGTCGCTTCACGTTTCGCTCGCGTATCGGTGATTTCATTCTGCTAA GACGTTTTGCAGCAGCAGTCGCTTCACGTTTCGCTCGCGTATCGGTGATTTCATTCTGCTAA
pET23_2T7 pET23a+gene pET23_2T7-terRC	----- CCAGTAAGGCAACCCCGCCAGCCTAGCCGGTCTCAACGACAGGAGCAGCATCATGCGC CCAGTAAGGCAACCCCGCCAGCCTAGCCGGTCTCAACGACAGGAGCAGCATCATGCGC

Figure 3 continued

```

pET23_2T7 -----
pET23a+gene ACCCGTGGCCAGGACCCAACGCTGCCCGAGATCTCGATCCCGCGAAATTAATACGACTCA
pET23_2T7-terRC ACCCGTGGCCAGGACCCAACGCTGCCCGAGATCTCGATCCCGCGAAATTAATACGACTCA

pET23_2T7 -----ACCGGAATTCCTCTAGATAATTTTGTCTTACTTTAAG AAGGAG
pET23a+gene CTATAGGGGAGACCACAACGGTTTCCTCTAGAAAATAATTTTGTCTTACTTTAAG AAGGAG
pET23_2T7-terRC CTATAGGGGAGACCACAACGGTTTCCTCTAGAAAATAATTTTGTCTTACTTTAAG AAGGAG
                **      *      *      *      *      *      *      *      *      *      *

pET23_2T7 ATATA CATATG TATACACCATAAAAGTTCTTTACGAATGAGATGATAAAAAACGTATCGA
pET23a+gene ATATA CATATG TATACACCATAAAAGTTCTTTACGAATGAGATGATAAAAAACGTATCGA
pET23_2T7-terRC ATATA CATATG TATACACCATAAAAGTTCTTTACGAATGAGATGATAAAAAACGTATCGA
                *****

pET23_2T7 A TACTGTGAAAGAGGTCTCATCCTTTATATATCCACCAGTCACGTTATATCAAGATAGCT
pET23a+gene A TACTGTGAAAGAGGTCTCATCCTTTATATATCCACCAGTCACGTTATATCAAGATAGCT
pET23_2T7-terRC A TACTGTGAAAGAGGTCTCATCCTTTATATATCCACCAGTCACGTTATATCAAGATAGCT
                *****

pET23_2T7 CTGATCTGGTATTGGAAGCAGAAATGGCCGGGTTTGACAAGAAAAACATAAAGGTCTCGG
pET23a+gene CTGATCTGGTATTGGAAGCAGAAATGGCCGGGTTTGACAAGAAAAACATAAAGGTCTCGG
pET23_2T7-terRC CTGATCTGGTATTGGAAGCAGAAATGGCCGGGTTTGACAAGAAAAACATAAAGGTCTCGG
                *****

pET23_2T7 TAAATAAGAATGTACTCACTATAAGTGCGGAGAGAAAAGAGAGAATACTCTACCGTATATA
pET23a+gene TAAATAAGAATGTACTCACTATAAGTGCGGAGAGAAAAGAGAGAATACTCTACCGTATATA
pET23_2T7-terRC TAAATAAGAATGTACTCACTATAAGTGCGGAGAGAAAAGAGAGAATACTCTACCGTATATA
                *****

pET23_2T7 TCGATCAGCGCGTTGACAAAGTGTATAAAGTAGTTAAGCTGCCCGTAGAGATTGAGCAGC
pET23a+gene TCGATCAGCGCGTTGACAAAGTGTATAAAGTAGTTAAGCTGCCCGTAGAGATTGAGCAGC
pET23_2T7-terRC TCGATCAGCGCGTTGACAAAGTGTATAAAGTAGTTAAGCTGCCCGTAGAGATTGAGCAGC
                *****

pET23_2T7 AGGACATATCTGCTAAGTATAGTGAAGGCATACTTACAGTTAGAATGAAAACCAAGAACA
pET23a+gene AGGACATATCTGCTAAGTATAGTGAAGGCATACTTACAGTTAGAATGAAAACCAAGAACA
pET23_2T7-terRC AGGACATATCTGCTAAGTATAGTGAAGGCATACTTACAGTTAGAATGAAAACCAAGAACA
                *****

pET23_2T7 TAAAGAACGTAGAAATAGAATAAAATCATTTTTTAATAATAATATATATGAAAAGTATTG
pET23a+gene TAAAGAACGTAGAAATAGAATAAAATCATTTTTTAATAATAATATATATGAAAAGTATTG
pET23_2T7-terRC TAAAGAACGTAGAAATAGAATAAAATCATTTTTTAATAATAATATATATGAAAAGTATTG
                *****

pET23_2T7 CATCTATTGCTATAAGTATGCTTGATGTGATTGGGTGGGATCCGAATTCGAGCTCCGTCG
pET23a+gene CATCTATTGCTATAAGTATGCTTGATGTGATTGGGTGGGATCCGAATTCGAGCTCCGTCG
pET23_2T7-terRC CATCTATTGCTATAAGTATGCTTGATGTGATTGGGTGGGATCCGAATTCGAGCTCCGTCG
                *****

pET23_2T7 ACAAGCTTGCGGCCGCACTCGAG CACCACCACCACCACCAC TGAGATCCGGCTGCTAACA
pET23a+gene ACAAGCTTGCGGCCGCACTCGAG CACCACCACCACCACCAC TGAGATCCGGCTGCTAACA
pET23_2T7-terRC ACAAGCTTGCGGCCGCACTCGAG CACCACCACCACCACCAC TGAGATCCGGCTGCTAACA
                *****

pET23_2T7 AAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCCTGAGCAATAACTAGCATAACCC
pET23a+gene AAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCCTGAGCAATAACTAGCATAACCC
pET23_2T7-terRC AAGCCCGAAAGGAAGCTATGGGAC-----
                ***** *

pET23_2T7 TTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACCTATATCCGGAT
pET23a+gene TTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACCTATATCCGGAT
pET23_2T7-terRC -----

pET23_2T7 TGCGGAATGGGACGCGCCCTGTAGCGCGCATTAAGCGCGCGGGTGTGGTGTACGCG
pET23a+gene -----
pET23_2T7-terRC -----
    
```

Figure 3 continued

pET23 2T7	CAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTC
pET23a+gene	-----
pET23 2T7-terRC	-----
pET23 2T7	CTTCTCGCCACGTTTCGCCGGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGG
pET23a+gene	-----
pET23 2T7-terRC	-----
pET23_2T7	GTTCCGATTTAGTGCTTTACGGCACCCCTCGACCCCAAAAACTTGATAGGTGATGTTTAC
pET23a+gene	-----
pET23 2T7-terRC	-----
pET23_2T7	GTAGTGGGCCATCGCCCT
pET23a+gene	-----
pET23_2T7-terRC	-----

Figure 3 MSA of sequencing result of *Tpv* Hsp 14.3 cloned into pET23a(+) vector with the gene sequence available online at National Centre for Biotechnology Information (NCBI). Grey: T7 promoter; Pink:RBS; Light Blue: *Nde*I cut-site; Yellow: TVN0775 gene, Green: *Bam*HI cut-site; Red: His-tag

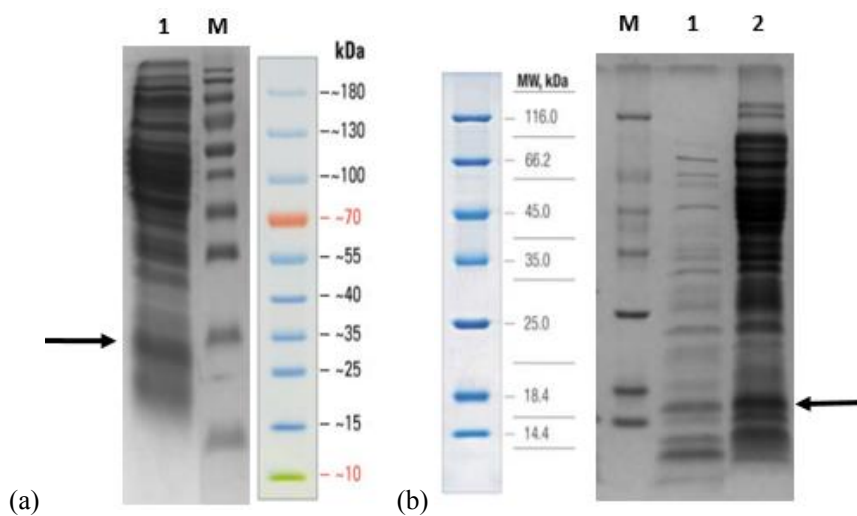


Figure 4 (a) Cell free extract of pure BL21(DE3) cells, without any transformation was done for control. The arrow shows the position where the protein for gene TVN0775 is present for recombinant plasmid containing cells in the cell lysate. M = marker SM Page Ruler Prestained protein Ladder; (b) the cell free extract of T7-Express competent cells containing *Tpv* Hsp 14.3 WT gene plasmid. M = marker unstained protein Ladder.

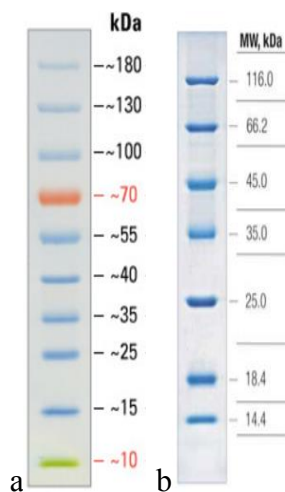


Figure 5: PageRuler™ Prestained Protein Ladder (Fermentas, Lithuania) (a) and PageRuler™ Unstained Protein Ladder (Fermentas, Lithuania) (b)

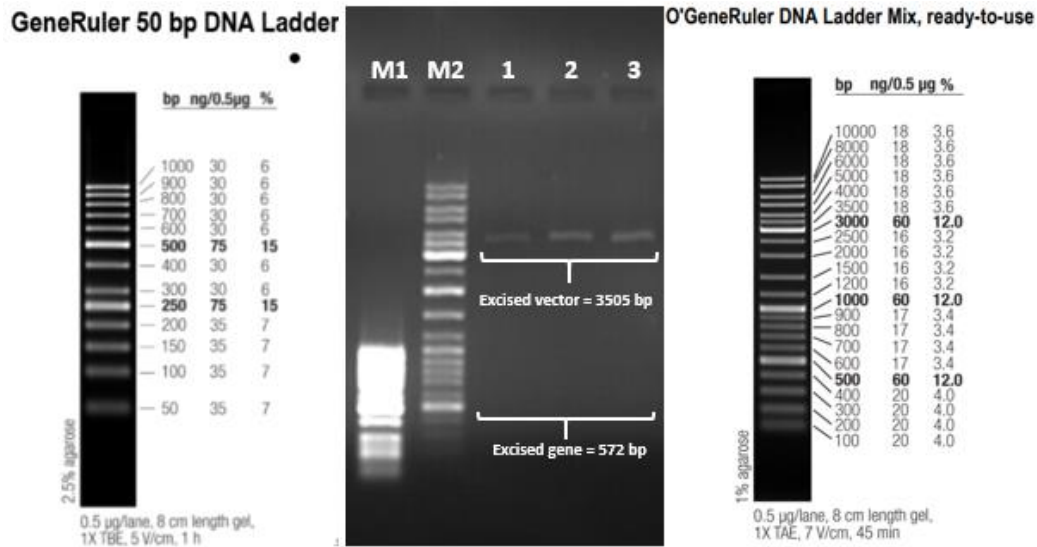


Figure 9: Isolation of possibly mutated plasmids of F8YV31G mutant, intended to be sent for sequencing. However, could not be sent due to low concentration.

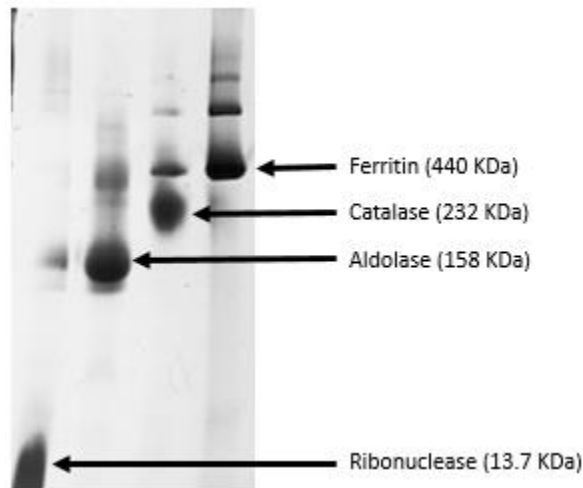


Figure 10: Resolution of pure proteins, Ferritin, Catalase, Aldolase and Ribonuclease, on Blue Native-PAGE, to be used for the generation of standard curve and calculation of unknown molecular weight of the oligomeric species of *Tpv* Hsp 14,3 WT and its various NTD mutants.

E. Tables

Table 1 Percentage alterations in nature of NTD following substitutions.

Name	Hydrophobic	Acidic	Basic	Neutral	Comment on hydrophobicity
WT	46.88%	6.25%	9.38%	37.5%	-
I5T	43.75%	6.25%	9.38%	40.63%	3.13% decrease in hydrophobicity
F8Y	43.75%	6.25%	9.38%	40.63%	3.13% decrease in hydrophobicity
I5TF8Y	42.42%	6.06%	9.09%	42.42%	4.46% decrease in hydrophobicity
F7SF8Y	40.63%	6.25%	9.38%	43.75%	6.25% decrease in hydrophobicity
I5TF7SF8Y	37.5%	6.25%	9.38%	46.88%	9.38% decrease in hydrophobicity
F8YV31G	40.63%	6.25%	9.38%	43.75%	6.25% decrease in hydrophobicity
E11V	50%	3.13%	9.38%	37.5%	3.12% increase in hydrophobicity 3.12% decrease in acidic nature
E22G	46.88%	3.13%	9.38%	40.63%	3.13% increase in neutrality 3.12% decrease in acidic nature
M12T	43.75%	6.25%	9.38%	40.63%	3.13% decrease in hydrophobicity
F26YI27T	40.63%	6.25%	9.38%	43.75%	6.25% decrease in hydrophobicity
V23GF26YI27T	37.5%	6.25%	9.38%	46.88%	9.38% decrease in hydrophobicity
V31G	43.75%	6.25%	9.38%	40.63%	3.13% decrease in hydrophobicity
V31IL33I	48.48%	6.06%	9.09%	36.36%	1.6% increase in hydrophobicity

Table 2 Intramolecular Hydrogen Bonds in *Tpv* sHSP 14.3 WT and I5T variant at position 5.

<i>Tpv</i> sHSP 14.3 WT					I5T variant				
Interacting AA	Interaction With	Category	Types	Distance	Interacting AA	Interaction With	Category	Types	Distance
A:THR3:C B	A:ILE5:O	Hydrogen Bond	Carbon Hydrogen Bond	3.602792	-				
-					B:PHE8:N	B:THR5:O	Hydrogen Bond	Conventional Hydrogen Bond	3.330606
B:THR3:C B	B:ILE5:O	Hydrogen Bond	Carbon Hydrogen Bond	3.602969	-				
-					A:PHE8:N	A:THR5:O	Hydrogen Bond	Conventional Hydrogen Bond	3.330588

Table 3 Inter and intramolecular bonds in *Tpv* sHSP 14.3 WT and F8Y variant at position 8.

<i>Tpv</i> sHSP 14.3 WT					F8Y variant				
Interacting AA	Interaction With	Category	Types	Distance	Interacting AA	Interaction With	Category	Types	Distance
Intramolecular Hydrophobic Interactions									
-					A:TYR8	A:ILE5	Hydrophobic	Pi-Alkyl	5.451190
A:PHE8	A:MET12	Hydrophobic	Pi-Alkyl	5.132089	-				
-					B:TYR8	B:ILE5	Hydrophobic	Pi-Alkyl	5.451036
B:PHE8 -	B:MET12	Hydrophobic	Pi-Alkyl	5.131935	-				
Intramolecular Electrostatic Interactions									
-					A:MET1:N	A:TYR8	Electrostatic	Pi-Cation	4.042867
A:GLU11:OE1	A:PHE8	Electrostatic	Pi-Anion	3.973179	-				
-					B:MET1:N	B:TYR8	Electrostatic	Pi-Cation	4.043074
B:GLU11:OE1	B:PHE8	Electrostatic	Pi-Anion	3.972929	-				
Intramolecular Hydrogen Bonds									
-					A:MET1:N	A:TYR8	Hydrogen Bond	Pi-Donor Hydrogen Bond	4.042867
A:GLU11:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.887148	-				
A:MET12:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186769	A:MET12:N	A:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.911169
-					B:MET1:N	B:TYR8	Hydrogen Bond	Pi-Donor Hydrogen Bond	4.043074
B:GLU11:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.887269	-				
B:MET12:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186399	B:MET12:N	B:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.910648

Table 4 Intramolecular bonds in *Tpv* sHSP 14.3 WT and I5TF8Y variant at positions 5 and 8.

<i>Tpv</i> sHSP 14.3 WT					I5TF8Y variant				
Interacting AA	Interaction With	Category	Types	Distance	Interacting AA	Interaction With	Category	Types	Distance
Intramolecular Hydrogen Bonds									
-					A:THR5:N	A:TYR2:O	Hydrogen Bond	Conventional Hydrogen Bond	3.398931
A:THR3:CB	A:ILE5:O	Hydrogen Bond	Carbon Hydrogen Bond	3.602792	-				
-					A:THR9:OG1	A:THR5:O	Hydrogen Bond	Conventional Hydrogen Bond	2.732859
-					A:THR9:N	A:THR5:O	Hydrogen Bond	Conventional Hydrogen Bond	3.270969
-					A:TYR8:N	A:THR5:OG1	Hydrogen Bond	Conventional Hydrogen Bond	2.715057
-					A:TYR8:CA	A:GLU11:OE1	Hydrogen Bond	Carbon Hydrogen Bond	3.755284
A:GLU11:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.887148	A:GLU11:N	A:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.016700
A:MET12:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186769	A:MET12:N	A:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.757703
-					B:THR5:N	B:TYR2:O	Hydrogen Bond	Conventional Hydrogen Bond	3.398113
B:THR3:CB	B:ILE5:O	Hydrogen Bond	Carbon Hydrogen Bond	3.602969	-				
-					B:THR9:OG1	B:THR5:O	Hydrogen Bond	Conventional Hydrogen Bond	2.732897
-					B:THR9:N	B:THR5:O	Hydrogen Bond	Conventional Hydrogen Bond	3.271326
-					B:TYR8:N	B:THR5:OG1	Hydrogen Bond	Conventional Hydrogen Bond	2.714510
-					B:TYR8:CA	B:GLU11:OE1	Hydrogen Bond	Carbon Hydrogen Bond	3.755171
B:GLU11:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.887269	B:GLU11:N	B:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.016584
B:MET12:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186399	B:MET12:N	B:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.757846
Intramolecular Electrostatic Interactions									
A:GLU11:OE1	A:PHE8	Electrostatic	Pi-Anion	3.973179	A:GLU11:OE1	A:TYR8	Electrostatic	Pi-Anion	3.838625
B:GLU11:OE1	B:PHE8	Electrostatic	Pi-Anion	3.972929	B:GLU11:OE1	B:TYR8	Electrostatic	Pi-Anion	3.838511
Intramolecular Hydrophobic Interactions									
A:PHE8	A:MET12	Hydrophobic	Pi-Alkyl	5.132089	-				
B:PHE8	B:MET12	Hydrophobic	Pi-Alkyl	5.131935	-				

Table 5 Intramolecular bonds in *Tpv* sHSP 14.3 WT and F7SF8Y variant at positions 7 and 8.

<i>Tpv</i> sHSP 14.3 WT					F7SF8Y variant				
Interacting AA	Interaction With	Category	Types	Distance	Interacting AA	Interaction With	Category	Types	Distance
Intramolecular Hydrophobic Interactions									
-					A:TYR8	A:LYS6	Hydrophobic	Pi-Alkyl	5.421456
A:PHE8	A:MET12	Hydrophobic	Pi-Alkyl	5.132089	A:TYR8	A:MET12	Hydrophobic	Pi-Alkyl	4.172348
-					B:TYR8	B:LYS6	Hydrophobic	Pi-Alkyl	5.421378
B:PHE8	B:MET12	Hydrophobic	Pi-Alkyl	5.131935	B:TYR8	B:MET12	Hydrophobic	Pi-Alkyl	4.172689
Intramolecular Electrostatic Interactions									
A:GLU11:OE1	A:PHE8	Electrostatic	Pi-Anion	3.973179	-				
B:GLU11:OE1	B:PHE8	Electrostatic	Pi-Anion	3.972929	-				
Intramolecular Hydrogen Bonds									
A:GLU11:N	A:PHE7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.952824	A:GLU11:N	A:SER7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.628570
A:ASN10:N	A:PHE7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.874697	A:ASN10:N	A:SER7:O	Hydrogen Bond	Conventional Hydrogen Bond	3.097215
A:GLU11:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.887148	-				
A:MET12:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186769	A:MET12:N	A:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.954456
B:ASN10:N	B:PHE7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.875026	B:ASN10:N	B:SER7:O	Hydrogen Bond	Conventional Hydrogen Bond	3.097724
B:GLU11:N	B:PHE7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.951727	B:GLU11:N	B:SER7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.629450
B:GLU11:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.887269	-				
B:MET12:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186399	B:MET12:N	B:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.954435

Table 6 Intramolecular bonds in *Tpv* sHSP 14.3 WT and I5TF7SF8Y variant at positions 5, 7 and 8.

<i>Tpv</i> sHSP 14.3 WT					I5TF7SF8Y variant				
Interacting AA	Interaction With	Category	Types	Distance	Interacting AA	Interaction With	Category	Types	Distance
Intramolecular Hydrophobic Interactions									
A:PHE8	A:MET12	Hydrophobic	Pi-Alkyl	5.132089	A:TYR8	A:MET12	Hydrophobic	Pi-Alkyl	4.956308
B:PHE8	B:MET12	Hydrophobic	Pi-Alkyl	5.131935	B:TYR8	B:MET12	Hydrophobic	Pi-Alkyl	4.956263
Intramolecular Hydrogen Bonds									
-					A:MET1:CA	A:THR5:OG1	Hydrogen Bond	Carbon Hydrogen Bond	2.978061
A:THR3:CB	A:ILE5:O	Hydrogen Bond	Carbon Hydrogen Bond	3.602792	-				
-					A:THR5:N	A:MET1:O	Hydrogen Bond	Conventional Hydrogen Bond	3.051020
A:ASN10:N	A:PHE7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.874697	A:ASN10:N	A:SER7:OG	Hydrogen Bond	Conventional Hydrogen Bond	2.922599
-					A:SER7:OG	A:ASN10:OD1	Hydrogen Bond	Conventional Hydrogen Bond	3.044098
A:GLU11:N	A:PHE7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.952824	A:GLU11:N	A:SER7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.796432
A:GLU11:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.887148	-				
A:MET12:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186769	A:MET12:N	A:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.885743
-					B:MET1:CA	B:THR5:OG1	Hydrogen Bond	Carbon Hydrogen Bond	2.977798
B:THR3:CB	B:ILE5:O	Hydrogen Bond	Carbon Hydrogen Bond	3.602969	-				
-					B:THR5:N	B:MET1:O	Hydrogen Bond	Conventional Hydrogen Bond	3.050620
B:ASN10:N	B:PHE7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.875026	B:ASN10:N	B:SER7:OG	Hydrogen Bond	Conventional Hydrogen Bond	2.922750
-					B:SER7:OG	B:ASN10:OD1	Hydrogen Bond	Conventional Hydrogen Bond	3.043819
B:GLU11:N	B:PHE7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.951727	B:GLU11:N	B:SER7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.796981
B:GLU11:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.887269	-				
B:MET12:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186399	B:MET12:N	B:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.885753

Table 7 Intramolecular bonds in *Tpv* sHSP 14.3 WT and E11V variant at position 11.

<i>Tpv</i> sHSP 14.3 WT					E11V variant				
Interacting AA	Interaction With	Category	Types	Distance	Interacting AA	Interaction With	Category	Types	Distance
Intramolecular Electrostatic Interactions									
A:GLU11:OE1	A:PHE8	Electrostatic	Pi-Anion	3.973179	-				
B:GLU11:OE1	B:PHE8	Electrostatic	Pi-Anion	3.972929	-				
Intramolecular Hydrogen Bonds									
A:GLU11:N	A:PHE7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.952824	A:VAL11:N	A:PHE7:O	Hydrogen Bond	Conventional Hydrogen Bond	3.046140
A:GLU11:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.887148	A:VAL11:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.150260
A:LYS14:N	A:GLU11:O	Hydrogen Bond	Conventional Hydrogen Bond	3.191906	A:LYS14:N	A:VAL11:O	Hydrogen Bond	Conventional Hydrogen Bond	3.181861
A:ASN15:N	A:GLU11:O	Hydrogen Bond	Conventional Hydrogen Bond	3.070697	A:ASN15:N	A:VAL11:O	Hydrogen Bond	Conventional Hydrogen Bond	3.160643
A:ASN15:ND2	A:GLU11:O	Hydrogen Bond	Conventional Hydrogen Bond	3.111119	A:ASN15:ND2	A:VAL11:O	Hydrogen Bond	Conventional Hydrogen Bond	3.234069
B:GLU11:N	B:PHE7:O	Hydrogen Bond	Conventional Hydrogen Bond	2.951727	B:VAL11:N	B:PHE7:O	Hydrogen Bond	Conventional Hydrogen Bond	3.045470
B:GLU11:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.887269	B:VAL11:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.149996
B:LYS14:N	B:GLU11:O	Hydrogen Bond	Conventional Hydrogen Bond	3.191682	B:LYS14:N	B:VAL11:O	Hydrogen Bond	Conventional Hydrogen Bond	3.181374
B:ASN15:N	B:GLU11:O	Hydrogen Bond	Conventional Hydrogen Bond	3.070294	B:ASN15:N	B:VAL11:O	Hydrogen Bond	Conventional Hydrogen Bond	3.160829
B:ASN15:ND2	B:GLU11:O	Hydrogen Bond	Conventional Hydrogen Bond	3.111758	B:ASN15:ND2	B:VAL11:O	Hydrogen Bond	Conventional Hydrogen Bond	3.233737

Table 8 Intramolecular hydrogen bonds in *Tpv* sHSP 14.3 WT and E22G variant at position 22.

<i>Tpv</i> sHSP 14.3 WT					E22G variant				
Interacting AA	Interaction With	Category	Types	Distance	Interacting AA	Interaction With	Category	Types	Distance
A:GLU2 2:N	A:ASN18:O	Hydrogen Bond	Conventional Hydrogen Bond	2.857062	A:GLY2 2:N	A:ASN18:O	Hydrogen Bond	Conventional Hydrogen Bond	2.752455
A:GLU2 2:N	A:THR19:O	Hydrogen Bond	Conventional Hydrogen Bond	2.879527	A:GLY2 2:N	A:THR19:O	Hydrogen Bond	Conventional Hydrogen Bond	3.039120
A:THR1 9:CA	A:GLU22:OE1	Hydrogen Bond	Carbon Hydrogen Bond	3.487750	A:GLY2 2:CA	A:SERT25:OG	Hydrogen Bond	Conventional Hydrogen Bond	3.460498
A:SER2 5:CB	A:GLU22:O	Hydrogen Bond	Carbon Hydrogen Bond	3.237537	-				
A:SER2 5:N	A:GLU22:O	Hydrogen Bond	Conventional Hydrogen Bond	3.300164	A:SER2 5:N	A:GLY22:O	Hydrogen Bond	Conventional Hydrogen Bond	2.933260
A:PHE2 6:N	A:GLU22:O	Hydrogen Bond	Conventional Hydrogen Bond	3.091639	A:PHE2 6:N	A:GLY22:O	Hydrogen Bond	Conventional Hydrogen Bond	3.012011
B:GLU2 2:N	B:ASN18:O	Hydrogen Bond	Conventional Hydrogen Bond	2.857256	B:GLY2 2:N	B:ASN18:O	Hydrogen Bond	Conventional Hydrogen Bond	2.752705
B:GLU2 2:N	B:THR19:O	Hydrogen Bond	Conventional Hydrogen Bond	2.879778	B:GLY2 2:N	B:THR19:O	Hydrogen Bond	Conventional Hydrogen Bond	3.039567
B:THR1 9:CA	B:GLU22:OE1	Hydrogen Bond	Carbon Hydrogen Bond	3.488520	B:GLY2 2:CA	B:SERT25:OG	Hydrogen Bond	Conventional Hydrogen Bond	3.459904
B:SER2 5:CB	B:GLU22:O	Hydrogen Bond	Carbon Hydrogen Bond	3.236773	-				
B:SER2 5:N	B:GLU22:O	Hydrogen Bond	Conventional Hydrogen Bond	3.299949	B:SER2 5:N	B:GLY22:O	Hydrogen Bond	Conventional Hydrogen Bond	2.933794
B:PHE2 6:N	B:GLU22:O	Hydrogen Bond	Conventional Hydrogen Bond	3.091763	B:PHE2 6:N	B:GLY22:O	Hydrogen Bond	Conventional Hydrogen Bond	3.011819

Table 9 Intramolecular bonds in *Tpv* sHSP 14.3 WT and M12T variant at position 12.

<i>Tpv</i> sHSP 14.3 WT					M12T variant				
Interacting AA	Interaction With	Category	Types	Distance	Interacting AA	Interaction With	Category	Types	Distance
Intramolecular Hydrophobic Interactions									
A:PHE8	A:MET12	Hydrophobic	Pi-Alkyl	5.132089	-				
B:PHE8	B:MET12	Hydrophobic	Pi-Alkyl	5.131935	-				
Intramolecular Hydrogen Bonds									
A:MET12:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186769	A:THR12:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.184312
A:MET12:N	A:THR9:O	Hydrogen Bond	Conventional Hydrogen Bond	2.996088	A:THR12:N	A:THR9:O	Hydrogen Bond	Conventional Hydrogen Bond	3.076693
A:ASN15:N	A:MET12:O	Hydrogen Bond	Conventional Hydrogen Bond	3.033410	A:ASN15:N	A:THR12:O	Hydrogen Bond	Conventional Hydrogen Bond	2.851917
A:VAL16:N	A:MET12:O	Hydrogen Bond	Conventional Hydrogen Bond	2.623762	A:VAL16:N	A:THR12:O	Hydrogen Bond	Conventional Hydrogen Bond	3.134791
B:MET12:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186399	B:THR12:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.184443
B:MET12:N	B:THR9:O	Hydrogen Bond	Conventional Hydrogen Bond	2.996095	B:THR12:N	B:THR9:O	Hydrogen Bond	Conventional Hydrogen Bond	3.077078
B:ASN15:N	B:MET12:O	Hydrogen Bond	Conventional Hydrogen Bond	3.033268	B:ASN15:N	B:THR12:O	Hydrogen Bond	Conventional Hydrogen Bond	2.852392
B:VAL16:N	B:MET12:O	Hydrogen Bond	Conventional Hydrogen Bond	2.624255	B:VAL16:N	B:THR12:O	Hydrogen Bond	Conventional Hydrogen Bond	3.134512

Table 10 Intramolecular bonds in *Tpv* sHSP 14.3 WT and F26YI27T variant at positions 26 and 27.

<i>Tpv</i> sHSP 14.3 WT					F26YI27T variant				
Interacting AA	Interaction With	Category	Types	Distance	Interacting AA	Interaction With	Category	Types	Distance
Intramolecular Hydrophobic Interactions									
-					A:TYR26	A:VAL23	Hydrophobic	Pi-Alkyl	4.884494
-					B:TYR26	B:VAL23	Hydrophobic	Pi-Alkyl	4.884392
Intermolecular Hydrophobic Interactions									
A:ILE27	B:ILE27	Hydrophobic	Alkyl	3.980927	-				
Intramolecular Hydrogen Bonds									
A:PHE26:N	A:GLU22:O	Hydrogen Bond	Conventional Hydrogen Bond	3.091639	A:TYR26:N	A:GLU22:O	Hydrogen Bond	Conventional Hydrogen Bond	3.073575
-					A:TYR26:N	A:VAL23:O	Hydrogen Bond	Conventional Hydrogen Bond	3.159980
A:LYS87:NZ	A:PHE26:O	Hydrogen Bond	Conventional Hydrogen Bond	2.932154	-				
-					A:THR27:N	A:VAL23:O	Hydrogen Bond	Conventional Hydrogen Bond	3.036949
-					A:THR27:OG1	A:VAL23:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186357
A:ILE27:N	A:SER24:O	Hydrogen Bond	Conventional Hydrogen Bond	3.238292	A:THR27:N	A:SER24:O	Hydrogen Bond	Conventional Hydrogen Bond	3.284529
A:LYS87:NZ	A:ILE27:O	Hydrogen Bond	Conventional Hydrogen Bond	2.709873	A:LYS87:CE	A:THR27:O	Hydrogen Bond	Carbon Hydrogen Bond	3.524348
-					A:THR27:OG1	A:TYR26	Hydrogen Bond	Pi-Donor Hydrogen Bond	3.463525
B:PHE26:N	B:GLU22:O	Hydrogen Bond	Conventional Hydrogen Bond	3.091763	B:TYR26:N	B:GLU22:O	Hydrogen Bond	Conventional Hydrogen Bond	3.073335
-					B:TYR26:N	B:VAL23:O	Hydrogen Bond	Conventional Hydrogen Bond	3.159093
B:LYS87:NZ	B:PHE26:O	Hydrogen Bond	Conventional Hydrogen Bond	2.932315	-				
-					B:THR27:N	B:VAL23:O	Hydrogen Bond	Conventional Hydrogen Bond	3.036913
-					B:THR27:OG1	B:VAL23:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186722
B:ILE27:N	B:SER24:O	Hydrogen Bond	Conventional Hydrogen Bond	3.238045	B:THR27:N	B:SER24:O	Hydrogen Bond	Conventional Hydrogen Bond	3.283899
B:LYS87:NZ	B:ILE27:O	Hydrogen Bond	Conventional Hydrogen Bond	2.710162	B:LYS87:CE	B:THR27:O	Hydrogen Bond	Carbon Hydrogen Bond	3.524574
-					B:THR27:OG1	B:TYR26	Hydrogen Bond	Pi-Donor Hydrogen Bond	3.463306

Table 11 Intramolecular bonds in *Tpv* sHSP 14.3 WT and V23GF26YI27T variant at positions 23, 26 and 27.

<i>Tpv</i> sHSP 14.3 WT					V23GF26YI27T variant				
Interacting AA	Interacting With	Category	Types	Distance	Interacting AA	Interacting With	Category	Types	Distance
Intermolecular Hydrophobic Interactions									
A:ILE27	B:ILE27	Hydrophobic	Alkyl	3.980927	-				
Intermolecular Hydrogen Bonds									
-					A:THR27:OG1	B:THR27:OG1	Hydrogen Bond	Conventional Hydrogen Bond	3.345167
Intramolecular Hydrogen Bonds									
A:VAL23:N	A:THR19:O	Hydrogen Bond	Conventional Hydrogen Bond	2.750696	A:GLY23:N	A:THR19:O	Hydrogen Bond	Conventional Hydrogen Bond	2.963906
A:VAL23:N	A:VAL20:O	Hydrogen Bond	Conventional Hydrogen Bond	3.152604	A:GLY23:N	A:VAL20:O	Hydrogen Bond	Conventional Hydrogen Bond	2.983810
A:PHE26:N	A:GLU22:O	Hydrogen Bond	Conventional Hydrogen Bond	3.091639	A:TYR26:N	A:GLU22:O	Hydrogen Bond	Conventional Hydrogen Bond	2.997255
-					A:TYR26:N	A:GLY23:O	Hydrogen Bond	Conventional Hydrogen Bond	3.117683
A:LYS87:NZ	A:PHE26:O	Hydrogen Bond	Conventional Hydrogen Bond	2.932154	-				
-					A:THR27:N	A:GLY23:O	Hydrogen Bond	Conventional Hydrogen Bond	3.247533
A:ILE27:N	A:SER24:O	Hydrogen Bond	Conventional Hydrogen Bond	3.238292	A:THR27:N	A:SER24:O	Hydrogen Bond	Conventional Hydrogen Bond	3.217646
A:LYS87:NZ	A:ILE27:O	Hydrogen Bond	Conventional Hydrogen Bond	2.709873	-				
B:VAL23:N	B:THR19:O	Hydrogen Bond	Conventional Hydrogen Bond	2.751612	B:GLY23:N	B:THR19:O	Hydrogen Bond	Conventional Hydrogen Bond	2.963718
B:VAL23:N	B:VAL20:O	Hydrogen Bond	Conventional Hydrogen Bond	3.152867	B:GLY23:N	B:VAL20:O	Hydrogen Bond	Conventional Hydrogen Bond	2.983944
B:PHE26:N	B:GLU22:O	Hydrogen Bond	Conventional Hydrogen Bond	3.091763	B:TYR26:N	B:GLU22:O	Hydrogen Bond	Conventional Hydrogen Bond	2.997320
-					B:TYR26:N	B:GLY23:O	Hydrogen Bond	Conventional Hydrogen Bond	3.117907
B:LYS87:NZ	B:PHE26:O	Hydrogen Bond	Conventional Hydrogen Bond	2.932315	-				
-					B:THR27:N	B:GLY23:O	Hydrogen Bond	Conventional Hydrogen Bond	3.247025
B:ILE27:N	B:SER24:O	Hydrogen Bond	Conventional Hydrogen Bond	3.238045	B:THR27:N	B:SER24:O	Hydrogen Bond	Conventional Hydrogen Bond	3.217285
B:LYS87:NZ	B:ILE27:O	Hydrogen Bond	Conventional Hydrogen Bond	2.710162	-				

Table 12: Intramolecular bonds in *Tpv* sHSP 14.3 WT and V31G variant at position 31.

<i>Tpv</i> sHSP 14.3 WT					V31G variant				
Interacting AA	Interaction With	Category	Types	Distance	Interacting AA	Interaction With	Category	Types	Distance
Intramolecular Hydrophobic Interactions									
A:ALA44	A:VAL31	Hydrophobic	Alkyl	4.400209	-				
B:ALA44	B:VAL31	Hydrophobic	Alkyl	4.400720	-				
Intermolecular Hydrogen Bonds									
-					B:GLN80:NE2	A:GLY31:O	Hydrogen Bond	Conventional Hydrogen Bond	2.746526
A:GLN80:NE2	B:VAL31:O	Hydrogen Bond	Conventional Hydrogen Bond	3.190572	A:GLN80:NE2	B:GLY31:O	Hydrogen Bond	Conventional Hydrogen Bond	2.882050
A:GLN80:N	B:VAL31:O	Hydrogen Bond	Conventional Hydrogen Bond	3.247263	-				

Table 13: Intramolecular bonds in *Tpv* sHSP 14.3 WT and V31IL33I variant at position 31 and 33.

<i>Tpv</i> sHSP 14.3 WT					V31IL33I variant				
Interacting AA	Interaction With	Category	Types	Distance	Interacting AA	Interaction With	Category	Types	Distance
Intramolecular Hydrophobic Interactions									
A:ALA44	A:VAL31	Hydrophobic	Alkyl	4.400209	A:ALA44	A:ILE31	Hydrophobic	Alkyl	5.102092
-					A:ILE31	A:LEU42	Hydrophobic	Alkyl	4.437880
A:LEU33	A:LEU40	Hydrophobic	Alkyl	4.310526	A:ILE33	A:LEU40	Hydrophobic	Alkyl	5.411559
A:LEU33	A:LEU42	Hydrophobic	Alkyl	4.673365	A:ILE33	A:LEU42	Hydrophobic	Alkyl	4.858525
A:PRO92	A:LEU33	Hydrophobic	Alkyl	5.008810	-				
B:ALA44	B:VAL31	Hydrophobic	Alkyl	4.400720	B:ALA44	B:ILE31	Hydrophobic	Alkyl	5.102319
-					B:ILE31	B:LEU42	Hydrophobic	Alkyl	4.438164
B:LEU33	B:LEU40	Hydrophobic	Alkyl	4.310476	B:ILE33	B:LEU40	Hydrophobic	Alkyl	5.411217
B:LEU33	B:LEU42	Hydrophobic	Alkyl	4.673178	B:ILE33	B:LEU42	Hydrophobic	Alkyl	4.858412
B:PRO92	B:LEU33	Hydrophobic	Alkyl	5.009041	-				
Intramolecular Hydrogen Bonds									
-					A:ILE33:CA	A:VAL41:O	Hydrogen Bond	Carbon Hydrogen Bond	3.265791
-					B:ILE33:CA	B:VAL41:O	Hydrogen Bond	Carbon Hydrogen Bond	3.265144
Intermolecular Hydrogen Bonds									
-					B:GLN80:N	A:ILE31:O	Hydrogen Bond	Conventional Hydrogen Bond	3.317943
-					B:GLN80:NE2	A:ILE31:O	Hydrogen Bond	Conventional Hydrogen Bond	2.419318
-					A:ILE33:N	B:ILE78:O	Hydrogen Bond	Conventional Hydrogen Bond	3.138515
A:GLN80:NE2	B:VAL31:O	Hydrogen Bond	Conventional Hydrogen Bond	3.190572	A:GLN80:N	B:ILE31:O	Hydrogen Bond	Conventional Hydrogen Bond	3.075022
A:GLN80:N	B:VAL31:O	Hydrogen Bond	Conventional Hydrogen Bond	3.247263	A:GLN80:NE2	B:ILE31:O	Hydrogen Bond	Conventional Hydrogen Bond	2.470822
-					B:ILE33:N	A:ILE78:O	Hydrogen Bond	Conventional Hydrogen Bond	2.857412

Table 14: Inter and intramolecular bonds in *Tpv* sHSP 14.3 WT and F8YV31G variant at positions 8 and 31.

<i>Tpv</i> sHSP 14.3 WT					F8YV31G variant				
Interacting AA	Interaction With	Category	Types	Distance	Interacting AA	Interaction With	Category	Types	Distance
Intramolecular Hydrophobic Interactions									
A:PHE8	A:MET12	Hydrophobic	Pi-Alkyl	5.132089	-				
A:ALA44	A:VAL31	Hydrophobic	Alkyl	4.400209	-				
B:PHE8	B:MET12	Hydrophobic	Pi-Alkyl	5.131935	-				
B:ALA44	B:VAL31	Hydrophobic	Alkyl	4.400720	-				
Intramolecular Electrostatic Interactions									
A:GLU11:OE1	A:PHE8	Electrostatic	Pi-Anion	3.973179	A:GLU11:OE2	A:TYR8	Electrostatic	Pi-Anion	4.309719
B:GLU11:OE1	B:PHE8	Electrostatic	Pi-Anion	3.972929	B:GLU11:OE2	B:TYR8	Electrostatic	Pi-Anion	4.310039
Intermolecular Hydrogen Bonds									
-					B:GLN80:N	A:GLY31:O	Hydrogen Bond	Conventional Hydrogen Bond	3.171077
A:GLN80:NE2	B:VAL31:O	Hydrogen Bond	Conventional Hydrogen Bond	3.190572	-				
A:GLN80:N	B:VAL31:O	Hydrogen Bond	Conventional Hydrogen Bond	3.247263	A:GLN80:N	B:GLY31:O	Hydrogen Bond	Conventional Hydrogen Bond	3.205830
Intramolecular Hydrogen Bonds									
A:GLU11:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.887148	A:GLU11:N	A:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.843748
A:MET12:N	A:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186769	A:MET12:N	A:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.237162
B:GLU11:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.887269	B:GLU11:N	B:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	2.843234
B:MET12:N	B:PHE8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.186399	B:MET12:N	B:TYR8:O	Hydrogen Bond	Conventional Hydrogen Bond	3.236218

Table 15: PONDR Score of *Tpv* sHSP 14.3 WT and its proximal NTD mutant variants

Residue Number	Amino Acid	WT	I5T	F8Y	I5TF8Y	F7SF8Y	I5TF7SF8Y
1	M	0.60256	0.79974	0.60256	0.79974	0.60256	0.79974
2	Y	0.17793	0.53636	0.17793	0.53636	0.66602	0.84008
3	T	0.02826	0.22811	0.02660	0.21693	0.28837	0.58491
4	P	0.05291	0.27203	0.05031	0.26180	0.32512	0.55968
5	I	0.13865	0.36799	0.13616	0.35954	0.40361	0.65063
6	K	0.07384	0.29285	0.07126	0.28375	0.35383	0.61556
7	F	0.05618	0.23536	0.05349	0.22616	0.29492	0.53745
8	F	0.05410	0.21109	0.05155	0.20309	0.27173	0.48136
9	T	0.05132	0.18401	0.04877	0.17683	0.23866	0.42241
10	N	0.06659	0.18400	0.06403	0.17739	0.23486	0.38655
11	E	0.05986	0.15347	0.05743	0.14775	0.19855	0.32026
12	M	0.05293	0.12262	0.05078	0.11804	0.16183	0.25315
13	I	0.04601	0.09044	0.04390	0.08680	0.12428	0.18485
14	K	0.02782	0.03966	0.02653	0.03767	0.06518	0.10286
15	N	0.02795	0.02821	0.02676	0.02687	0.05028	0.05135
16	V	0.03012	0.03037	0.02903	0.02913	0.03946	0.04039
17	S	0.03228	0.03252	0.03124	0.03132	0.03383	0.03471
18	N	0.03009	0.03028	0.02934	0.02941	0.03169	0.03238
19	T	0.01259	0.01270	0.01195	0.01199	0.01395	0.01435
20	V	0.01236	0.01236	0.01192	0.01192	0.01342	0.01342
21	K	0.01306	0.01306	0.01266	0.01266	0.01403	0.01403
22	E	0.01465	0.01465	0.01458	0.01458	0.01458	0.01458
23	V	0.01907	0.01907	0.01907	0.01907	0.01907	0.01907
24	S	0.02154	0.02154	0.02154	0.02154	0.02154	0.02154
25	S	0.02346	0.02346	0.02346	0.02346	0.02346	0.02346
26	F	0.02420	0.02420	0.02420	0.02420	0.02420	0.02420
27	I	0.02726	0.02726	0.02726	0.02726	0.02726	0.02726
28	Y	0.03601	0.03601	0.03601	0.03601	0.03601	0.03601
29	P	0.04521	0.04521	0.04521	0.04521	0.04521	0.04521
30	P	0.05699	0.05699	0.05699	0.05699	0.05699	0.05699
31	I	0.07131	0.07131	0.07131	0.07131	0.07131	0.07131
32	T	0.07977	0.07977	0.07977	0.07977	0.07977	0.07977
33	I	0.11662	0.11662	0.11662	0.11662	0.11662	0.11662
34	Y	0.16539	0.16539	0.16539	0.16539	0.16539	0.16539
35	Q	0.23794	0.23794	0.23794	0.23794	0.23794	0.23794
36	D	0.29837	0.29837	0.29837	0.29837	0.29837	0.29837
37	S	0.31692	0.31692	0.31692	0.31692	0.31692	0.31692
38	S	0.31761	0.31761	0.31761	0.31761	0.31761	0.31761
39	D	0.31491	0.31491	0.31491	0.31491	0.31491	0.31491
40	L	0.30852	0.30852	0.30852	0.30852	0.30852	0.30852
41	V	0.32925	0.32925	0.32925	0.32925	0.32925	0.32925
42	L	0.31518	0.31518	0.31518	0.31518	0.31518	0.31518
43	E	0.27968	0.27968	0.27968	0.27968	0.27968	0.27968
44	A	0.21516	0.21516	0.21516	0.21516	0.21516	0.21516
45	E	0.15964	0.15964	0.15964	0.15964	0.15964	0.15964
46	M	0.13206	0.13206	0.13206	0.13206	0.13206	0.13206
47	A	0.12245	0.12245	0.12245	0.12245	0.12245	0.12245
48	G	0.11291	0.11291	0.11291	0.11291	0.11291	0.11291
49	F	0.10308	0.10308	0.10308	0.10308	0.10308	0.10308
50	D	0.06950	0.06950	0.06950	0.06950	0.06950	0.06950

Table 15 (Continued)

Residue Number	Amino Acid	WT	I5T	F8Y	I5TF8Y	F7SF8Y	I5TF7SF8Y
51	K	0.04224	0.04224	0.04224	0.04224	0.04224	0.04224
52	K	0.02413	0.02413	0.02413	0.02413	0.02413	0.02413
53	N	0.01344	0.01344	0.01344	0.01344	0.01344	0.01344
54	I	0.00487	0.00487	0.00487	0.00487	0.00487	0.00487
55	K	0.01960	0.01960	0.01960	0.01960	0.01960	0.01960
56	V	0.11811	0.11811	0.11811	0.11811	0.11811	0.11811
57	S	0.22823	0.22823	0.22823	0.22823	0.22823	0.22823
58	V	0.33368	0.33368	0.33368	0.33368	0.33368	0.33368
59	N	0.40071	0.40071	0.40071	0.40071	0.40071	0.40071
60	K	0.47484	0.47484	0.47484	0.47484	0.47484	0.47484
61	N	0.55579	0.55579	0.55579	0.55579	0.55579	0.55579
62	V	0.62967	0.62967	0.62967	0.62967	0.62967	0.62967
63	L	0.68084	0.68084	0.68084	0.68084	0.68084	0.68084
64	T	0.70134	0.70134	0.70134	0.70134	0.70134	0.70134
65	I	0.64647	0.64647	0.64647	0.64647	0.64647	0.64647
66	S	0.58167	0.58167	0.58167	0.58167	0.58167	0.58167
67	A	0.53026	0.53026	0.53026	0.53026	0.53026	0.53026
68	E	0.51329	0.51329	0.51329	0.51329	0.51329	0.51329
69	R	0.48195	0.48195	0.48195	0.48195	0.48195	0.48195
70	K	0.48540	0.48540	0.48540	0.48540	0.48540	0.48540
71	R	0.49277	0.49277	0.49277	0.49277	0.49277	0.49277
72	E	0.46303	0.46303	0.46303	0.46303	0.46303	0.46303
73	Y	0.50231	0.50231	0.50231	0.50231	0.50231	0.50231
74	S	0.52915	0.52915	0.52915	0.52915	0.52915	0.52915
75	T	0.56774	0.56774	0.56774	0.56774	0.56774	0.56774
76	V	0.60580	0.60580	0.60580	0.60580	0.60580	0.60580
77	Y	0.56482	0.56482	0.56482	0.56482	0.56482	0.56482
78	I	0.52317	0.52317	0.52317	0.52317	0.52317	0.52317
79	D	0.43761	0.43761	0.43761	0.43761	0.43761	0.43761
80	Q	0.35715	0.35715	0.35715	0.35715	0.35715	0.35715
81	R	0.33637	0.33637	0.33637	0.33637	0.33637	0.33637
82	V	0.26427	0.26427	0.26427	0.26427	0.26427	0.26427
83	D	0.19675	0.19675	0.19675	0.19675	0.19675	0.19675
84	K	0.13916	0.13916	0.13916	0.13916	0.13916	0.13916
85	V	0.08045	0.08045	0.08045	0.08045	0.08045	0.08045
86	Y	0.09769	0.09769	0.09769	0.09769	0.09769	0.09769
87	K	0.12676	0.12676	0.12676	0.12676	0.12676	0.12676
88	V	0.12937	0.12937	0.12937	0.12937	0.12937	0.12937
89	V	0.12871	0.12871	0.12871	0.12871	0.12871	0.12871
90	K	0.12794	0.12794	0.12794	0.12794	0.12794	0.12794
91	L	0.12574	0.12574	0.12574	0.12574	0.12574	0.12574
92	P	0.12442	0.12442	0.12442	0.12442	0.12442	0.12442
93	V	0.10471	0.10471	0.10471	0.10471	0.10471	0.10471
94	E	0.09735	0.09735	0.09735	0.09735	0.09735	0.09735
95	I	0.09455	0.09455	0.09455	0.09455	0.09455	0.09455
96	E	0.09176	0.09176	0.09176	0.09176	0.09176	0.09176
97	Q	0.15081	0.15081	0.15081	0.15081	0.15081	0.15081
98	Q	0.23334	0.23334	0.23334	0.23334	0.23334	0.23334
99	D	0.31330	0.31330	0.31330	0.31330	0.31330	0.31330
100	I	0.37998	0.37998	0.37998	0.37998	0.37998	0.37998

Table 15 (Continued)

Residue Number	Amino Acid	WT	I5T	F8Y	I5TF8Y	F7SF8Y	I5TF7SF8Y
101	S	0.43782	0.43782	0.43782	0.43782	0.43782	0.43782
102	A	0.46400	0.46400	0.46400	0.46400	0.46400	0.46400
103	K	0.44956	0.44956	0.44956	0.44956	0.44956	0.44956
104	Y	0.43259	0.43259	0.43259	0.43259	0.43259	0.43259
105	S	0.41520	0.41520	0.41520	0.41520	0.41520	0.41520
106	E	0.36423	0.36423	0.36423	0.36423	0.36423	0.36423
107	G	0.29950	0.29950	0.29950	0.29950	0.29950	0.29950
108	I	0.23174	0.23174	0.23174	0.23174	0.23174	0.23174
109	L	0.17290	0.17290	0.17290	0.17290	0.17290	0.17290
110	T	0.13199	0.13199	0.13199	0.13199	0.13199	0.13199
111	V	0.11861	0.11861	0.11861	0.11861	0.11861	0.11861
112	R	0.13088	0.13088	0.13088	0.13088	0.13088	0.13088
113	M	0.15356	0.15356	0.15356	0.15356	0.15356	0.15356
114	K	0.16281	0.16281	0.16281	0.16281	0.16281	0.16281
115	T	0.18508	0.18508	0.18508	0.18508	0.18508	0.18508
116	K	0.19432	0.19432	0.19432	0.19432	0.19432	0.19432
117	N	0.21742	0.21742	0.21742	0.21742	0.21742	0.21742
118	I	0.23890	0.23890	0.23890	0.23890	0.23890	0.23890
119	K	0.24118	0.24118	0.24118	0.24118	0.24118	0.24118
120	N	0.26858	0.26858	0.26858	0.26858	0.26858	0.26858
121	V	0.31940	0.31940	0.31940	0.31940	0.31940	0.31940
122	E	0.26808	0.26808	0.26808	0.26808	0.26808	0.26808
123	I	0.19078	0.19078	0.19078	0.19078	0.19078	0.19078
124	E	0.42432	0.42432	0.42432	0.42432	0.42432	0.42432

Table 16: PONDR Score of *Tpv* sHSP 14.3 WT and its middle NTD mutant variants

Residue Number	Amino Acid	WT	M12T	E11V	E22G
1	M	0.60256	0.60256	0.60256	0.60256
2	Y	0.17793	0.17793	0.17793	0.17793
3	T	0.02826	0.02826	0.02826	0.02826
4	P	0.05291	0.05291	0.05291	0.05291
5	I	0.13865	0.12891	0.10804	0.13865
6	K	0.07384	0.06362	0.04169	0.07384
7	F	0.05618	0.04537	0.02242	0.05618
8	F	0.05410	0.04302	0.01959	0.05408
9	T	0.05132	0.03937	0.01441	0.05125
10	N	0.06659	0.05011	0.01579	0.06646
11	E	0.05986	0.04337	0.01432	0.05963
12	M	0.05293	0.03845	0.01270	0.05264
13	I	0.04601	0.03372	0.01102	0.04517
14	K	0.02782	0.02102	0.00656	0.02670
15	N	0.02795	0.02151	0.00643	0.02530
16	V	0.03012	0.02404	0.00683	0.02462
17	S	0.03228	0.02630	0.00720	0.02463
18	N	0.03009	0.02494	0.00741	0.02791
19	T	0.01259	0.01196	0.00379	0.01267
20	V	0.01236	0.01174	0.00370	0.01271
21	K	0.01306	0.01244	0.00448	0.01446
22	E	0.01465	0.01408	0.00665	0.01480
23	V	0.01907	0.01851	0.01136	0.01623
24	S	0.02154	0.02110	0.01562	0.01696
25	S	0.02346	0.02324	0.02091	0.01738
26	F	0.02420	0.02415	0.02420	0.01716
27	I	0.02726	0.02726	0.02726	0.01165
28	Y	0.03601	0.03601	0.03601	0.01217
29	P	0.04521	0.04521	0.04521	0.02119
30	P	0.05699	0.05699	0.05699	0.03198
31	I	0.07131	0.07131	0.07131	0.04812
32	T	0.07977	0.07977	0.07977	0.05984
33	I	0.11662	0.11662	0.11662	0.09996
34	Y	0.16539	0.16539	0.16539	0.15307
35	Q	0.23794	0.23794	0.23794	0.22876
36	D	0.29837	0.29837	0.29837	0.29233
37	S	0.31692	0.31692	0.31692	0.31692
38	S	0.31761	0.31761	0.31761	0.31761
39	D	0.31491	0.31491	0.31491	0.31491
40	L	0.30852	0.30852	0.30852	0.30852
41	V	0.32925	0.32925	0.32925	0.32925
42	L	0.31518	0.31518	0.31518	0.31518
43	E	0.27968	0.27968	0.27968	0.27968
44	A	0.21516	0.21516	0.21516	0.21516
45	E	0.15964	0.15964	0.15964	0.15964
46	M	0.13206	0.13206	0.13206	0.13206
47	A	0.12245	0.12245	0.12245	0.12245
48	G	0.11291	0.11291	0.11291	0.11291
49	F	0.10308	0.10308	0.10308	0.10308
50	D	0.06950	0.06950	0.06950	0.06950

Table 16 (Continued)

Residue Number	Amino Acid	WT	M12T	E11V	E22G
51	K	0.04224	0.04224	0.04224	0.04224
52	K	0.02413	0.02413	0.02413	0.02413
53	N	0.01344	0.01344	0.01344	0.01344
54	I	0.00487	0.00487	0.00487	0.00487
55	K	0.01960	0.01960	0.01960	0.01960
56	V	0.11811	0.11811	0.11811	0.11811
57	S	0.22823	0.22823	0.22823	0.22823
58	V	0.33368	0.33368	0.33368	0.33368
59	N	0.40071	0.40071	0.40071	0.40071
60	K	0.47484	0.47484	0.47484	0.47484
61	N	0.55579	0.55579	0.55579	0.55579
62	V	0.62967	0.62967	0.62967	0.62967
63	L	0.68084	0.68084	0.68084	0.68084
64	T	0.70134	0.70134	0.70134	0.70134
65	I	0.64647	0.64647	0.64647	0.64647
66	S	0.58167	0.58167	0.58167	0.58167
67	A	0.53026	0.53026	0.53026	0.53026
68	E	0.51329	0.51329	0.51329	0.51329
69	R	0.48195	0.48195	0.48195	0.48195
70	K	0.48540	0.48540	0.48540	0.48540
71	R	0.49277	0.49277	0.49277	0.49277
72	E	0.46303	0.46303	0.46303	0.46303
73	Y	0.50231	0.50231	0.50231	0.50231
74	S	0.52915	0.52915	0.52915	0.52915
75	T	0.56774	0.56774	0.56774	0.56774
76	V	0.60580	0.60580	0.60580	0.60580
77	Y	0.56482	0.56482	0.56482	0.56482
78	I	0.52317	0.52317	0.52317	0.52317
79	D	0.43761	0.43761	0.43761	0.43761
80	Q	0.35715	0.35715	0.35715	0.35715
81	R	0.33637	0.33637	0.33637	0.33637
82	V	0.26427	0.26427	0.26427	0.26427
83	D	0.19675	0.19675	0.19675	0.19675
84	K	0.13916	0.13916	0.13916	0.13916
85	V	0.08045	0.08045	0.08045	0.08045
86	Y	0.09769	0.09769	0.09769	0.09769
87	K	0.12676	0.12676	0.12676	0.12676
88	V	0.12937	0.12937	0.12937	0.12937
89	V	0.12871	0.12871	0.12871	0.12871
90	K	0.12794	0.12794	0.12794	0.12794
91	L	0.12574	0.12574	0.12574	0.12574
92	P	0.12442	0.12442	0.12442	0.12442
93	V	0.10471	0.10471	0.10471	0.10471
94	E	0.09735	0.09735	0.09735	0.09735
95	I	0.09455	0.09455	0.09455	0.09455
96	E	0.09176	0.09176	0.09176	0.09176
97	Q	0.15081	0.15081	0.15081	0.15081
98	Q	0.23334	0.23334	0.23334	0.23334
99	D	0.31330	0.31330	0.31330	0.31330
100	I	0.37998	0.37998	0.37998	0.37998

Table 16 (Continued)

Residue Number	Amino Acid	WT	M12T	E11V	E22G
101	S	0.43782	0.43782	0.43782	0.43782
102	A	0.46400	0.46400	0.46400	0.46400
103	K	0.44956	0.44956	0.44956	0.44956
104	Y	0.43259	0.43259	0.43259	0.43259
105	S	0.41520	0.41520	0.41520	0.41520
106	E	0.36423	0.36423	0.36423	0.36423
107	G	0.29950	0.29950	0.29950	0.29950
108	I	0.23174	0.23174	0.23174	0.23174
109	L	0.17290	0.17290	0.17290	0.17290
110	T	0.13199	0.13199	0.13199	0.13199
111	V	0.11861	0.11861	0.11861	0.11861
112	R	0.13088	0.13088	0.13088	0.13088
113	M	0.15356	0.15356	0.15356	0.15356
114	K	0.16281	0.16281	0.16281	0.16281
115	T	0.18508	0.18508	0.18508	0.18508
116	K	0.19432	0.19432	0.19432	0.19432
117	N	0.21742	0.21742	0.21742	0.21742
118	I	0.23890	0.23890	0.23890	0.23890
119	K	0.24118	0.24118	0.24118	0.24118
120	N	0.26858	0.26858	0.26858	0.26858
121	V	0.31940	0.31940	0.31940	0.31940
122	E	0.26808	0.26808	0.26808	0.26808
123	I	0.19078	0.19078	0.19078	0.19078
124	E	0.42432	0.42432	0.42432	0.42432

Table 17: PONDR Score of *Tpv* sHSP 14.3 WT and its distal NTD mutant variants

Residue Number	Amino Acid	WT	V23GF26YI27T	F26YI27T
1	M	0.60256	0.60256	0.60256
2	Y	0.17793	0.17793	0.17793
3	T	0.02826	0.02826	0.02826
4	P	0.05291	0.05291	0.05291
5	I	0.13865	0.13865	0.13865
6	K	0.07384	0.07384	0.07384
7	F	0.05618	0.05618	0.05618
8	F	0.05410	0.05410	0.05410
9	T	0.05132	0.05137	0.05132
10	N	0.06659	0.06671	0.06659
11	E	0.05986	0.06009	0.05986
12	M	0.05293	0.05314	0.05289
13	I	0.04601	0.04634	0.04583
14	K	0.02782	0.02835	0.02768
15	N	0.02795	0.03104	0.02833
16	V	0.03012	0.03987	0.03248
17	S	0.03228	0.04637	0.03576
18	N	0.03009	0.05855	0.03995
19	T	0.01259	0.04642	0.02470
20	V	0.01236	0.04916	0.02588
21	K	0.01306	0.08144	0.04188
22	E	0.01465	0.09733	0.05047
23	V	0.01907	0.13368	0.07189
24	S	0.02154	0.16572	0.09088
25	S	0.02346	0.16694	0.09248
26	F	0.02420	0.16677	0.09296
27	I	0.02726	0.15884	0.09049
28	Y	0.03601	0.16765	0.09647
29	P	0.04521	0.17926	0.10375
30	P	0.05699	0.15965	0.10044
31	I	0.07131	0.14855	0.09691
32	T	0.07977	0.11606	0.07950
33	I	0.11662	0.13102	0.10954
34	Y	0.16539	0.17384	0.15665
35	Q	0.23794	0.24295	0.22835
36	D	0.29837	0.29995	0.28793
37	S	0.31692	0.31301	0.30700
38	S	0.31761	0.30819	0.30819
39	D	0.31491	0.30532	0.30532
40	L	0.30852	0.30994	0.30994
41	V	0.32925	0.33949	0.33949
42	L	0.31518	0.31518	0.31518
43	E	0.27968	0.27968	0.27968
44	A	0.21516	0.21516	0.21516
45	E	0.15964	0.15964	0.15964
46	M	0.13206	0.13206	0.13206
47	A	0.12245	0.12245	0.12245
48	G	0.11291	0.11291	0.11291
49	F	0.10308	0.10308	0.10308
50	D	0.06950	0.06950	0.06950

Table 17 (Continued)

Residue Number	Amino Acid	WT	V23GF26YI27T	F26YI27T
51	K	0.04224	0.04224	0.04224
52	K	0.02413	0.02413	0.02413
53	N	0.01344	0.01344	0.01344
54	I	0.00487	0.00487	0.00487
55	K	0.01960	0.01960	0.01960
56	V	0.11811	0.11811	0.11811
57	S	0.22823	0.22823	0.22823
58	V	0.33368	0.33368	0.33368
59	N	0.40071	0.40071	0.40071
60	K	0.47484	0.47484	0.47484
61	N	0.55579	0.55579	0.55579
62	V	0.62967	0.62967	0.62967
63	L	0.68084	0.68084	0.68084
64	T	0.70134	0.70134	0.70134
65	I	0.64647	0.64647	0.64647
66	S	0.58167	0.58167	0.58167
67	A	0.53026	0.53026	0.53026
68	E	0.51329	0.51329	0.51329
69	R	0.48195	0.48195	0.48195
70	K	0.48540	0.48540	0.48540
71	R	0.49277	0.49277	0.49277
72	E	0.46303	0.46303	0.46303
73	Y	0.50231	0.50231	0.50231
74	S	0.52915	0.52915	0.52915
75	T	0.56774	0.56774	0.56774
76	V	0.60580	0.60580	0.60580
77	Y	0.56482	0.56482	0.56482
78	I	0.52317	0.52317	0.52317
79	D	0.43761	0.43761	0.43761
80	Q	0.35715	0.35715	0.35715
81	R	0.33637	0.33637	0.33637
82	V	0.26427	0.26427	0.26427
83	D	0.19675	0.19675	0.19675
84	K	0.13916	0.13916	0.13916
85	V	0.08045	0.08045	0.08045
86	Y	0.09769	0.09769	0.09769
87	K	0.12676	0.12676	0.12676
88	V	0.12937	0.12937	0.12937
89	V	0.12871	0.12871	0.12871
90	K	0.12794	0.12794	0.12794
91	L	0.12574	0.12574	0.12574
92	P	0.12442	0.12442	0.12442
93	V	0.10471	0.10471	0.10471
94	E	0.09735	0.09735	0.09735
95	I	0.09455	0.09455	0.09455
96	E	0.09176	0.09176	0.09176
97	Q	0.15081	0.15081	0.15081
98	Q	0.23334	0.23334	0.23334
99	D	0.31330	0.31330	0.31330
100	I	0.37998	0.37998	0.37998

Table 17 (Continued)

Residue Number	Amino Acid	WT	V23GF26YI27T	F26YI27T
101	S	0.43782	0.43782	0.43782
102	A	0.46400	0.46400	0.46400
103	K	0.44956	0.44956	0.44956
104	Y	0.43259	0.43259	0.43259
105	S	0.41520	0.41520	0.41520
106	E	0.36423	0.36423	0.36423
107	G	0.29950	0.29950	0.29950
108	I	0.23174	0.23174	0.23174
109	L	0.17290	0.17290	0.17290
110	T	0.13199	0.13199	0.13199
111	V	0.11861	0.11861	0.11861
112	R	0.13088	0.13088	0.13088
113	M	0.15356	0.15356	0.15356
114	K	0.16281	0.16281	0.16281
115	T	0.18508	0.18508	0.18508
116	K	0.19432	0.19432	0.19432
117	N	0.21742	0.21742	0.21742
118	I	0.23890	0.23890	0.23890
119	K	0.24118	0.24118	0.24118
120	N	0.26858	0.26858	0.26858
121	V	0.31940	0.31940	0.31940
122	E	0.26808	0.26808	0.26808
123	I	0.19078	0.19078	0.19078
124	E	0.42432	0.42432	0.42432

Table 18: PONDR Score of *Tpv* sHSP 14.3 WT and its distal NTD mutant variants

Residue Number	Amino Acid	WT	V31G	V31IL33I	F8YV31G
1	M	0.60256	0.60256	0.60256	0.60256
2	Y	0.17793	0.17793	0.17793	0.17793
3	T	0.02826	0.02826	0.02826	0.02660
4	P	0.05291	0.05291	0.05291	0.05031
5	I	0.13865	0.13865	0.13865	0.13616
6	K	0.07384	0.07384	0.07384	0.07126
7	F	0.05618	0.05618	0.05618	0.05349
8	F	0.05410	0.05410	0.05410	0.05155
9	T	0.05132	0.05132	0.05132	0.04877
10	N	0.06659	0.06659	0.06659	0.06403
11	E	0.05986	0.05986	0.05986	0.05743
12	M	0.05293	0.05293	0.05293	0.05078
13	I	0.04601	0.04601	0.04601	0.04390
14	K	0.02782	0.02782	0.02782	0.02653
15	N	0.02795	0.02795	0.02795	0.02676
16	V	0.03012	0.03012	0.03012	0.02903
17	S	0.03228	0.03399	0.03181	0.03295
18	N	0.03009	0.03235	0.02949	0.03159
19	T	0.01259	0.01512	0.01193	0.01449
20	V	0.01236	0.01494	0.01169	0.01450
21	K	0.01306	0.01636	0.01226	0.01596
22	E	0.01465	0.01946	0.01356	0.01939
23	V	0.01907	0.02689	0.01737	0.02689
24	S	0.02154	0.03236	0.01923	0.03236
25	S	0.02346	0.03705	0.02054	0.03705
26	F	0.02420	0.03799	0.02134	0.03799
27	I	0.02726	0.04241	0.02412	0.04241
28	Y	0.03601	0.05573	0.03186	0.05573
29	P	0.04521	0.06974	0.04001	0.06974
30	P	0.05699	0.08687	0.05056	0.08687
31	I	0.07131	0.10583	0.06372	0.10583
32	T	0.07977	0.11661	0.07155	0.11661
33	I	0.11662	0.16024	0.10639	0.16024
34	Y	0.16539	0.21586	0.15310	0.21586
35	Q	0.23794	0.29218	0.22435	0.29218
36	D	0.29837	0.36028	0.28238	0.36028
37	S	0.31692	0.38298	0.29974	0.38298
38	S	0.31761	0.37880	0.30177	0.37880
39	D	0.31491	0.37004	0.30071	0.37004
40	L	0.30852	0.35750	0.29578	0.35750
41	V	0.32925	0.37291	0.31775	0.37291
42	L	0.31518	0.34905	0.30629	0.34905
43	E	0.27968	0.30394	0.27347	0.30394
44	A	0.21516	0.23373	0.21066	0.23373
45	E	0.15964	0.16864	0.15796	0.16864
46	M	0.13206	0.13206	0.13263	0.13206
47	A	0.12245	0.12245	0.12273	0.12245
48	G	0.11291	0.11291	0.11291	0.11291
49	F	0.10308	0.10308	0.10308	0.10308
50	D	0.06950	0.06950	0.06950	0.06950

Table 18 (Continued)

Residue Number	Amino Acid	WT	V31G	V31IL33I	F8YV31G
51	K	0.04224	0.04224	0.04224	0.04224
52	K	0.02413	0.02413	0.02413	0.02413
53	N	0.01344	0.01344	0.01344	0.01344
54	I	0.00487	0.00487	0.00487	0.00487
55	K	0.01960	0.01960	0.01960	0.01960
56	V	0.11811	0.11811	0.11811	0.11811
57	S	0.22823	0.22823	0.22823	0.22823
58	V	0.33368	0.33368	0.33368	0.33368
59	N	0.40071	0.40071	0.40071	0.40071
60	K	0.47484	0.47484	0.47484	0.47484
61	N	0.55579	0.55579	0.55579	0.55579
62	V	0.62967	0.62967	0.62967	0.62967
63	L	0.68084	0.68084	0.68084	0.68084
64	T	0.70134	0.70134	0.70134	0.70134
65	I	0.64647	0.64647	0.64647	0.64647
66	S	0.58167	0.58167	0.58167	0.58167
67	A	0.53026	0.53026	0.53026	0.53026
68	E	0.51329	0.51329	0.51329	0.51329
69	R	0.48195	0.48195	0.48195	0.48195
70	K	0.48540	0.48540	0.48540	0.48540
71	R	0.49277	0.49277	0.49277	0.49277
72	E	0.46303	0.46303	0.46303	0.46303
73	Y	0.50231	0.50231	0.50231	0.50231
74	S	0.52915	0.52915	0.52915	0.52915
75	T	0.56774	0.56774	0.56774	0.56774
76	V	0.60580	0.60580	0.60580	0.60580
77	Y	0.56482	0.56482	0.56482	0.56482
78	I	0.52317	0.52317	0.52317	0.52317
79	D	0.43761	0.43761	0.43761	0.43761
80	Q	0.35715	0.35715	0.35715	0.35715
81	R	0.33637	0.33637	0.33637	0.33637
82	V	0.26427	0.26427	0.26427	0.26427
83	D	0.19675	0.19675	0.19675	0.19675
84	K	0.13916	0.13916	0.13916	0.13916
85	V	0.08045	0.08045	0.08045	0.08045
86	Y	0.09769	0.09769	0.09769	0.09769
87	K	0.12676	0.12676	0.12676	0.12676
88	V	0.12937	0.12937	0.12937	0.12937
89	V	0.12871	0.12871	0.12871	0.12871
90	K	0.12794	0.12794	0.12794	0.12794
91	L	0.12574	0.12574	0.12574	0.12574
92	P	0.12442	0.12442	0.12442	0.12442
93	V	0.10471	0.10471	0.10471	0.10471
94	E	0.09735	0.09735	0.09735	0.09735
95	I	0.09455	0.09455	0.09455	0.09455
96	E	0.09176	0.09176	0.09176	0.09176
97	Q	0.15081	0.15081	0.15081	0.15081
98	Q	0.23334	0.23334	0.23334	0.23334
99	D	0.31330	0.31330	0.31330	0.31330
100	I	0.37998	0.37998	0.37998	0.37998

Table 18 (Continued)

Residue Number	Amino Acid	WT	V31G	V31IL33I	F8YV31G
101	S	0.43782	0.43782	0.43782	0.43782
102	A	0.46400	0.46400	0.46400	0.46400
103	K	0.44956	0.44956	0.44956	0.44956
104	Y	0.43259	0.43259	0.43259	0.43259
105	S	0.41520	0.41520	0.41520	0.41520
106	E	0.36423	0.36423	0.36423	0.36423
107	G	0.29950	0.29950	0.29950	0.29950
108	I	0.23174	0.23174	0.23174	0.23174
109	L	0.17290	0.17290	0.17290	0.17290
110	T	0.13199	0.13199	0.13199	0.13199
111	V	0.11861	0.11861	0.11861	0.11861
112	R	0.13088	0.13088	0.13088	0.13088
113	M	0.15356	0.15356	0.15356	0.15356
114	K	0.16281	0.16281	0.16281	0.16281
115	T	0.18508	0.18508	0.18508	0.18508
116	K	0.19432	0.19432	0.19432	0.19432
117	N	0.21742	0.21742	0.21742	0.21742
118	I	0.23890	0.23890	0.23890	0.23890
119	K	0.24118	0.24118	0.24118	0.24118
120	N	0.26858	0.26858	0.26858	0.26858
121	V	0.31940	0.31940	0.31940	0.31940
122	E	0.26808	0.26808	0.26808	0.26808
123	I	0.19078	0.19078	0.19078	0.19078
124	E	0.42432	0.42432	0.42432	0.42432

Table 19: Disorder Probability of the residue for WT and NTD proximal mutant variants

Residue number	Amino acid	Disorder probability of the residue (WT)	Disorder probability of the residue (I5T)	Disorder probability of the residue (F8Y)	Disorder probability of the residue (I5TF8Y)	Disorder probability of the residue (F7SF8Y)	Disorder probability of the residue (I5TF7SF8Y)
1	M	0.66	0.73	0.69	0.75	0.72	0.79
2	Y	0.59	0.68	0.64	0.71	0.67	0.71
3	T	0.48	0.63	0.57	0.64	0.61	0.65
4	P	0.44	0.59	0.52	0.59	0.61	0.65
5	I	0.42	0.58	0.48	0.58	0.58	0.66
6	K	0.44	0.59	0.53	0.57	0.59	0.65
7	F	0.43	0.56	0.48	0.53	0.57	0.63
8	F	0.45	0.57	0.51	0.53	0.58	0.64
9	T	0.46	0.58	0.52	0.53	0.58	0.64
10	N	0.46	0.57	0.49	0.51	0.56	0.61
11	E	0.45	0.54	0.46	0.47	0.53	0.57
12	M	0.44	0.53	0.45	0.46	0.49	0.54
13	I	0.43	0.52	0.43	0.45	0.47	0.52
14	K	0.43	0.51	0.43	0.45	0.47	0.51
15	N	0.43	0.51	0.43	0.45	0.48	0.51
16	V	0.44	0.49	0.43	0.46	0.48	0.51
17	S	0.45	0.52	0.44	0.47	0.52	0.52
18	N	0.46	0.53	0.45	0.48	0.53	0.53
19	T	0.49	0.55	0.48	0.52	0.56	0.55
20	V	0.53	0.57	0.52	0.54	0.58	0.57
21	K	0.53	0.56	0.52	0.53	0.57	0.56
22	E	0.49	0.52	0.47	0.48	0.54	0.53
23	V	0.45	0.46	0.43	0.44	0.47	0.46
24	S	0.4	0.41	0.38	0.39	0.42	0.41
25	S	0.34	0.35	0.32	0.33	0.35	0.35
26	F	0.28	0.29	0.26	0.27	0.29	0.28
27	I	0.22	0.22	0.2	0.21	0.22	0.22
28	Y	0.16	0.16	0.14	0.15	0.16	0.16
29	P	0.12	0.12	0.1	0.11	0.12	0.12
30	P	0.09	0.1	0.1	0.1	0.1	0.1
31	V	0.1	0.1	0.1	0.1	0.1	0.1
32	T	0.1	0.09	0.1	0.1	0.1	0.1
33	L	0.12	0.11	0.12	0.11	0.11	0.11
34	Y	0.15	0.13	0.14	0.13	0.14	0.14
35	Q	0.16	0.15	0.16	0.14	0.15	0.15
36	D	0.16	0.14	0.16	0.14	0.14	0.14
37	S	0.14	0.13	0.15	0.13	0.13	0.13
38	S	0.12	0.11	0.13	0.12	0.11	0.11
39	D	0.1	0.1	0.1	0.1	0.1	0.1
40	L	0.09	0.09	0.09	0.09	0.09	0.09
41	V	0.08	0.08	0.08	0.08	0.07	0.08
42	L	0.07	0.07	0.07	0.07	0.06	0.07
43	E	0.06	0.06	0.06	0.06	0.06	0.06
44	A	0.05	0.05	0.05	0.05	0.05	0.05
45	E	0.04	0.05	0.05	0.05	0.04	0.05

Table 19 (Continued)

Residue number	Amino acid	Disorder probability of the residue (WT)	Disorder probability of the residue (I5T)	Disorder probability of the residue (F8Y)	Disorder probability of the residue (I5TF8Y)	Disorder probability of the residue (F7SF8Y)	Disorder probability of the residue (I5TF7SF8Y)
46	M	0.04	0.05	0.05	0.05	0.04	0.05
47	A	0.05	0.05	0.05	0.05	0.04	0.05
48	G	0.05	0.05	0.05	0.05	0.05	0.05
49	F	0.05	0.05	0.05	0.06	0.05	0.05
50	D	0.06	0.06	0.06	0.06	0.06	0.06
51	K	0.06	0.07	0.07	0.07	0.06	0.06
52	K	0.07	0.07	0.07	0.07	0.07	0.07
53	N	0.08	0.08	0.08	0.07	0.07	0.07
54	I	0.08	0.08	0.08	0.08	0.08	0.08
55	K	0.09	0.09	0.09	0.09	0.09	0.09
56	V	0.09	0.09	0.09	0.09	0.09	0.09
57	S	0.09	0.09	0.09	0.09	0.09	0.09
58	V	0.1	0.1	0.1	0.1	0.09	0.1
59	N	0.1	0.1	0.1	0.1	0.1	0.09
60	K	0.11	0.11	0.1	0.11	0.09	0.1
61	N	0.12	0.11	0.11	0.11	0.1	0.11
62	V	0.11	0.1	0.1	0.1	0.09	0.1
63	L	0.11	0.1	0.1	0.1	0.09	0.1
64	T	0.12	0.1	0.1	0.1	0.1	0.1
65	I	0.14	0.12	0.12	0.13	0.12	0.12
66	S	0.19	0.17	0.17	0.17	0.17	0.17
67	A	0.24	0.23	0.23	0.23	0.23	0.23
68	E	0.3	0.29	0.29	0.29	0.29	0.29
69	R	0.36	0.35	0.35	0.35	0.35	0.35
70	K	0.42	0.42	0.42	0.42	0.42	0.41
71	R	0.47	0.47	0.47	0.47	0.47	0.46
72	E	0.53	0.53	0.53	0.52	0.53	0.51
73	Y	0.54	0.53	0.53	0.53	0.53	0.52
74	S	0.54	0.53	0.53	0.52	0.53	0.51
75	T	0.54	0.53	0.53	0.52	0.53	0.49
76	V	0.54	0.53	0.53	0.52	0.53	0.51
77	Y	0.56	0.55	0.55	0.54	0.55	0.53
78	I	0.58	0.57	0.57	0.55	0.56	0.54
79	D	0.59	0.58	0.58	0.57	0.57	0.55
80	Q	0.58	0.56	0.57	0.55	0.55	0.54
81	R	0.56	0.55	0.55	0.53	0.53	0.52
82	V	0.53	0.52	0.52	0.48	0.49	0.48
83	D	0.47	0.47	0.47	0.45	0.45	0.44
84	K	0.42	0.42	0.42	0.41	0.41	0.4
85	V	0.37	0.37	0.37	0.35	0.36	0.35
86	Y	0.33	0.33	0.33	0.32	0.32	0.32
87	K	0.29	0.3	0.29	0.29	0.29	0.29
88	V	0.27	0.27	0.27	0.26	0.26	0.26
89	V	0.24	0.25	0.24	0.24	0.24	0.24
90	K	0.22	0.23	0.22	0.21	0.22	0.22

Table 19 (Continued)

Residue number	Amino acid	Disorder probability of the residue (WT)	Disorder probability of the residue (I5T)	Disorder probability of the residue (F8Y)	Disorder probability of the residue (I5TF8Y)	Disorder probability of the residue (F7SF8Y)	Disorder probability of the residue (I5TF7SF8Y)
91	L	0.21	0.21	0.21	0.2	0.21	0.21
92	P	0.2	0.2	0.2	0.19	0.2	0.2
93	V	0.2	0.21	0.2	0.19	0.2	0.2
94	E	0.21	0.22	0.21	0.2	0.21	0.21
95	I	0.21	0.22	0.22	0.2	0.22	0.21
96	E	0.21	0.22	0.21	0.2	0.22	0.21
97	Q	0.21	0.21	0.21	0.2	0.21	0.21
98	Q	0.21	0.2	0.2	0.19	0.21	0.2
99	D	0.19	0.18	0.19	0.17	0.19	0.18
100	I	0.16	0.15	0.16	0.14	0.16	0.15
101	S	0.12	0.11	0.12	0.11	0.12	0.11
102	A	0.1	0.09	0.1	0.09	0.1	0.1
103	K	0.09	0.08	0.09	0.08	0.09	0.09
104	Y	0.08	0.07	0.08	0.07	0.08	0.08
105	S	0.07	0.06	0.07	0.06	0.07	0.07
106	E	0.06	0.06	0.06	0.06	0.06	0.06
107	G	0.05	0.05	0.05	0.05	0.05	0.05
108	I	0.05	0.05	0.05	0.05	0.05	0.05
109	L	0.05	0.05	0.05	0.05	0.06	0.06
110	T	0.06	0.06	0.06	0.06	0.07	0.06
111	V	0.07	0.07	0.07	0.08	0.08	0.08
112	R	0.09	0.09	0.09	0.09	0.09	0.09
113	M	0.13	0.13	0.13	0.14	0.14	0.14
114	K	0.21	0.22	0.22	0.23	0.22	0.22
115	T	0.3	0.3	0.3	0.31	0.3	0.31
116	K	0.39	0.39	0.39	0.4	0.39	0.39
117	N	0.46	0.46	0.46	0.47	0.46	0.46
118	I	0.55	0.55	0.55	0.56	0.54	0.55
119	K	0.64	0.64	0.64	0.65	0.64	0.64
120	N	0.71	0.71	0.71	0.72	0.7	0.7
121	V	0.69	0.69	0.69	0.7	0.68	0.69
122	E	0.73	0.73	0.73	0.73	0.71	0.72
123	I	0.71	0.71	0.71	0.7	0.69	0.7
124	E	0.8	0.8	0.8	0.8	0.78	0.79

Table 20: Disorder Probability of the residue for WT and NTD middle mutant variants

Residue number	Amino acid	Disorder probability of the residue (WT)	Disorder probability of the residue (E11V)	Disorder probability of the residue (M12T)	Disorder probability of the residue (E22G)
1	M	0.66	0.69	0.73	0.66
2	Y	0.59	0.63	0.66	0.58
3	T	0.48	0.52	0.56	0.51
4	P	0.44	0.48	0.53	0.46
5	I	0.42	0.45	0.48	0.44
6	K	0.44	0.49	0.52	0.46
7	F	0.43	0.46	0.52	0.44
8	F	0.45	0.48	0.55	0.47
9	T	0.46	0.51	0.57	0.48
10	N	0.46	0.51	0.58	0.48
11	E	0.45	0.48	0.56	0.47
12	M	0.44	0.46	0.55	0.47
13	I	0.43	0.46	0.53	0.46
14	K	0.43	0.45	0.51	0.46
15	N	0.43	0.45	0.48	0.45
16	V	0.44	0.45	0.47	0.45
17	S	0.45	0.45	0.47	0.46
18	N	0.46	0.45	0.47	0.47
19	T	0.49	0.47	0.48	0.48
20	V	0.53	0.51	0.51	0.53
21	K	0.53	0.51	0.48	0.52
22	E	0.49	0.47	0.46	0.48
23	V	0.45	0.43	0.42	0.43
24	S	0.4	0.39	0.38	0.38
25	S	0.34	0.33	0.33	0.33
26	F	0.28	0.27	0.28	0.27
27	I	0.22	0.21	0.22	0.20
28	Y	0.16	0.15	0.16	0.15
29	P	0.12	0.11	0.12	0.10
30	P	0.09	0.10	0.10	0.10
31	V	0.1	0.10	0.09	0.10
32	T	0.1	0.09	0.10	0.10
33	L	0.12	0.11	0.13	0.12
34	Y	0.15	0.14	0.15	0.15
35	Q	0.16	0.15	0.16	0.16
36	D	0.16	0.14	0.16	0.16
37	S	0.14	0.13	0.14	0.15
38	S	0.12	0.11	0.13	0.13
39	D	0.1	0.10	0.09	0.09
40	L	0.09	0.09	0.09	0.09
41	V	0.08	0.07	0.08	0.08
42	L	0.07	0.06	0.07	0.07
43	E	0.06	0.06	0.06	0.06
44	A	0.05	0.05	0.05	0.05
45	E	0.04	0.04	0.05	0.04
46	M	0.04	0.04	0.05	0.04
47	A	0.05	0.04	0.05	0.04
48	G	0.05	0.04	0.05	0.04
49	F	0.05	0.05	0.05	0.05
50	D	0.06	0.06	0.06	0.05

Table 20 (Continued)

Residue number	Amino acid	Disorder probability of the residue (WT)	Disorder probability of the residue (E11V)	Disorder probability of the residue (M12T)	Disorder probability of the residue (E22G)
51	K	0.06	0.06	0.07	0.06
52	K	0.07	0.07	0.07	0.06
53	N	0.08	0.07	0.08	0.07
54	I	0.08	0.08	0.08	0.08
55	K	0.09	0.09	0.09	0.08
56	V	0.09	0.09	0.09	0.08
57	S	0.09	0.09	0.09	0.09
58	V	0.1	0.10	0.10	0.09
59	N	0.1	0.09	0.09	0.10
60	K	0.11	0.11	0.10	0.09
61	N	0.12	0.11	0.10	0.10
62	V	0.11	0.11	0.09	0.10
63	L	0.11	0.11	0.09	0.10
64	T	0.12	0.11	0.10	0.10
65	I	0.14	0.13	0.12	0.10
66	S	0.19	0.18	0.18	0.15
67	A	0.24	0.24	0.24	0.20
68	E	0.3	0.30	0.30	0.26
69	R	0.36	0.36	0.36	0.32
70	K	0.42	0.43	0.43	0.38
71	R	0.47	0.48	0.47	0.44
72	E	0.53	0.55	0.53	0.48
73	Y	0.54	0.56	0.53	0.49
74	S	0.54	0.56	0.52	0.49
75	T	0.54	0.56	0.51	0.49
76	V	0.54	0.56	0.52	0.52
77	Y	0.56	0.58	0.53	0.54
78	I	0.58	0.59	0.55	0.55
79	D	0.59	0.6	0.56	0.56
80	Q	0.58	0.59	0.55	0.54
81	R	0.56	0.56	0.53	0.52
82	V	0.53	0.53	0.49	0.47
83	D	0.47	0.47	0.45	0.43
84	K	0.42	0.43	0.41	0.39
85	V	0.37	0.37	0.36	0.33
86	Y	0.33	0.33	0.32	0.29
87	K	0.29	0.30	0.29	0.25
88	V	0.27	0.27	0.26	0.23
89	V	0.24	0.24	0.24	0.20
90	K	0.22	0.22	0.22	0.17
91	L	0.21	0.21	0.20	0.16
92	P	0.2	0.20	0.20	0.15
93	V	0.2	0.21	0.20	0.16
94	E	0.21	0.21	0.21	0.17
95	I	0.21	0.22	0.21	0.17
96	E	0.21	0.21	0.21	0.18
97	Q	0.21	0.21	0.21	0.18
98	Q	0.21	0.20	0.20	0.17
99	D	0.19	0.18	0.18	0.15
100	I	0.16	0.15	0.15	0.12

Table 20 (Continued)

Residue number	Amino acid	Disorder probability of the residue (WT)	Disorder probability of the residue (E11V)	Disorder probability of the residue (M12T)	Disorder probability of the residue (E22G)
101	S	0.12	0.11	0.11	0.10
102	A	0.1	0.09	0.09	0.09
103	K	0.09	0.08	0.08	0.08
104	Y	0.08	0.07	0.07	0.07
105	S	0.07	0.07	0.06	0.06
106	E	0.06	0.06	0.05	0.05
107	G	0.05	0.05	0.05	0.05
108	I	0.05	0.05	0.05	0.05
109	L	0.05	0.06	0.05	0.05
110	T	0.06	0.07	0.06	0.06
111	V	0.07	0.08	0.07	0.08
112	R	0.09	0.10	0.09	0.09
113	M	0.13	0.15	0.13	0.15
114	K	0.21	0.23	0.22	0.23
115	T	0.3	0.32	0.30	0.32
116	K	0.39	0.41	0.39	0.40
117	N	0.46	0.48	0.46	0.48
118	I	0.55	0.57	0.54	0.57
119	K	0.64	0.66	0.63	0.66
120	N	0.71	0.74	0.7	0.73
121	V	0.69	0.72	0.69	0.71
122	E	0.73	0.76	0.72	0.72
123	I	0.71	0.73	0.7	0.71
124	E	0.8	0.82	0.79	0.79

Table 21: Disorder Probability of the residue for WT and NTD distal mutant variants

Residue number	Amino acid	Disorder probability of the residue (WT)	Disorder probability of the residue (V23GF26YI27T)	Disorder probability of the residue (F26YI27T)
1	M	0.66	0.66	0.73
2	Y	0.59	0.62	0.67
3	T	0.48	0.54	0.59
4	P	0.44	0.48	0.55
5	I	0.42	0.46	0.52
6	K	0.44	0.46	0.52
7	F	0.43	0.46	0.48
8	F	0.45	0.47	0.52
9	T	0.46	0.48	0.52
10	N	0.46	0.48	0.52
11	E	0.45	0.46	0.47
12	M	0.44	0.44	0.46
13	I	0.43	0.42	0.45
14	K	0.43	0.41	0.44
15	N	0.43	0.40	0.43
16	V	0.44	0.41	0.44
17	S	0.45	0.41	0.44
18	N	0.46	0.43	0.45
19	T	0.49	0.46	0.48
20	V	0.53	0.48	0.52
21	K	0.53	0.51	0.52
22	E	0.49	0.48	0.48
23	V	0.45	0.45	0.44
24	S	0.4	0.41	0.40
25	S	0.34	0.36	0.34
26	F	0.28	0.30	0.29
27	I	0.22	0.24	0.23
28	Y	0.16	0.19	0.18
29	P	0.12	0.14	0.14
30	P	0.09	0.11	0.12
31	V	0.1	0.10	0.11
32	T	0.1	0.11	0.12
33	L	0.12	0.13	0.14
34	Y	0.15	0.15	0.16
35	Q	0.16	0.16	0.17
36	D	0.16	0.16	0.16
37	S	0.14	0.14	0.15
38	S	0.12	0.12	0.13
39	D	0.1	0.09	0.09
40	L	0.09	0.09	0.09
41	V	0.08	0.08	0.08
42	L	0.07	0.07	0.07
43	E	0.06	0.06	0.06
44	A	0.05	0.05	0.05
45	E	0.04	0.05	0.05
46	M	0.04	0.04	0.04
47	A	0.05	0.04	0.04
48	G	0.05	0.04	0.05
49	F	0.05	0.05	0.05
50	D	0.06	0.05	0.06

Table 21 (Continued)

Residue number	Amino acid	Disorder probability of the residue (WT)	Disorder probability of the residue (V23GF26YI27T)	Disorder probability of the residue (F26YI27T)
51	K	0.06	0.06	0.06
52	K	0.07	0.06	0.07
53	N	0.08	0.07	0.07
54	I	0.08	0.07	0.08
55	K	0.09	0.08	0.09
56	V	0.09	0.08	0.09
57	S	0.09	0.08	0.09
58	V	0.1	0.09	0.10
59	N	0.1	0.09	0.10
60	K	0.11	0.10	0.10
61	N	0.12	0.09	0.11
62	V	0.11	0.10	0.10
63	L	0.11	0.10	0.10
64	T	0.12	0.10	0.11
65	I	0.14	0.11	0.14
66	S	0.19	0.15	0.19
67	A	0.24	0.20	0.24
68	E	0.3	0.26	0.30
69	R	0.36	0.32	0.37
70	K	0.42	0.39	0.43
71	R	0.47	0.44	0.48
72	E	0.53	0.48	0.54
73	Y	0.54	0.49	0.55
74	S	0.54	0.49	0.54
75	T	0.54	0.49	0.55
76	V	0.54	0.51	0.55
77	Y	0.56	0.53	0.57
78	I	0.58	0.54	0.58
79	D	0.59	0.55	0.59
80	Q	0.58	0.53	0.58
81	R	0.56	0.49	0.56
82	V	0.53	0.46	0.53
83	D	0.47	0.42	0.47
84	K	0.42	0.37	0.42
85	V	0.37	0.31	0.36
86	Y	0.33	0.27	0.32
87	K	0.29	0.23	0.29
88	V	0.27	0.20	0.25
89	V	0.24	0.18	0.23
90	K	0.22	0.15	0.21
91	L	0.21	0.14	0.20
92	P	0.2	0.14	0.19
93	V	0.2	0.14	0.20
94	E	0.21	0.15	0.21
95	I	0.21	0.16	0.21
96	E	0.21	0.16	0.21
97	Q	0.21	0.16	0.21
98	Q	0.21	0.16	0.20
99	D	0.19	0.15	0.18
100	I	0.16	0.12	0.15

Table 21 (Continued)

Residue number	Amino acid	Disorder probability of the residue (WT)	Disorder probability of the residue (V23GF26YI27T)	Disorder probability of the residue (F26YI27T)
101	S	0.12	0.10	0.11
102	A	0.1	0.09	0.09
103	K	0.09	0.08	0.08
104	Y	0.08	0.07	0.07
105	S	0.07	0.06	0.06
106	E	0.06	0.05	0.06
107	G	0.05	0.05	0.05
108	I	0.05	0.05	0.05
109	L	0.05	0.05	0.06
110	T	0.06	0.06	0.07
111	V	0.07	0.08	0.08
112	R	0.09	0.09	0.10
113	M	0.13	0.13	0.16
114	K	0.21	0.21	0.25
115	T	0.3	0.30	0.33
116	K	0.39	0.38	0.41
117	N	0.46	0.45	0.48
118	I	0.55	0.54	0.57
119	K	0.64	0.63	0.65
120	N	0.71	0.7	0.72
121	V	0.69	0.69	0.71
122	E	0.73	0.72	0.73
123	I	0.71	0.69	0.7
124	E	0.8	0.77	0.78

Table 22: Disorder Probability of the residue for WT and NTD distal mutant variants

Residue number	Amino acid	Disorder probability of the residue (WT)	Disorder probability of the residue (V31G)	Disorder probability of the residue (V31IL33I)	Disorder probability of the residue (F8YV31G)
1	M	0.66	0.72	0.68	0.71
2	Y	0.59	0.69	0.62	0.68
3	T	0.48	0.62	0.55	0.62
4	P	0.44	0.56	0.48	0.57
5	I	0.42	0.52	0.47	0.53
6	K	0.44	0.55	0.48	0.57
7	F	0.43	0.51	0.47	0.52
8	F	0.45	0.52	0.49	0.53
9	T	0.46	0.52	0.52	0.52
10	N	0.46	0.48	0.51	0.48
11	E	0.45	0.46	0.48	0.45
12	M	0.44	0.43	0.47	0.42
13	I	0.43	0.41	0.46	0.40
14	K	0.43	0.40	0.45	0.38
15	N	0.43	0.39	0.44	0.38
16	V	0.44	0.39	0.44	0.38
17	S	0.45	0.40	0.45	0.38
18	N	0.46	0.41	0.46	0.39
19	T	0.49	0.43	0.49	0.42
20	V	0.53	0.46	0.53	0.44
21	K	0.53	0.46	0.54	0.44
22	E	0.49	0.45	0.49	0.42
23	V	0.45	0.42	0.45	0.39
24	S	0.4	0.38	0.40	0.35
25	S	0.34	0.33	0.34	0.30
26	F	0.28	0.29	0.28	0.26
27	I	0.22	0.24	0.21	0.21
28	Y	0.16	0.20	0.15	0.17
29	P	0.12	0.17	0.11	0.15
30	P	0.09	0.15	0.10	0.13
31	V	0.1	0.15	0.10	0.14
32	T	0.1	0.16	0.10	0.15
33	L	0.12	0.18	0.11	0.18
34	Y	0.15	0.20	0.14	0.19
35	Q	0.16	0.20	0.16	0.20
36	D	0.16	0.19	0.16	0.18
37	S	0.14	0.16	0.15	0.16
38	S	0.12	0.13	0.13	0.13
39	D	0.1	0.09	0.09	0.10
40	L	0.09	0.09	0.09	0.09
41	V	0.08	0.08	0.08	0.08
42	L	0.07	0.07	0.07	0.07
43	E	0.06	0.06	0.06	0.06
44	A	0.05	0.05	0.05	0.05
45	E	0.04	0.05	0.05	0.05
46	M	0.04	0.04	0.05	0.04
47	A	0.05	0.04	0.05	0.04
48	G	0.05	0.04	0.05	0.05
49	F	0.05	0.05	0.06	0.05
50	D	0.06	0.06	0.06	0.06

Table 22: Disorder Probability of the residue for WT and NTD distal mutant variants

Residue number	Amino acid	Disorder probability of the residue (WT)	Disorder probability of the residue (V31G)	Disorder probability of the residue (V31IL33I)	Disorder probability of the residue (F8YV31G)
51	K	0.06	0.06	0.07	0.07
52	K	0.07	0.07	0.07	0.07
53	N	0.08	0.08	0.08	0.08
54	I	0.08	0.08	0.08	0.09
55	K	0.09	0.09	0.09	0.09
56	V	0.09	0.09	0.09	0.09
57	S	0.09	0.09	0.09	0.10
58	V	0.1	0.10	0.10	0.10
59	N	0.1	0.09	0.10	0.10
60	K	0.11	0.10	0.11	0.10
61	N	0.12	0.10	0.11	0.10
62	V	0.11	0.09	0.10	0.10
63	L	0.11	0.09	0.10	0.10
64	T	0.12	0.09	0.10	0.10
65	I	0.14	0.11	0.13	0.10
66	S	0.19	0.16	0.18	0.15
67	A	0.24	0.21	0.23	0.20
68	E	0.3	0.27	0.29	0.26
69	R	0.36	0.33	0.36	0.32
70	K	0.42	0.39	0.42	0.38
71	R	0.47	0.44	0.48	0.44
72	E	0.53	0.48	0.54	0.48
73	Y	0.54	0.52	0.55	0.52
74	S	0.54	0.53	0.55	0.53
75	T	0.54	0.53	0.55	0.54
76	V	0.54	0.54	0.55	0.55
77	Y	0.56	0.57	0.56	0.57
78	I	0.58	0.58	0.57	0.6
79	D	0.59	0.6	0.57	0.61
80	Q	0.58	0.58	0.55	0.6
81	R	0.56	0.56	0.52	0.58
82	V	0.53	0.52	0.47	0.54
83	D	0.47	0.46	0.44	0.48
84	K	0.42	0.42	0.39	0.44
85	V	0.37	0.37	0.35	0.39
86	Y	0.33	0.33	0.31	0.35
87	K	0.29	0.31	0.29	0.32
88	V	0.27	0.28	0.27	0.29
89	V	0.24	0.27	0.25	0.27
90	K	0.22	0.25	0.22	0.25
91	L	0.21	0.24	0.21	0.24
92	P	0.2	0.24	0.20	0.23
93	V	0.2	0.24	0.20	0.24
94	E	0.21	0.25	0.20	0.24
95	I	0.21	0.25	0.20	0.24
96	E	0.21	0.24	0.20	0.24
97	Q	0.21	0.24	0.19	0.23
98	Q	0.21	0.23	0.19	0.23
99	D	0.19	0.21	0.17	0.21
100	I	0.16	0.17	0.14	0.17

Table 22: Disorder Probability of the residue for WT and NTD distal mutant variants

Residue number	Amino acid	Disorder probability of the residue (WT)	Disorder probability of the residue (V31G)	Disorder probability of the residue (V31IL33I)	Disorder probability of the residue (F8YV31G)
101	S	0.12	0.13	0.11	0.13
102	A	0.1	0.09	0.09	0.09
103	K	0.09	0.09	0.08	0.09
104	Y	0.08	0.08	0.07	0.08
105	S	0.07	0.07	0.07	0.07
106	E	0.06	0.06	0.06	0.06
107	G	0.05	0.06	0.05	0.06
108	I	0.05	0.05	0.05	0.05
109	L	0.05	0.06	0.05	0.06
110	T	0.06	0.07	0.06	0.06
111	V	0.07	0.08	0.08	0.08
112	R	0.09	0.10	0.09	0.10
113	M	0.13	0.15	0.14	0.14
114	K	0.21	0.24	0.23	0.22
115	T	0.3	0.32	0.31	0.31
116	K	0.39	0.41	0.40	0.39
117	N	0.46	0.48	0.48	0.47
118	I	0.55	0.57	0.57	0.55
119	K	0.64	0.65	0.66	0.64
120	N	0.71	0.73	0.73	0.73
121	V	0.69	0.71	0.71	0.70
122	E	0.73	0.73	0.74	0.73
123	I	0.71	0.70	0.71	0.69
124	E	0.8	0.79	0.80	0.79

CURRICULUM VITAE

Surname, Name: Rafiq, Azra

E-mail: azrarafiq22@yahoo.com Phone: 0090-507-1619655

Room No: C3-342-4, Lisansüstü Konukevi, Middle
East Technical University, Çankaya, Ankara,
Turkey

RESEARCH INTERESTS and OBJECTIVE

To focus on chronic illnesses especially causes and treatment of Alzheimer's disease and Cancer. To work in team and achieve the planned goals in the minimum time possible. To be punctual and hardworking for any assigned task. To spend my life for the greater good of humanity.

ACHIEVEMENTS

Best Debater in English and Urdu of School, College and Bachelor level University.

Scholarship holder for MS and PhD, awarded by Higher Education Commission, Government of Pakistan.

Certificate of least Plagiarism (3%) in MS thesis writing.

Holder of Erasmus Exchange Program Scholarship.

Student Representative of Department of Biochemistry at METU, Ankara, Turkey. (2015-2016).

Certificate for Best Doctorate Student in PhD course work.

RESEARCH EXPERIENCE

METU, Ankara, Department of Chemistry

(Laboratory of Assoc. Prof. Dr. Nursen ÇORUH – Mar, 2013 to Nov, 2016)

METU, Ankara, Department of Molecular Biology and Genetics

(Laboratory of Prof. Dr. Semra KOCABIYIK – Jan, 2017 to May, 2022)

PUBLISHED ABSTRACT

A. Rafiq, S. Zabci, Y. Kaçer, S. Kocabiyik,, Temperature dependent variations in oligomer status of an archaeal small heat shock protein, TpvHSP14.3. *"BioMicroWorld Abstract Book"*, (2017), p.335.

EDUCATION

Degree	Institution	Year of Graduation
PhD	METU, Biochemistry	2022
MS	METU, Biochemistry	2015
BS	BZU, Doctor in Pharmacy	2011
High School	Pakistan Higher Secoondary School, U.A.E.	2005

FOREIGN LANGUAGES

Advanced English, Fluent Urdu, Basic Turkish, Basic Arabic