

STRUCTURAL MODIFICATIONS AND NOVEL MINI CHAPERONE  
DESIGNS OF AN ARCHAEAL SMALL HEAT SHOCK PROTEIN, *TPV*  
sHSP14.3

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

BY

SEMA ZABCI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR  
THE DEGREE OF DOCTOR OF PHILOSOPHY  
IN  
MOLECULAR BIOLOGY AND GENETICS

JUNE 2022



Approval of the thesis:

**STRUCTURAL MODIFICATIONS AND NOVEL MINI CHAPERONE  
DESIGNS OF AN ARCHAEL SMALL HEAT SHOCK PROTEIN, TPV  
sHSP14.3**

submitted by **SEMA ZABCI** in partial fulfillment of the requirements for the degree  
of **Doctor of Philosophy in Molecular Biology and Genetics, Middle East  
Technical University** by,

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

Prof. Dr. Ayşe Gül Gözen  
Head of the Department, **Biological Sciences**

Prof. Dr. Semra Kocabıyık  
Supervisor, **Biological Sciences, METU**

**Examining Committee Members:**

Assoc.Prof. Dr. Tülin Yanık  
Biological Sciences, METU

Prof. Dr. Semra Kocabıyık  
Biological Sciences, METU

Prof. Dr. Pınar Çalık  
Chemical Eng., METU

Prof. Dr. Günay Balta  
Child Health Institute, Hacettepe University

Asst.Prof. Dr. Çiğdem Akın Pekşen  
Mol. Biology and Gen., Başkent University

Date: 23.06.2022

**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: Sema Zabcı

Signature:

## ABSTRACT

### **STRUCTURAL MODIFICATIONS AND NOVEL MINI CHAPERONE DESIGNS OF AN ARCHAEOAL SMALL HEAT SHOCK PROTEIN, *TPV* sHSP14.3**

Zabcı, Sema

Doctor of Philosophy, Molecular Biology and Genetics

Supervisor: Prof. Dr. Semra Kocabiyık

June 2022, 348 pages

Small heat shock proteins (sHSPs) are characterized by a conserved,  $\alpha$ -crystallin domain (ACD) which is the structural hallmark and considered as main functional unit in their chaperone activity. In this study, targeted residues in the ACD of *Tpv* sHSP 14.3 were replaced by amino acids of different biochemical features by site directed mutagenesis to analyze their effects on structure and function of the *Tpv* sHSP14.3. Results from chaperone activity assays showed that *Tpv* sHSP 14.3 ACD mutants and wild type prevented the thermal inactivation of the model substrates; pig heart citrate synthase (CS) and yeast alcohol dehydrogenase (ADH). The extent of the protection against the heat induced inactivation was found to be different among the mutants and specific to the substrate proteins. Increasing hydrophobicity at the junction of ACD and N-terminal domain by Y34F mutation enhanced protection efficiency of the sHSP against heat induced inactivation of the ADH and CS. To investigate the importance of the charge residues in the ACD for the chaperone function, when the negative charge at position 45 (E45G) of the  $\beta$ 3 strand was eliminated, the chaperone activity was compromised. On the other hand, introduction of charged groups in place of hydrophobic residues (*i.e.*, G48E, I108K,

Y34FG48E, G48EI108K) resulted in higher chaperone activities against heat induced inactivation of ADH. This result indicated that both hydrophobic and electrostatic interactions can be crucial for the chaperone activity of *Tpv* sHSP 14.3. In accordance with the previous reports for other sHSPs, this study provided additional evidences for the critical roles of two conserved motifs; P/A-G doublet and G-x-L motif. Single point mutations, A47D, G107A and G107D, on those motifs, altered the stability and activity of the resultant mutant proteins. As deduced from changes in Rosetta Energy levels, substitution with the negatively charged residues more reduced than the hydrophobic amino acids the propensity of the sHSP variant for fibrillation. The contacts between alpha carbon atoms, as well as the RR distances, also varied as a result of the mutations. The extent of this variation was depended on the substituted amino acid, and the hydrophobic ones mostly preferred long-range contacts. On the other hand, there was a correlation between long range contacts and thermodynamic stability of the mutant proteins. Thermodynamic stabilities of the all mutant sHSP variants were reduced, except E43V mutant protein that had slight increase. The *Tpv* sHSP WT and most of the mutant variants possessed 24-mer oligomeric form which is typical for the sHSPs as revealed by BN-PAGE analysis. As an exception, two mutants (Y34F and I108K) instead formed 36-meric oligomers. The equilibrium of the oligomeric specie shifted toward larger species which was particularly obvious for the K87E mutant variant. Our results also indicated that the novel mini chaperones we designed regarding their superior chaperone activities as compared to WT *Tpv* sHSP14.3 can be considered promising for chaperone therapy.

Keywords: Archaea, *Thermoplasma volcanium*, Small Heat Shock Protein, Site Directed Mutagenesis, Mini Chaperone, Structural Modifications

## ÖZ

### **BİR ARKEA KÜÇÜK ISI ŞOKU PROTEİNİN (TPV-sHSP 14.3) YAPISAL MODİFİKASYONLAR VE ÖZGÜN KÜÇÜK ŞAPERON TASARIMLARI**

Zabcı, Sema  
Doktora, Moleküler Biyoloji ve Genetik  
Tez Yöneticisi: Prof. Dr. Semra Kocabıyık

Haziran 2022, 348 sayfa

Küçük ısı şoku proteinleri (sHSP) tipik bir yapısal öge olan ve şaperon aktivitesinde ana işlevsel birim olarak kabul edilen korunmuş, alpha-kristalin bölgeleri (ACD) ile karakterize edilirler. Bu çalışmada, mutasyonların yapı ve fonksiyon üzerindeki etkilerini araştırmak için *Tpv* sHSP 14.3'ün ACD'sinin hedef pozisyonlardaki gruplar, bölgeye özel mutajenez yoluyla farklı biyokimyasal özellikteki aminoasitlerle değiştirilmiştir. Şaperon aktivite deneylerinden elde edilen sonuçlar, *Tpv* sHSP 14.3 ACD mutantlarının ve yabancı sHSP'nin, model substratlar, domuz kalbi sitrat sentazı (CS) ve maya alkol dehidrogenaz (ADH)'nin termal inaktivasyonunu önlediğini göstermiştir. Isı kaynaklı inaktivasyona karşı korumanın boyutunun, mutantlara göre değiştiği ve substrat proteinlere özgü olduğu bulunmuştur. Y34F mutasyonu ile ACD ve N-terminal bölgesinin kesişim yerinde hidrofobikliğin artırılması, ADH ve CS'nin ısı kaynaklı inaktivasyonuna karşı sHSP proteinin koruma etkinliğini güçlendirmiştir. Ayrıca, ACD'nin yüklü gruplarının şaperon işlevindeki önemini incelemek için  $\beta$ 3 ipliğinin 45. pozisyonundaki (E45G) negatif yük elimine edildiğinde, şaperon aktivitesi kötüye gitmiştir. Öte yandan, hidrofobik kalıntılar yerine yüklü grupların eklenmesi (yani, G48E, I108K, Y34FG48E, G48EI108K), ADH'nin ısı kaynaklı inaktivasyonuna karşı şaperon aktivitesinin artması ile sonuçlanmıştır. Bu sonuç, hem hidrofobik hem de

elektrostatik etkileşimlerin, *Tpv* sHSP 14.3'ün şaperon aktivitesi için çok önemli olabileceğine işaret etmektedir. Çalışmamız aynı zamanda korunmuş iki motifin önemini de göstermiştir. Bu çalışma diğer sHSP'ler için daha önce yapılan bildirimlerle uyumlu olarak P/A-G ve G-x-L motiflerinin kritik rolleri hakkında ek kanıtlar ortaya koymuştur. Bu motiflerdeki tek nokta mutasyonları, A47D, G107A ve G107D, ortaya çıkan mutant proteinlerin stabilitesini ve aktivitesini değiştirmiştir. Rosetta Enerji düzeyindeki değişikliklere dayanarak, hidrofobik amino asitlerle yer değiştirmeye göre negatif yüklü kalıntılarla ikame durumunda fibrilasyon eğiliminde azalmanın daha fazla olduğu bulunmuştur. Mutasyonlar sonucunda, RR mesafesi yanı sıra, alfa karbon atomları arasındaki temaslar da değişmiştir. Değişim boyutu ikame edilen amino aside bağlı olup genellikle hidrofobik olanların uzun mesafeli temasları tercih ettiği görülmüştür. Diğer taraftan, uzun mesafeli temaslar ile mutant proteinin termodinamik stabilitesi arasında bir korelasyon olduğu bulunmuştur. Az artışın görüldüğü E43V mutant varyantı dışında, tüm mutant proteinlerin termodinamik stabilitesi azalmıştır. BN-PAGE analizi ile gösterildiği üzere, *Tpv* sHSP WT ve mutant varyantlarının büyük çoğunluğu, sHSPler için karakteristik olan 24-mer oligomere sahiptir. İstisna olarak, iki mutant (Y34F ve I108K) bunun yerine 36-merlik oligomerler oluşturmuştur. Özellikle, K87E mutant varyantında belirgin olarak oligomerik formların dengesi daha büyük türlerin oluşumu yönünde bir eğilim göstermiştir. Sonuçlarımız ayrıca, WT *Tpv* sHSP14.3'e kıyasla üstün şaperon aktivitelerine sahip oldukları dikkate alındığında, tasarladığımız özgün küçük şaperonların şaperon tedavisi için umut verici olabileceğini göstermiştir.

Anahtar Kelimeler: Arkeabakteri, *Thermoplasma volcanium*, Küçük Isı Şoku Proteini, Bölgeye Özgü Mutagenез, Küçük Şaperonlar, Yapısal Modifikasyonlar



To My Parents,

## ACKNOWLEDGMENTS

In the list of people, I wish to acknowledge for the completion of this piece, the first is my supervisor, Prof. Dr. Semra KOCABIYIK. Words cannot express my gratitude to my professor for her invaluable patience and feedback. I could not have undertaken this journey without her efforts and generously provided knowledge and expertise. I entered her laboratory as an under-graduate student and completing master study and Ph.D. research with her is a huge honour for me. Being far from family was difficult during this time period, but this difficulty eased as my professor always treated me like one of her family members. Moreover, I felt my laboratory as my home. The trust and encouragement she gave me from under graduate study helped me in achieving an unshakeable confidence in myself. I have grown up in the light of her ideas. Her advices will always serve as a source of light for me and the courage and strength she built in me will definitely help me in the times to come. I am grateful for her efforts to bringing me one step further. Although, completing my thesis study gives me happiness on one hand, however leaving my professor and our laboratory makes me feel sad, on the other. I am extremely thankful to her for this good chapter of my life, which spans through almost 10 years. I will miss all good times with her, but my connection will always continue. I truly wish that all supervisors come with such qualities in them.

I am really indebted to Prof. Dr. Emel Arınç, who introduced me to my advisor. Her wonderful instructions in the field of biochemistry and molecular pharmacology improved my academic background. She was always kind to me and loved me as her grandchild. I wish she could see me graduating. Her memories will always live with me and whenever her name passes my thoughts, I am remembering her with all respect. I know she can feel each of my success in heaven.

My gratitude to the committee members, Prof. Dr. Günay Balta, Prof. Dr. Pınar Çalık, Assoc. Prof. Dr. Tülin Yanık and Asst. Prof. Dr. Çiğdem Akın Pekşen for their time, remarks and advices regarding this research.

I would be remiss in not mentioning my family, especially my parents, my lovely mommy Sevil Zabcı and my dear father Kadir Zabcı. Prior to thanking my parents, my biggest thanks to my uncle Osman Zabcı, for the love and support, for he has always treated me as his child. I will forever be indebted to his kindness and courage. I am also extremely grateful to my parents for the trust they had in me and for making me what I am today. I totally owe my courage, strength, confidence. The love they instilled in me will always make me feel unique, secure and emotionally complete, no matter where I go and what I pursue. Their belief in me has kept my spirits and motivation high during the research process and their words always encouraged me when I was at my lowest. My mommy used to become happy when I succeeded in experiments and sad when the experiment did not work. I shared all my stresses with her and she never complained. Thanks to her for having this big patience. I would also like to express my sincere thanks to my sibling, my sweetest friend and toughest competitor at house, my brother, Ertuğrul Zabcı (My Eko) for his love and support, which was definitely not limited to the Ph.D. research but extends way beyond. His motivational gifts, small courage giving talks and surprising celebrations were a source of mood elevation for me and it was always comforting to have a thought that I have been blessed with the best brother. My gratitude is also extended to my cousin, Kaan Günal, for making me realize the true meaning and how to value health and that, to place health as first priority. Lastly, thanks to my uncle, Assoc. Prof. Dr. Selami Günal. I thank him for helping me survive through this stressful time and never letting me give up besides his guidance in academics.

Besides my family, I would like to express my deepest appreciation to my friends, who were never less than a family to me. Shahrzad Nikghadam, my Şehom, for her editing help, late-night feedback sessions, and moral support and guidance and for not making the distance be felt at all. She was always there in my time of need and her advices have truly helped me many a times to come out of stressful issues. Thanks also to her mom, Jila annem for having such a big heart filled with good intentions.

My house-mate Özge Demirdoğan, thank you for sharing good and bad times together. She was always like a younger sister to me. Spending time with her was invaluable. I enjoyed going theatre, opera and musical with her. Not only social activities, but also sometimes sitting with her without doing anything and smiling to the irrelevant things was indefinitely enjoyable for me. She is the witness of all my stress during thesis writing process but her support and motivation and sleepless night with me and picking me from department, late in the night, are some of her favours that I will always remain indebted to.

Thanks to Merve Akkulak, who never left me alone in the hard times. Her positive energy surrounded me and alleviated my mood. Also, the social activities that we had together contributed to accumulation of good memories. Furthermore, she behaved as a -80 deep freezer controller when I was on holidays. She, many a times, saved my samples due to the -80 troubles.

Extending my gratitude to Didem Mimirolu, who is more than a friend to me. The time we spent from morning to evening in department and outside of the campus, visiting antique stores, and Ankara Castle, tasting different foods and beverages, long hour walking, long hour talks, and in short, all the times spent together, is so valuable for me. I am really happy to meet her at the beginning of my Ph.D. study. Special thanks to her to be always on my side.

Thanking next is my dear friend, Sibel Öztürk. She is the strongest lady which I have ever met. She has grown up in the light of modern thoughts and she is the women's rights activist. She enlightened me with these ideas. I am admired from her strength for coping with the difficult conditions in the life. Thanks to her for providing me with the best company. I have great memories with her when sharing the same house. I would like to thank my close friends (sisters), Seda Çepkenlioğlu Kılıç and Dilara Fidan, for their friendship and motivation. Thank you for sharing many happy and memorable moments with me from prep school. We were known as 'trio. Now I am waiting excitedly to meet Ada Kılıç (my niece) in the coming days and I am happy that I will spend more time in the future with my little princess. I can never forget

Dilara's delicious Adana dolma when I was in quarantine during pandemic times. She never left me alone.

Evrin Tandođan, Eda Sezer, Simge Tutku Yıldırım, Yađmur Kaçer, Mehtap Yonca, Can Özaslangöz, Deniz Ak, Nastaran Deljavan and Rümeysa Fayetorbay, thank you for providing me with the best company, chit chat sessions, celebrations and also the courage and strength in the hard times. Not for a moment I felt alone or lonely and the warmth of your presence, always made me feel like home.

I want to thank Yusuf Melih Güner (YMG) for his support and motivation in the last 6 month of this stressful period. His psychological guidance, encouragement and friendship were extremely valuable to me.

In the end I would like to thank my labmate, Azra Rafiq. Crying, working, fighting, we did all, together. I am so lucky to have such a caring sister, best friend and colleague. She has been helpful and jolly and at times, angry and aggressive. Nevertheless, she was always gentle to me. I thank her for relieving me of my stress in difficult times. Additionally, she helped me a lot during thesis writing process by preparing delicious Pakistani foods. In short, I was always entertained and being in lab with her, throughout my research made me feel like we were twins at times, and two sides of a coin at other, thus making memories that are unforgettable. I would like to also thank Dr. Farah Rafiq to reply my questions regarding health issues without getting annoyed. Moreover, the extended Rafiq family, I thank them to treat me as their family members.

My friends from beyond the borders also deserve my great appreciation and these include, Kemal Jacob Alis, Zouheir Bellal, Elbay Malikmammadov, Sonum Chaudhary and Hira Tariq for the support and the warmth of their love that was always there.

I would like to thank Turkish Scientific and Technical Research Council (TÜBİTAK) (KBAG 113T085) and METU-BAP (TEZ-D-108-2020-10187, GAP-108-2019-10177) for the grants and fellowship to support this work.

## TABLE OF CONTENTS

|   |       |
|---|-------|
| ABSTRACT .....  | v     |
| ÖZ.....   | vii   |
| ACKNOWLEDGMENTS .....   | x     |
| TABLE OF CONTENTS .....   | xiv   |
| LIST OF TABLES .....  | xx    |
| LIST OF FIGURES .....   | xxiii |
| CHAPTERS  |       |
| 1 INTRODUCTION .....  | 1     |
| 1.1 Small Heat Shock Proteins .....   | 1     |
| 1.1.1 Structure of Small Heat Shock Proteins .....  | 2     |
| 1.1.2 Dimerization and Oligomerization of sHSPs .....   | 4     |
| 1.1.3 Recognition and Substrate Interactions of sHSPs .....   | 6     |
| 1.1.4 Mechanism of sHSPs, Action and Their Role in the Proteostasis.....                                      | 10    |
| 1.1.5 Identification of the Critical Residues for Chaperone Activity on the<br>ACD of sHSPs from Archaea..... | 11    |
| 1.1.6 Identification of the Important Residues on the ACD of sHSPs from<br>Prokaryotes .....                  | 14    |
| 1.1.7 Identification of the Important Residues on the ACD of sHSPs from<br>Eukaryotes .....                   | 17    |
| 1.1.8 Chaperone Activity of $\alpha$ -Crystallin Derived Mini Chaperones.....                                 | 20    |

|         |   |    |
|---------|---|----|
| 1.1.9   | Interaction of sHSPs with Other Proteins and Their Therapeutics         |    |
|         | Importance .....  | 23 |
| 1.2     | Scope and Aim of the Study.....   | 24 |
| 2       | MATERIALS AND METHODS.....  | 29 |
| 2.1     | Materials.....  | 29 |
| 2.1.1   | Chemicals, Enzymes, and Kits.....                                       | 29 |
| 2.1.2   | Buffer Solutions .....  | 30 |
| 2.1.3   | Molecular Size Markers and Plasmid Vectors.....                         | 30 |
| 2.2     | Strains and Growth Media.....   | 31 |
| 2.2.1   | Growth Medium and Culture Conditions .....                              | 31 |
| 2.3     | Methods.....  | 31 |
| 2.3.1   | Expression of Recombinant <i>Tpv</i> sHSP14.3 Protein from its Own      |    |
|         | AUG .....   | 31 |
| 2.3.1.1 | Isolation of the Target DNA for Amplification.....                      | 31 |
| 2.3.1.2 | Insert and Vector Preparation.....                                      | 33 |
| 2.3.1.3 | Ligation and Transformation into Chemically Competent <i>E.coli</i>     |    |
|         | (BL21(DE3)) Cells.....  | 33 |
| 2.3.1.4 | Screening for Recombinant Colonies by Plasmid Isolations and            |    |
|         | Restriction Enzyme Digestions.....                                      | 34 |
| 2.3.1.5 | Confirmation of Cloning by DNA Sequencing .....                         | 34 |
| 2.3.2   | Overexpression and Purification of the Wild Type <i>Tpv</i> sHSP 14.3 . | 36 |
| 2.3.3   | Site Directed Mutagenesis .....   | 37 |
| 2.3.3.1 | Constructing Single Site Mutants .....                                  | 38 |
| 2.3.3.2 | Construction of Double Site Mutants .....                               | 40 |

|        |   |    |
|--------|---|----|
| 2.3.4  | Transformation into the Competent <i>E.coli</i> Cells for Protein Expression .....                                      | 42 |
| 2.3.5  | Expression and Initial Purification of Mutant Proteins .....  | 42 |
| 2.3.6  | Chaperone Activity Assay with Pig Heart Citrate Synthase .....  | 43 |
| 2.3.7  | Thermal Aggregation Assay using CS as Substrate .....   | 43 |
| 2.3.8  | Chaperone Activity Assay with Alcohol Dehydrogenase .....   | 44 |
| 2.3.9  | Thermal Aggregation Assay with ADH .....  | 45 |
| 2.3.10 | Native Polyacrylamide Gel Electrophoresis .....   | 45 |
| 2.3.11 | Blue Native PAGE Analysis .....   | 45 |
| 2.3.12 | Designing Alpha Crystallin Derived Novel Mini-Chaperone Peptides .....  | 46 |
| 2.3.13 | Bioinformatics Analyses and Three-Dimensional Structure Modelling of <i>Tpv</i> sHSP14.3 .....                          | 47 |
| 3      | RESULTS .....   | 49 |
| 3.1    | Sub-cloning of the <i>Tpv</i> sHSP14.3 Gene for Expression of the Recombinant Protein from its Own AUG .....            | 49 |
| 3.1.1  | Screening of Transformant Colonies for Recombinant Clones by Plasmid Isolations and Restriction Enzyme Digestions ..... | 51 |
| 3.1.2  | Verification of Restriction Mapping Results by DNA Sequencing ..  | 54 |
| 3.2    | Selection of the Amino Acids to be Targeted in Mutagenesis .....  | 56 |
| 3.2.1  | Structure Superimpositions of <i>Tpv</i> sHSP 14.3 with Well Known Small Heat Shock Proteins .....                      | 61 |
| 3.3    | Overexpression of the Wild Type <i>Tpv</i> sHSP 14.3 and Heat Treatment ...   | 64 |
| 3.3.1  | Large Scale Cell Lysate Preparation for <i>Tpv</i> sHSP14.3 Protein Purification .....                                  | 66 |



|       |  |     |
|-------|--|-----|
| 3.4   | Site Directed Mutagenesis.....   | 69  |
| 3.4.1 | Preparation of plasmids.....   | 69  |
| 3.4.2 | Generating single site mutants .....   | 70  |
| 3.4.3 | Generating Double Site Mutants.....  | 74  |
| 3.5   | Transformation into the Competent BL21(DE3) <i>E.coli</i> for Protein Expression.....  | 75  |
| 3.6   | Transformation into the T7 Express <i>E.coli</i> cells for protein expression ..   | 77  |
| 3.7   | Expression and Initial Purification of Mutant Proteins .....   | 79  |
| 3.8   | HPLC Purification of the <i>Tpv</i> sHSP14.3 ACD Mutants .....   | 82  |
| 3.9   | Thermal Aggregation Measurements with Pig Heart Citrate Synthase ...   | 85  |
| 3.9.1 | Effect of the <i>Tpv</i> sHSP14.3 Mutants with Altered Hydrophobicity on CS Thermal Aggregation .....                        | 86  |
| 3.9.2 | Effect of the <i>Tpv</i> sHSP14.3 Mutants that Lost Charge by Introduction of Hydrophobicity on CS Thermal Aggregation ..... | 88  |
| 3.9.3 | Effect of the <i>Tpv</i> sHSP14.3 Mutants with Gain of Charge (Loss of Hydrophobicity) on CS Thermal Aggregation .....       | 90  |
| 3.9.4 | Effect of the <i>Tpv</i> sHSP14.3 Double mutants on CS Thermal Aggregation.....  | 92  |
| 3.9.5 | Effect of the <i>Tpv</i> sHSP14.3 Mutant with Reversing of the Charge on CS Thermal Aggregation .....                        | 93  |
| 3.10  | CS Activity Assay .....  | 94  |
| 3.11  | Yeast Alcohol Dehydrogenase Aggregation Assay.....   | 108 |
| 3.12  | Chaperone Activity Assay with Yeast Alcohol Dehydrogenase .....  | 117 |
| 3.13  | Native Polyacrylamide Gel Electrophoresis.....   | 121 |
| 3.14  | Blue Native PAGE Analysis.....   | 123 |

|         |   |     |
|---------|---|-----|
| 3.15    | Primary and Secondary Structure Analysis of <i>Tpv</i> sHSP14.3 Protein.  | 126 |
| 3.16    | Computational Modelling of 3-D Structure of <i>Tpv</i> sHSP14.3 .....   | 126 |
| 3.17    | Comparative Analysis of Intra- and Inter- Molecular Interactions of<br>Wild Type <i>Tpv</i> sHSP 14.3 protein and its mutant variants ..... | 129 |
| 3.17.1  | <i>Tpv</i> sHSP14.3 L33S Mutant Variant .....   | 129 |
| 3.17.2  | <i>Tpv</i> sHSP14.3 Y34F Mutant Variant.....  | 130 |
| 3.17.3  | <i>Tpv</i> sHSP14.3 G107A Mutant Variant.....   | 131 |
| 3.17.4  | <i>Tpv</i> sHSP14.3 E43V Mutant Variant .....   | 132 |
| 3.17.5  | <i>Tpv</i> sHSP14.3 E45G Mutant Variant .....   | 133 |
| 3.17.6  | <i>Tpv</i> sHSP14.3 A47D Mutant Variant.....  | 135 |
| 3.17.7  | <i>Tpv</i> sHSP14.3 G48E Mutant Variant .....   | 137 |
| 3.17.8  | <i>Tpv</i> sHSP14.3 G107D Mutant Variant.....   | 139 |
| 3.17.9  | <i>Tpv</i> sHSP14.3 I108K Mutant Variant .....  | 140 |
| 3.17.10 | <i>Tpv</i> sHSP14.3 Y34FG48E Double Mutant Variant .....  | 141 |
| 3.17.11 | <i>Tpv</i> sHSP14.3 G48EI108K Double Mutant Variant .....   | 143 |
| 3.17.12 | <i>Tpv</i> sHSP14.3 K87E Mutant Variant .....   | 146 |
| 3.18    | Prediction of the Fibril Forming Units.....   | 147 |
| 3.19    | Comparision of the Hydrophobic Surface of the WT and Mutants ....   | 154 |
| 3.20    | Prediction of <i>Tpv</i> sHSP14.3 WT and its Mutants Protein Stability After<br>Mutation .....  | 164 |
| 3.21    | Generation of Residue-Residue Distance Map Using UCSF Chimera<br>Tool.....  | 171 |
| 3.22    | Residue Contact Map Prediction of Mutants and WT .....  | 177 |
| 3.23    | Chaperone Activity of <i>Tpv</i> ACD Derived Mini Chaperones .....  | 199 |
| 4       | DISCUSSION.....   | 205 |

|     |  |     |
|-----|--|-----|
| 4.1 | Effect of the <i>Tpv</i> sHSP14.3 Mutants with Altered Hydrophobicity on Structure and Function .....                  | 208 |
| 4.2 | Effect of the <i>Tpv</i> sHSP14.3 Mutations that Remove Charge but Introduce Hydrophobicity .....                      | 215 |
| 4.3 | Effect of the <i>Tpv</i> sHSP14.3 Mutants with Gain of Charge (loss of hydrophobicity) on Structure and Function ..... | 218 |
| 4.4 | Effect of the <i>Tpv</i> sHSP14.3 Double Mutants on Structure and Function .....                                       | 225 |
| 4.5 | Effect of the <i>Tpv</i> sHSP14.3 Mutant with Reversing of the Charge on Structure and Function .....                  | 229 |
| 4.6 | Chaperone Activity of <i>Tpv</i> ACD Derived Mini Chaperones .....   | 231 |
| 5   | CONCLUSION .....   | 235 |
|     | REFERENCES .....   | 239 |
|     | APPENDICES   |     |
| A.  | Buffer Solutions and Compositions .....  | 255 |
| B.  | Molecular Size Markers .....   | 257 |
| C.  | Expression Vector Map .....  | 261 |
| D.  | pET 21-9 Sequencing Results .....  | 262 |
| E.  | Wild Type Sequencing Results .....   | 265 |
| F.  | Standart Proteins used in the BN-PAGE Analysis .....   | 268 |
| G.  | Structural Bond Analysis of the <i>Tpv</i> sHSP 14.3 WT and its Mutants .....  | 269 |
| H.  | Rosetta Energy Level Analysis .....  | 276 |
| I.  | Contact Map Analysis .....   | 279 |
|     | CURRICULUM VITAE .....   | 347 |

## LIST OF TABLES

### TABLES

|  |     |
|--|-----|
| Table 2.1. List of the mutagenic oligonucleotide primers.....  | 39  |
| Table 2.2. Cycling parameters for the QuickChange Site-Directed Mutagenesis Method.....  | 39  |
| Table 3.1. List of chosen residues and their equivalents as determined by MSA in the ACDs of the sHSP from top ten blast hit species archaea, eubacteria, and eukarya sHSPs..... | 56  |
| Table 3.2. Transformation results of single mutants .....  | 71  |
| Table 3.3. Summary of the sequencing results of the single amino acid substitutions .....  | 73  |
| Table 3.4. Transformation efficiency with <i>Tpv</i> sHSP mutant plasmids. ....  | 77  |
| Table 3.5. Thermodynamic stability of <i>Tpv</i> sHSP14.3 mutants.....   | 165 |
| Table 3.6. Thermodynamic stability of <i>Tpv</i> sHSP14.3 mutants using DynaMut2 Programme .....   | 170 |
| Table 3.7. Summary of contact prediction results of <i>Tpv</i> sHSP 14.3 and its mutants .....   | 178 |
| Table H. 1. Comparative analysis of the rosetta energy levels of <i>Tpv</i> sHSP WT and mutants (L33S, Y34F, E43V, E45G, A47D, G48E).....  | 276 |
| Table H. 2. Comparative analysis of the rosetta energy levels of <i>Tpv</i> sHSP WT and mutants (G107A, G107D, I108K, K87E).....   | 277 |
| Table H. 3. Comparative analysis of the rosetta energy levels of <i>Tpv</i> sHSP WT and mutants (Y34FG48E, and G48EI108K).....   | 278 |
| Table I. 4. List of residues forming contacts with the residue G107 in WT <i>Tpv</i> sHSP14.3 and A107 in G107A <i>Tpv</i> sHSP14.3 mutant.....                                  | 315 |
| Table I. 5. List of the specific contacts which are present in the WT <i>Tpv</i> sHSP14.3, but not for the L33S <i>Tpv</i> sHSP14.3 mutant and vice versa.....                   | 316 |
| Table I. 6. List of residues forming contacts with the residue L33 in WT <i>Tpv</i> sHSP14.3 and Ser33 in L33S <i>Tpv</i> sHSP14.3 mutant. ....                                  | 319 |

|   |     |
|---|-----|
| Table I. 7. List of the specific contacts which are present in the WT <i>Tpv</i> sHSP14.3, but not for the G48E <i>Tpv</i> sHSP14.3 mutant and vice versa.....        | 320 |
| Table I. 8. List of residues forming contacts with the residue Gly48 in WT <i>Tpv</i> sHSP14.3 and Glu48 in L33S <i>Tpv</i> sHSP14.3 mutant .....                     | 323 |
| Table I. 9. List of the specific contacts which are present in the WT <i>Tpv</i> sHSP14.3, but not for the I108K <i>Tpv</i> sHSP14.3 mutant and vice versa.....       | 323 |
| Table I. 10. List of residues forming contacts with the residue I108 in WT <i>Tpv</i> sHSP14.3 and K108 in I108K <i>Tpv</i> sHSP14.3 mutant.....                      | 325 |
| Table I. 11. List of the specific contacts which are present in the WT <i>Tpv</i> sHSP14.3, but not for the G107D <i>Tpv</i> sHSP14.3 mutant and vice versa .....     | 326 |
| Table I. 12. List of residues forming contacts with the residue G107 in WT <i>Tpv</i> sHSP14.3 and D107 in G107D <i>Tpv</i> sHSP14.3 mutant .....                     | 328 |
| Table I. 13. List of the specific contacts which are present in the WT <i>Tpv</i> sHSP14.3, but not for the A47D <i>Tpv</i> sHSP14.3 mutant and vice versa .....      | 328 |
| Table I. 14. List of residues forming contacts with the residue A47 in WT <i>Tpv</i> sHSP14.3 and D47 in A47D <i>Tpv</i> sHSP14.3 Mutant.....                         | 330 |
| Table I. 15. List of the specific contacts which are present in the WT <i>Tpv</i> sHSP14.3, but not for the E45G <i>Tpv</i> sHSP14.3 mutant and vice versa.....       | 331 |
| Table I. 16. List of residues forming contacts with the residue E45 in WT <i>Tpv</i> sHSP14.3 and G45 in E45G <i>Tpv</i> sHSP14.3 mutant.....                         | 333 |
| Table I. 17. List of the specific contacts which are present in the WT <i>Tpv</i> sHSP14.3, but not for the E43V <i>Tpv</i> sHSP14.3 mutant and vice versa.....       | 334 |
| Table I. 18. List of residues forming contacts with the residue E43 in WT <i>Tpv</i> sHSP14.3 and V43 in E43V <i>Tpv</i> sHSP14.3 mutant.....                         | 337 |
| Table I. 19. List of the specific contacts which are present in the WT <i>Tpv</i> sHSP14.3, but not for the Y34FG48E <i>Tpv</i> sHSP14.3 mutant and vice versa .....  | 338 |
| Table I. 20. List of residues forming contacts with the residues Y34, G48 in WT <i>Tpv</i> sHSP14.3 and F34, E48 in Y34FG48E <i>Tpv</i> sHSP14.3 double mutant.....   | 340 |
| Table I. 21. List of the specific contacts which are present in the WT <i>Tpv</i> sHSP14.3, but not for the G48EI108K <i>Tpv</i> sHSP14.3 mutant and vice versa ..... | 341 |

|  |     |
|--|-----|
| Table I. 22. List of residues forming contacts with the residues G48, I108 in WT <i>Tpv</i> sHSP14.3 and E48, K108 in G48E1108K <i>Tpv</i> sHSP14.3 Mutant ..... | 343 |
| Table I. 23. List of the specific contacts which are present in the WT <i>Tpv</i> sHSP14.3, but not for the K87E <i>Tpv</i> sHSP14.3 mutant and vice versa ..... | 343 |
| Table I. 24. List of residues forming contacts with the residues K87 in WT <i>Tpv</i> sHSP14.3 and E87 in K87E <i>Tpv</i> sHSP14.3 Mutant .....                  | 345 |

## LIST OF FIGURES

### FIGURES

|  |    |
|--|----|
| Figure 1.1. Scheme of domain structure of small heat shock proteins .....  | 2  |
| Figure 1.2 Dimer interfaces of nonmetazoan and metazoan sHSPs .....  | 5  |
| Figure 1.3. The organization of the proteostasis network.....  | 10 |
| Figure 1.4. sHsps influence the substrate aggregation process.....   | 11 |
| Figure 1.5. Multiple sequence alignment of ACD of the sHSPs sequences from different organisms by CLUSTAL O multiple sequence alignment server.....                    | 14 |
| Figure 1.6. Putative dimer interface points of <i>Tpv</i> sHSP 14.3.....   | 26 |
| Figure 1.7. Multiple sequence alignment of <i>Tpv</i> sHSP14.3 among the sHSPs sequences from different organisms by CLUSTAL O multiple sequence alignment server..... | 27 |
| Figure 2.1. Recombinant plasmid diagram of cloned <i>tpv</i> -sHSP 14.3 gene in pDrive Cloning Vector.....   | 32 |
| Figure 2.2. Flow chart of the sub-cloning experiment.....  | 35 |
| Figure 2.3. HPLC purification system with a pre-packed anion exchange column.. ..  | 37 |
| Figure 2.4. Overview of the QuikChange II site-directed mutagenesis method.....  | 40 |
| Figure 2.5. Overview of the GeneArt site-directed mutagenesis method.....  | 41 |
| Figure 3.1. Isolation of the target plasmid DNA.....   | 49 |
| Figure 3.2. The sequence of the recombinant pDrive-sHSP14.3 as template DNA. ....  | 50 |
| Figure 3.3. Amplification of the TVN0775 gene by PCR .....   | 50 |
| Figure 3.4. Agarose gel images of double digested insert and vector after purification .....   | 51 |
| Figure 3.5. Screening of the colonies by double digestion of plasmids samples of 1 to 10 .....   | 52 |
| Figure 3.6. Screening of the colonies of the plasmid samples 4 and 9 by restriction enzyme digestions .....  | 53 |

|   |    |
|---|----|
| Figure 3.7. Screening of the colonies of the plasmid samples 2, 5, 6 and 7 by restriction enzyme digestions .....   | 53 |
| Figure 3.8. Agarose gel image <i>E.coli</i> pET21_tvshSP9 samples .....   | 54 |
| Figure 3.9. DNA quality determination for sequencing by agarose gel electrophoresis.....  | 55 |
| Figure 3.10. Agarose gel image of recombinants in NovaBlue Competent cells....  | 55 |
| Figure 3.11. Multiple sequence alignment of Alpha Crystallin Domain of <i>Tpv</i> sHSP 14.3 with the sHSPs from archaea.....                                    | 57 |
| Figure 3.12. Multiple sequence alignment of Alpha Crystallin Domain of <i>Tpv</i> sHSP 14.3 with the sHSPs from eubacteria .....                                | 58 |
| Figure 3.13. Multiple sequence alignment of Alpha Crystallin Domain of <i>Tpv</i> sHSP 14.3 with the sHSPs from eukaryotic organisms.....                       | 59 |
| Figure 3.14. Superimposition of the wheat sHSP, HSP16.9 (blue, PDB:1gme) and <i>Thermoplasma volcanium</i> sHSP14.3 (pink).....                                 | 62 |
| Figure 3.15. Superimposition of the dimeric <i>Xanthomonas axonopodis</i> sHSP XaHspA, (blue, PDB: 3GUF) and <i>Thermoplasma volcanium</i> sHSP14.3 (pink) .... | 62 |
| Figure 3.16. Superimposition of the dimeric human CRYAB (blue, PDB:2klr) and <i>Thermoplasma volcanium</i> sHSP14.3 (Pink).....                                 | 63 |
| Figure 3.17. Superimposition of the <i>Methanococcus jannaschii</i> HSP16.5 (blue, PDB:1shs) and <i>Thermoplasma volcanium</i> sHSP14.3 (Pink) .....            | 64 |
| Figure 3.18. SDS PAGE gel for the expression and heat stability of recombinant <i>E.coli</i> pET21_tvshSP2 and <i>E.coli</i> pET21_tvshSP6 clones. ....         | 65 |
| Figure 3.19. Screening of the colonies from transformed T7 competent <i>E.coli</i> cells by double digestion of plasmids .....                                  | 65 |
| Figure 3.20. SDS PAGE gel for the expression and heat stability of recombinant T7 Express <i>E.coli</i> pET21-tvshSP4 clones. ....                              | 66 |
| Figure 3.21. SDS PAGE analysis of the prepared cell lysate of WT <i>Tpv</i> sHSP14.3 protein.....   | 66 |
| Figure 3.22. SDS PAGE analysis of the buffer exchanged sample .....   | 67 |



|   |    |
|---|----|
| Figure 3.23. Chromatogram of the wild type <i>Tpv</i> sHSP14.3 purification and its SDS PAGE analysis .....   | 68 |
| Figure 3.24. SDS PAGE analysis of the concentrated fractions of WT <i>Tpv</i> sHSP14.3 protein after anion exchange .....   | 68 |
| Figure 3.25. The vector map of the recombinant pET21_tvshSP2 plasmid. ....  | 69 |
| Figure 3.26. Agarose gel image of plasmids for mutagenesis. ....  | 70 |
| Figure 3.27. Single mutation positions in the small heat shock sequence of the <i>Tpv</i> sHSP14.3.....   | 70 |
| Figure 3.28. Agarose gel image of 1st set putative mutant plasmids .....  | 72 |
| Figure 3.29. Agarose gel image of 2nd set putative mutant plasmids .....  | 72 |
| Figure 3.30. Agarose gel image of 3rd set of putative mutant plasmids. ....   | 72 |
| Figure 3.31. Chromatogram images of L33S, Y34F, E43V, E45G, and A47D single mutations.....  | 73 |
| Figure 3.32. Chromatogram images of G48E, G107A, G107D and I108K single mutations.....  | 73 |
| Figure 3.33. Double mutation positions in the small heat shock sequence of the <i>Tpv</i> sHSP14.3.....   | 74 |
| Figure 3.34. Agarose gel image of putative double mutant plasmids.....  | 75 |
| Figure 3.35. Chromatogram image of the G48EI108K and Y34FG48E double mutations.....   | 75 |
| Figure 3.36. Agarose gel image of plasmids DNA of <i>Tpv</i> sHSP14.3 single mutants .....  | 76 |
| Figure 3.37. Screening of the colonies by double digestion of plasmids 1 to 18. ...   | 77 |
| Figure 3.38. Agarose gel image of isolated plasmid DNA of <i>Tpv</i> sHSP14.3 double mutants.....   | 78 |
| Figure 3.39. Screening of the colonies by double digestion of plasmids 1 to 8 .....   | 78 |
| Figure 3.40. SDS PAGE analysis of <i>Tpv</i> sHSP14.3 L33S, Y34F, E43V, G48E, K87E, I108K and G107A single mutants before and after heat treatment at 60-80°C ..... | 80 |

|   |    |
|---|----|
| Figure 3.41. SDS PAGE analysis of <i>Tpv</i> sHSP14.3 E45G, A47D and G107D mutants before and after heat treatment at 60-80°C ..... | 81 |
| Figure 3.42. SDS PAGE analysis of <i>Tpv</i> sHSP14.3 double mutant proteins before and after heat treatment at 60-80°C .....       | 81 |
| Figure 3.43. SDS PAGE analysis of the prepared five cell lysates .....  | 82 |
| Figure 3.44. SDS PAGE analysis of the buffer exchanged samples .....  | 83 |
| Figure 3.45. Chromatogram of the <i>Tpv</i> sHSP14.3 Y34F mutant protein purification and its SDS-PAGE analysis.....                | 83 |
| Figure 3.46. Chromatogram of the <i>Tpv</i> sHSP14.3 G48E mutant protein purification and its SDS-PAGE analysis.....                | 84 |
| Figure 3.47. Chromatogram of the <i>Tpv</i> sHSP14.3 I108K mutant protein purification and its SDS-PAGE analysis.....               | 84 |
| Figure 3.48. Optimization of the CS (a) and <i>Tpv</i> sHSP14.3 (b) concentration for the thermal aggregation measurements .....    | 85 |
| Figure 3.49. Effect of the <i>Tpv</i> sHSP 14.3 WT on the thermal aggregation of the CS .....                                       | 86 |
| Figure 3.50. Effect of the <i>Tpv</i> sHSP 14.3 L33S mutant on the thermal aggregation of the CS .....                              | 87 |
| Figure 3.51. Effect of the <i>Tpv</i> sHSP 14.3 Y34F mutant on the thermal aggregation of the CS .....                              | 87 |
| Figure 3.52. Effect of the <i>Tpv</i> sHSP 14.3 G107A mutant on the thermal aggregation of the CS.....                              | 88 |
| Figure 3.53. Effect of the <i>Tpv</i> sHSP 14.3 E43V mutant on the thermal aggregation of the CS .....                              | 89 |
| Figure 3.54. Effect of the <i>Tpv</i> sHSP 14.3 E45G mutant on the thermal aggregation of the CS .....                              | 89 |
| Figure 3.55. Effect of the <i>Tpv</i> sHSP 14.3 A47D mutant on the thermal aggregation of the CS .....                              | 90 |
| Figure 3.56. Effect of the <i>Tpv</i> sHSP 14.3 G48E mutant on the thermal aggregation of the CS .....                              | 91 |

|  |     |
|--|-----|
| Figure 3.57. Effect of the <i>Tpv</i> sHSP 14.3 I108K mutant on the thermal aggregation of the CS .....  | 91  |
| Figure 3.58. Effect of the <i>Tpv</i> sHSP 14.3 Y34FG48E double mutant on the thermal aggregation of the CS.....   | 92  |
| Figure 3.59. Effect of the <i>Tpv</i> sHSP 14.3 G48EI108K double mutant on the thermal aggregation of the CS.....  | 93  |
| Figure 3.60. Effect of the <i>Tpv</i> sHSP 14.3 K87E mutant on the thermal aggregation of the CS .....   | 94  |
| Figure 3.61. Effect of <i>Tpv</i> sHSP 14.3 wild type on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C.....        | 95  |
| Figure 3.62. Effect of <i>Tpv</i> sHSP 14.3 L33S mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C .....     | 95  |
| Figure 3.63. Effect of <i>Tpv</i> sHSP 14.3 Y34F mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C .....     | 96  |
| Figure 3.64. Effect of <i>Tpv</i> sHSP 14.3 G48E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C .....     | 96  |
| Figure 3.65. Effect of <i>Tpv</i> sHSP 14.3 G48EI108K mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C..... | 97  |
| Figure 3.66. Effect of <i>Tpv</i> sHSP 14.3 K87E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C .....     | 97  |
| Figure 3.67. Effect of <i>Tpv</i> sHSP 14.3 I108K mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C .....    | 98  |
| Figure 3.68. Effect of <i>Tpv</i> sHSP 14.3 G107A mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C. ....    | 98  |
| Figure 3.69. Effect of <i>Tpv</i> sHSP 14.3 E43V mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C .....     | 99  |
| Figure 3.70. Effect of <i>Tpv</i> sHSP 14.3 Y34FG48E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C.....  | 99  |
| Figure 3.71. Effect of <i>Tpv</i> sHSP 14.3 WT mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C .....       | 100 |

|   |     |
|---|-----|
| Figure 3.72. Effect of <i>Tpv</i> sHSP 14.3 Y34F mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C .....      | 100 |
| Figure 3.73. Effect of <i>Tpv</i> sHSP 14.3 G48E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1/250 (w/w) at 47°C .....      | 101 |
| Figure 3.74. Effect of <i>Tpv</i> sHSP 14.3 L33S mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C .....      | 101 |
| Figure 3.75. Effect of <i>Tpv</i> sHSP 14.3 E43V mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C .....      | 102 |
| Figure 3.76. Effect of <i>Tpv</i> sHSP 14.3 G107A mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C .....     | 102 |
| Figure 3.77. Effect of <i>Tpv</i> sHSP 14.3 K87E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C .....      | 103 |
| Figure 3.78. Effect of <i>Tpv</i> sHSP 14.3 I108K mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C .....     | 103 |
| Figure 3.79. Effect of <i>Tpv</i> sHSP 14.3 Y34FG48E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C .....  | 104 |
| Figure 3.80. Effect of <i>Tpv</i> sHSP 14.3 G48EI108K mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C ..... | 104 |
| Figure 3.81. Effect of <i>Tpv</i> sHSP 14.3 WT on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:147 (w/w) at 47°C .....               | 105 |
| Figure 3.82. Effect of <i>Tpv</i> sHSP 14.3 L33S mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:147 (w/w) at 47°C .....      | 106 |
| Figure 3.83. Effect of <i>Tpv</i> sHSP 14.3 G48E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:147 (w/w) at 47°C .....      | 106 |
| Figure 3.84. Effect of <i>Tpv</i> sHSP 14.3 E43V mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:147 (w/w) at 47°C .....      | 107 |
| Figure 3.85. Effect of <i>Tpv</i> sHSP 14.3 G107A mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:147 (w/w) at 47°C .....     | 107 |
| Figure 3.86. Effect of <i>Tpv</i> sHSP 14.3 Y34F mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:147 (w/w) at 47°C .....      | 108 |

|   |     |
|---|-----|
| Figure 3.87. ADH aggregation assay with Y34F mutant sHSP.....   | 109 |
| Figure 3.88. ADH aggregation assay with E43V mutant sHSP.....   | 110 |
| Figure 3.89. ADH aggregation assay with L33S mutant sHSP .....  | 111 |
| Figure 3.90. ADH aggregation assay with G48E mutant sHSP.....   | 112 |
| Figure 3.91. ADH aggregation assay with K87E mutant sHSP.....   | 113 |
| Figure 3.92. ADH aggregation assay with I108K mutant sHSP.....  | 114 |
| Figure 3.93. ADH aggregation assay with Y34FG48E mutant sHSP .....  | 115 |
| Figure 3.94. ADH aggregation assay with G48EI108K mutant sHSP .....   | 116 |
| Figure 3.95. ADH activity assay in the absence and presence of WT <i>Tpv</i> sHSP 14.3<br>.....   | 117 |
| Figure 3.96. ADH activity assay in the absence and presence of I108K <i>Tpv</i> sHSP<br>14.3.....   | 118 |
| Figure 3.97. ADH activity assay in the absence and presence of G48E <i>Tpv</i> sHSP 14.3<br>.....   | 118 |
| Figure 3.98. ADH activity assay in the absence and presence of K87E <i>Tpv</i> sHSP 14.3<br>.....   | 119 |
| Figure 3.99. ADH activity assay in the absence and presence of Y34F <i>Tpv</i> sHSP 14.3<br>.....   | 119 |
| Figure 3.100. ADH activity assay in the absence and presence of Y34FG48E <i>Tpv</i><br>sHSP 14.3.....   | 120 |
| Figure 3.101. ADH activity assay in the absence and presence of G48EI108K <i>Tpv</i><br>sHSP 14.3.....  | 120 |
| Figure 3.102. ADH activity assay in the absence and presence of L33S <i>Tpv</i> sHSP<br>14.3.....   | 121 |
| Figure 3.103. ADH activity assay in the absence and presence of E43V <i>Tpv</i> sHSP<br>14.3.....   | 121 |
| Figure 3.104. Oligomeric state analysis by native gel .....   | 122 |
| Figure 3.105. SDS PAGE analysis of WT <i>Tpv</i> sHSP14.3 and G48E, I108K, K87E,<br>E43V, Y34F and G48EI108K mutant sHSPs after buffer exchange ..... | 123 |

|   |     |
|---|-----|
| Figure 3.106. Blue Native Gel Analysis for WT and selected <i>Tpv</i> sHSP14.3 mutants.<br>.....  | 125 |
| Figure 3.107. Secondary structure prediction of the <i>Tpv</i> sHSP14.3 protein .....   | 126 |
| Figure 3.108. 3D Structure of <i>Tpv</i> HSP 14.3 generated by homology modelling.  | 128 |
| Figure 3.109. Secondary structure analysis of the ACD of <i>Tpv</i> sHSP14.3 together<br>with multiple sequence alignment of sHSPs from archaeal species..... | 128 |
| Figure 3.110. Ribbon model of the dimeric <i>Tpv</i> sHSP 14.3.....   | 128 |
| Figure 3.111. Intra (a) and intermolecular (b) hydrophobic bonds of <i>Tpv</i> sHSP14.3<br>WT at residue Leu33 .....  | 129 |
| Figure 3.112. Intramolecular hydrogen bonds of <i>Tpv</i> sHSP14.3 WT (a) and L33S<br>mutant (b) .....  | 130 |
| Figure 3.113. Intermolecular hydrogen bonds of <i>Tpv</i> sHSP14.3 WT (a) and L33S<br>mutant (b) .....  | 130 |
| Figure 3.114. Intramolecular hydrogen bonds of <i>Tpv</i> sHSP14.3 WT (a) and Y34F<br>mutant (b) .....  | 131 |
| Figure 3.115. Intramolecular hydrophobic bonds of <i>Tpv</i> sHSP14.3 WT (a) and Y34F<br>mutant (b) .....   | 131 |
| Figure 3.116. Intramolecular hydrophobic interactions in G107A mutant. ....   | 132 |
| Figure 3.117. Intramolecular hydrogen bond in the WT at position Gly107 .....   | 132 |
| Figure 3.118. Intermolecular hydrogen (a) and electrostatic (b) bonds of <i>Tpv</i><br>sHSP14.3 WT at residue Glu43.....                                      | 133 |
| Figure 3.119. Intramolecular hydrogen bonds of <i>Tpv</i> sHSP14.3 WT (a) and E43V<br>mutant (b) .....  | 133 |
| Figure 3.120. Intramolecular hydrogen bonds of <i>Tpv</i> sHSP14.3 WT (a) and E45G<br>mutant (b) .....  | 134 |
| Figure 3.121. Intermolecular hydrogen (a) and electrostatic (b) bonds of <i>Tpv</i><br>sHSP14.3 WT at residue Glu45.....                                      | 135 |
| Figure 3.122. Intermolecular hydrophobic bonds of <i>Tpv</i> sHSP14.3 WT at residue<br>Ala47. ....  | 136 |

|  |     |
|--|-----|
| Figure 3.123. Intramolecular hydrogen bonds of <i>Tpv</i> sHSP14.3 WT (a) and A47D mutant (b).....                               | 136 |
| Figure 3.124. Intermolecular hydrogen bonds of <i>Tpv</i> sHSP14.3 A47D mutant at residue Asp47.....                             | 137 |
| Figure 3.125. Intramolecular hydrogen bonds of <i>Tpv</i> sHSP14.3 WT (a) and G48E mutant (b).....                               | 137 |
| Figure 3.126. Intermolecular hydrogen bonds of <i>Tpv</i> sHSP14.3 WT (a) and G48E mutant (b).....                               | 138 |
| Figure 3.127. Intramolecular electrostatic bonds of <i>Tpv</i> sHSP14.3 G48E mutant at residue Glu48.....                        | 138 |
| Figure 3.128. Intermolecular electrostatic (a) and hydrogen (b) bonds of <i>Tpv</i> sHSP14.3 G107D mutant at residue Asp107..... | 139 |
| Figure 3.129. Intramolecular hydrogen bonds in G107D mutant.....   | 140 |
| Figure 3.130. Intramolecular electrostatic bonds of <i>Tpv</i> sHSP14.3 I108K at residue Lys108.....                             | 140 |
| Figure 3.131. Intramolecular hydrogen bonds of <i>Tpv</i> sHSP14.3 WT (a) and I108K mutant (b).....                              | 141 |
| Figure 3.132. Intramolecular electrostatic interactions in the double mutant Y34FG48E.....                                       | 142 |
| Figure 3.133. Intramolecular hydrogen bonds in the <i>Tpv</i> sHSP14.3 WT (a) and Y34FG48E double mutant (b).....                | 142 |
| Figure 3.134. Intramolecular hydrophobic bonds in the <i>Tpv</i> sHSP14.3 WT (a) and Y34FG48E double mutant (b).....             | 143 |
| Figure 3.135. Intermolecular hydrogen bonds in the <i>Tpv</i> sHSP14.3 WT (a) and Y34FG48E double mutant (b).....                | 143 |
| Figure 3.136. Intramolecular (a) and intermolecular (b) electrostatic Bonds in G48EI108K double mutant.....                      | 144 |
| Figure 3.137. Intramolecular hydrogen bond in WT (a) and G48EI108K (b).....  | 145 |
| Figure 3.138. Intermolecular hydrogen bond in WT at position Gly48.....  | 145 |
| Figure 3.139. Intramolecular hydrogen bond in WT (a) and K87E mutant (b)....   | 146 |

|   |     |
|---|-----|
| Figure 3.140. Intramolecular hydrophobic bond in the WT at position Lys87.....  | 146 |
| Figure 3.141. Rosetta Energy bar chart of <i>Tpv</i> sHSP 14.3 WT.....  | 147 |
| Figure 3.142. Rosetta Energy bar chart of the ACDs of <i>Tpv</i> sHSP 14.3 WT (a) and L33S (b), Y34F (c), E43V (d) .....                  | 151 |
| Figure 3.143. Rosetta Energy bar chart of the ACDs of <i>Tpv</i> sHSP 14.3 WT (a) and E45G(b), A47D (c), G48E (d),G107A(e), G107D(f)..... | 152 |
| Figure 3.144. Rosetta Energy bar chart of the ACDs of <i>Tpv</i> sHSP 14.3 WT (a) and I108K(b) K87E(c), Y34FG48E (d), G48EI108K (e) ..... | 153 |
| Figure 3.145. Hydrophobic surface of <i>Tpv</i> sHSP14.3 WT and L33S mutant.....  | 155 |
| Figure 3.146. Hydrophobic surface of <i>Tpv</i> sHSP14.3 WT and Y34F mutant .....   | 155 |
| Figure 3.147. Hydrophobic surface of <i>Tpv</i> sHSP14.3 WT and E43V mutant .....   | 156 |
| Figure 3.148. Hydrophobic surface of <i>Tpv</i> sHSP14.3 WT and G48E mutant .....   | 157 |
| Figure 3.149. Hydrophobic surface of <i>Tpv</i> sHSP14.3 WT and K87E mutant .....   | 158 |
| Figure 3.150. Hydrophobic surface of <i>Tpv</i> sHSP14.3 WT and G107D mutant....  | 159 |
| Figure 3.150. Hydrophobic surface of <i>Tpv</i> sHSP14.3 WT and I108K mutant .....  | 160 |
| Figure 3.152. Hydrophobic surface of <i>Tpv</i> sHSP14.3 WT and Y34FG48E mutant .....   | 161 |
| Figure 3.153. Hydrophobic surface of <i>Tpv</i> sHSP14.3 WT and G48EI108K mutant .....  | 162 |
| Figure 3.154. Superimposition of the WT (Blue) and the E45G (a), A47D(b) and G107A(c) mutant structures .....                             | 163 |
| Figure 3.155. DynaMut2 result of the E45G mutant.....   | 166 |
| Figure 3.156. DynaMut2 result of the L33S mutant .....  | 166 |
| Figure 3.157. DynaMut2 result of the I108K mutant.....  | 167 |
| Figure 3.158. DynaMut2 result of the G107D mutant .....   | 167 |
| Figure 3.159. DynaMut2 result of the A47D mutant .....  | 168 |
| Figure 3.160. DynaMut2 result of the G48E mutant.....   | 168 |
| Figure 3.161.DynaMut2 result of the Y34F mutant .....   | 169 |
| Figure 3.162. DynaMut2 result of the E43V mutant.....   | 169 |
| Figure 3.163.DynaMut2 result of the G107A mutant .....  | 169 |



|   |     |
|---|-----|
| Figure 3.164. DynaMut2 result of the K87E mutant .....                                    | 170 |
| Figure 3.165. RRDist map analysis of <i>Tpv</i> sHSP14.3 WT and G107A mutant ...          | 172 |
| Figure 3.166. RRDist map analysis of <i>Tpv</i> sHSP14.3 WT and E43V mutant.....          | 172 |
| Figure 3.167. RRDist map analysis of <i>Tpv</i> sHSP14.3 WT and G107D mutant ...          | 173 |
| Figure 3.168. RRDist map analysis of <i>Tpv</i> sHSP14.3 WT and A47D mutant .....         | 173 |
| Figure 3.169. RRDist map analysis of <i>Tpv</i> sHSP14.3 WT and L33S mutant .....         | 174 |
| Figure 3.170. RRDist map analysis of <i>Tpv</i> sHSP14.3 WT and K87E mutant.....          | 174 |
| Figure 3.171. RRDist map analysis of <i>Tpv</i> sHSP14.3 WT and E45G mutant.....          | 175 |
| Figure 3.172. RRDist map analysis of <i>Tpv</i> sHSP14.3 WT and Y34F mutant.....          | 175 |
| Figure 3.173. RRDist map analysis of <i>Tpv</i> sHSP14.3 WT and I108K mutant.....         | 176 |
| Figure 3.174. RRDist map analysis of <i>Tpv</i> sHSP14.3 WT and G48EI108K mutant<br>..... | 176 |
| Figure 3.175. RRDist map analysis of <i>Tpv</i> sHSP14.3 WT and G48E mutant.....          | 177 |
| Figure 3.176. RRDist map analysis of <i>Tpv</i> sHSP14.3 WT and Y34FG48E mutant<br>.....  | 177 |
| Figure 3.177. Contact map analysis of WT and Y34F mutant. ....                            | 180 |
| Figure 3.178. Contact map analysis of WT and G107A mutant.....                            | 181 |
| Figure 3.179. Contact map analysis of WT and L33S mutant. ....                            | 183 |
| Figure 3.180. Contact map analysis of WT and G48E mutant .....                            | 184 |
| Figure 3.181. Contact map analysis of WT and I108K mutant .....                           | 185 |
| Figure 3.182. Contact map analysis of WT and G107D mutant.....                            | 187 |
| Figure 3.183. Contact map analysis of WT and A47D mutant.....                             | 188 |
| Figure 3.184. Contact map analysis of WT and E45G mutant. ....                            | 190 |
| Figure 3.185. Contact map analysis of WT and E43V mutant .....                            | 191 |
| Figure 3.186. Contact map analysis of WT and Y34FG48E mutant. ....                        | 193 |
| Figure 3.187. Contact map analysis of Y34F and Y34FG48E mutant. ....                      | 193 |
| Figure 3.188. Contact map analysis of G48E and Y34FG48E mutant. ....                      | 194 |
| Figure 3.189. Contact map analysis of WT and G48EI108K mutant.....                        | 196 |
| Figure 3.190. Contact map analysis of G48E and G48EI108K mutant.....                      | 196 |
| Figure 3.191. Contact map analysis of I108K and G48EI108K mutant.....                     | 197 |

|  |     |
|--|-----|
| Figure 3.192. Contact map analysis of WT and K87E mutant.....  | 198 |
| Figure 3.193. Suppression of ADH aggregation in the absence or presence of YIE mini-chaperone .....  | 199 |
| Figure 3.194. Suppression of ADH aggregation in the absence or presence of GIK mini-chaperone .....  | 200 |
| Figure 3.195. Suppression of ADH aggregation in the absence or presence of SDS mini-chaperone .....  | 200 |
| Figure 3.196. Suppression of ADH aggregation in the absence or presence of LVD mini-chaperone .....  | 201 |
| Figure 3.197. Suppression of CS aggregation in the absence or presence of YIE mini-chaperone.....  | 202 |
| Figure 3.198. Suppression of CS aggregation in the absence or presence of SDS mini-chaperone .....   | 202 |
| Figure 3.199. Suppression of CS aggregation in the absence or presence of GIK mini-chaperone.....  | 203 |
| Figure 3.200. Suppression of CS aggregation in the absence or presence of LVD mini-chaperone .....   | 203 |
| Figure 4.1. Superimposition of the <i>Methanococcus jannaschii</i> HSP16.5 (blue, PDB:1shs) and <i>Thermoplasma volcanium</i> sHSP14.3 (Pink). ..... | 210 |
| Figure C. 1. pET-21a (+) vector map .....  | 261 |
| Figure D. 1. Sequence result analysis of the pET21_tvshSP9 recombinant plasmid with nucleotide msa analysis and chromatogram analysis.....           | 264 |
| Figure E. 1. Sequence analysis of the cloned gene.....   | 267 |
| Figure F. 1. Protein Standards.....  | 268 |

## CHAPTER 1

### INTRODUCTION

The presence of the misfolded proteins in the cell is a significant danger for cellular integrity since it can affect multiple processes and consequently leads to loss of proper protein function. Environmental stress conditions (*e.g.*, heat, oxidative stress) and gene mutations during protein synthesis are the reasons for the production of defective proteins. In order to cope with such problems, protein quality control systems, including proteasomes and major chaperone systems, reduce protein aggregation and help in protein refolding. One of the members of the proteostasis system is the small heat shock proteins (Reinle *et al.*, 2021). They have important roles in protecting proteins from irreversible aggregation without using energy until the conditions become suitable for cell activity, and subsequent re-folding is mediated by ATP- dependent chaperones such as Hsp70 (Sun & MacRae, 2005). In addition to stress response, they are also implicated in the variety of functional activities, including signal transduction, transcription, translation, cell cycle, and cell differentiation, regulation of cytoskeleton architecture, apoptosis and even tumorigenicity and a metastasize (Arrigo, 2013; Kannan *et al.*, 2012; Boelens, 2020). Because of their diverse physiological roles, sHSPs are supposed to be a potential treatment avenue for treatment of several diseases associated with protein aggregation (Haslbeck *et al.*, 2019).

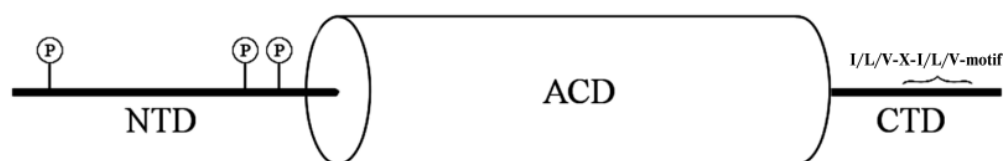
#### 1.1 Small Heat Shock Proteins

Small heat shock proteins are the widespread molecular chaperones found in the three domain of life. Eukaryotes generally have one or two cytosolic sHSP species; on the other hand, bacterium *Bradyrhizobium* spp has multiple sHSPs. The number of sHSPs increases in multicellular eukaryotes such that in humans and in plants, ten

copies and up to 50 copies of sHSPs are present, respectively. They are characterized by their small molecular weight (12-43 kDa) (Mogk *et al.*, 2019). Small heat shock proteins are the distinct molecular chaperones differentiated from other major chaperone families (Hsp60, Hsp70, Hsp90, and Hsp100) by their ATP independent function, which is termed as holdase activity (Mogk *et al.*, 2019). That means that they keep the damaged client proteins soluble but are unable to refold them. Release of bound substrate and subsequent refolding can be accomplished through participation of the ATP-dependent chaperones like HSP70/HSP27 (Haslbeck & Vierling, 2015).

### 1.1.1 Structure of Small Heat Shock Proteins

sHSPs have three partite domains consists of highly conserved alpha-crystallin domain flanked by variable and mostly disordered NTD and CTD terminal domains (Figure 1.1)(Reinle *et al.*, 2021).



**Figure 1.1. Scheme of domain structure of small heat shock proteins:** NTD, N-terminal domain; ACD,  $\alpha$ -crystallin domain; CTD, C -terminal domain (Tikhomirova *et al.*, 2017).

The ACD is the hallmark of the small heat shock protein consisting of 90-100 amino acids organized into six or seven beta strands resembling an immunoglobulin fold (Haslbeck *et al.*, 2019). A typical feature of all sHSPs is forming oligomers (*i.e.*, 9 to 40 subunits), and sHSP dimers are the basic building blocks of higher-order oligomers. sHSP oligomers are not static species and fluctuated by the environmental stress conditions. Subunits are continuously exchanged within minutes. This dynamic feature is very important for the function of sHSPs (Fu *et al.*, 2005; Mogk *et al.*, 2019).

The NTD is the most divergent region among sHSPs, both in sequence and length. The domain is located inside the oligomer structure. It is involved in substrate

interactions and interactions with the ACDs. Deletion of the NTD results in considerably smaller oligomers, showing that it stabilizes the overall structure of the oligomers (Boelens, 2020). The length of the N-terminus ranges from 24 residues in Hsp12.2 from *Caenorhabditis elegans* to 247 residues in Hsp42 from *Saccharomyces cerevisiae*. In the N-terminal domain, hydrophobic amino acids, e.g., phenylalanines and tryptophans, are predominantly found. In mammalian sHSPs, serine residues in the NTDs are phosphorylated by specific kinases in vivo. High flexibility of this region, as well as its susceptibility to proteolysis, limit the obtaining of the high quality of crystallographic data (Haslbeck *et al.*, 2019). Although NTDs are accepted as generally disordered, the N-termini resolved in the crystal structures of TaHSP 16.9 from wheat, TSP36 from the beef tapeworm *Taenia saginata*, and HSP 14.0 from *Sulfolobus tokodaii* showed a propensity to form helical secondary structure that affects oligomer organization and stability (Usui *et al.*, 2004; Stamler *et al.*, 2005).

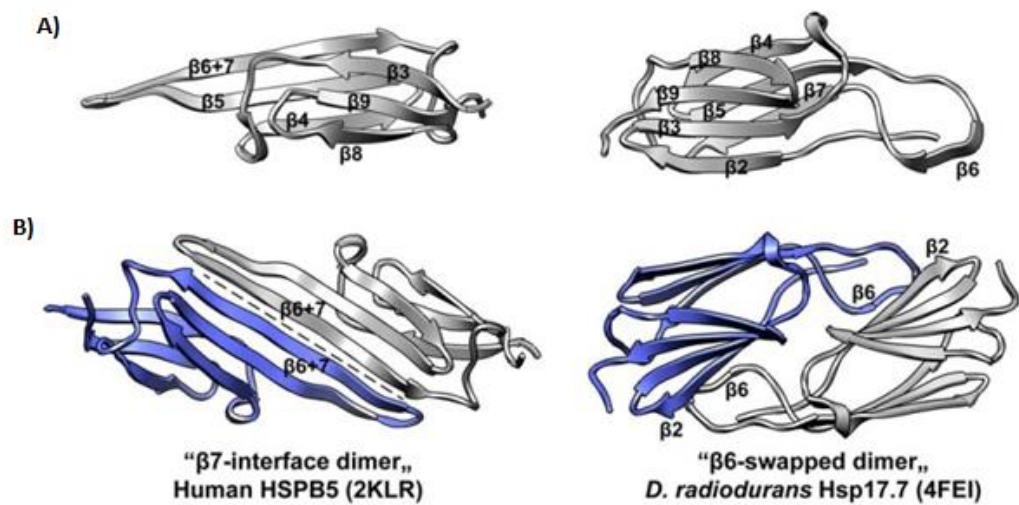
The highly disordered CTD is important for higher-order oligomerization. Also, polar and charged residues enriched in this region was thought to contribute solubilization of sHSPs in tissues. For example, a high concentration of sHSPs (>150 mg/mL) remains soluble in the eye lens (Janowska *et al.*, 2019). The conserved IXI/V motif of the C terminal region of many sHSPs is able to interact with a hydrophobic groove between strands  $\beta 4$  and  $\beta 8$  of an adjacent ACD, thus serves as a bridge between the dimers. By such bringing, ACD dimers interact with CTDs of cross-linked ACDs. This creates a meshwork and regulates oligomer formation in sHSPs. Studies on plant, Pea Hsp17.7, and bacterial *B.japonicum* sHSPs showed that the proper formation of high molecular weight complexes depended on the C-terminal extension. Deletion of CTEs mostly abolish oligomer formation and result in dimeric species formation (Mogk *et al.*, 2019; Boelens, 2020, Narberhaus, 2002).

### 1.1.2 Dimerization and Oligomerization of sHSPs

Oligomerization of sHSPs is a hierarchical blueprint and starts with the formation of dimers. sHSP dimers are formed either through interaction of the antiparallel  $\beta 6$  loop of one ACD with  $\beta 2$  strands of the partnering ACD in archaea, bacteria, fungi, and plants or interaction between extended  $\beta 6$ - $\beta 7$  strands of two partners in metazoan sHSPs (Figure 1.2). The  $\beta 6$ - $\beta 7$  interface is less stable; therefore, it helps dissociation into monomers, and this could be a reason that the mammalian sHSPs oligomers are consist of an odd number of subunits (Reinle *et al.*, 2021; Mogk *et al.*, 2019).

The characteristic feature of sHSPs is their dynamic quaternary structure, allowing them to form oligomeric species with different numbers of subunits (*e.g.*, 12-32 mers). Three sequence parts of the sHSPs, NTD, ACD, and CTD, take roles in forming the oligomers by making weak and intersubunit contacts (Riedl *et al.*, 2020). It is found that isolated ACDs alone are usually not enough for oligomer formation (Studer *et al.*, 2002). ACD dimer interactions in the core and the flanking regions determine the structure and allow dissociation and reassociation flexibly. Generally, the dimers formed by the ACD are associated with the first level of hierarchy. The second level of hierarchy is generated when the C-terminal I-X-I/V motif of one protomer interacts with the hydrophobic groove formed between  $\beta 4$ - $\beta 8$  strands of the other partner subunit, which leads to the generation of sub-oligomeric species (*e.g.*, tetramer and hexamer). The third level of hierarchy is regulated by the NTDs, which are required for the formation of higher-order of oligomers (Strauch & Haslbeck, 2016).

Under permissive conditions, sHSP oligomers are highly stable and show less subunit exchange. However, when they are exposed to stress, the rate of subunit exchange is enhanced significantly (Boelens, 2020). The equilibrium between dissociation and reassociation process is related to the regulation of the sHSPs activity (Haslbeck & Vierling, 2015).



**Figure 1.2 Dimer interfaces of nonmetazoan and metazoan sHSPs.** A) Immunoglobulin-like  $\beta$ -sandwich structure of the ACD. Note that ACD of mammalian and higher eukaryotic sHSPs (with the exceptions of plants) absent a distinct  $\beta 6$  strand but contain extended  $\beta 7$  strand (referred to as  $\beta 6+7$  strand) (left). In the ACD of bacterial sHSPs, there is a distinct  $\beta 6$  strand located in the loop connecting  $\beta 5$  and  $\beta 7$  strands (right). B) Structures of  $\beta 7$ -interface ACD dimers (*left*) and  $\beta 6$ -swapped ACD dimers (*right*) are shown. The ACDs of individual monomers are shown by *blue* and *gray* color. In human  $\alpha$ B-crystallin (*HSPB5*) (*left*) (PDB code 2KLR), the dimer interface is formed by the interaction of the extended  $\beta 6 + 7$  strands from neighboring monomers. The  $\beta 7$ -interface is pointed out by a *dashed line*. In Hsp17.7 from *D. radiodurans* (*right*) (PDB code 4FEI) (54), dimer is formed by the interactions of  $\beta 6$  strand with the  $\beta 2$  strand in the partner subunit (Haslbeck *et al.*, 2019).

In non-metazoans and plants, sHSPs become polydisperse upon exposure to heat stress; otherwise, they are usually monodisperse (Kampinga *et al.*, 2015; Collier & Benesch, 2020). This has allowed the determination of structural features of several sHSP oligomers by X-ray crystallography from yeast, wheat, archaea, and bacteria. Under physiological conditions, archaeal sHSPs form 24 mer oligomers as reported for sHSPs of *Methanococcus jannaschii*, *Sulfolobolus tokodaii*, and *Sulfolobolus solfataricus* (K. K. Kim *et al.*, 1998; Hanazono *et al.*, 2012; Liu *et al.*, 2015).

High resolution crystal structure of a plant small heat shock protein, Hsp16.9 from wheat showed a barrel-shaped structure which consists of two hexameric double disks with a total of 12-mers (R. L. van Montfort *et al.*, 2001). Similar to the wheat, Hsp21 from *Arabidopsis thaliana* have dodecameric oligomeric species of two hexameric disks (Ruttsdottir *et al.*, 2017). For the yeast SpHsp16.0 of

*Schizosaccharomyces pombe* and bacterium Hsp16.3 of *Mycobacterium tuberculosis* forms a 16 meric and nonameric oligomeric structures, respectively (Hanazono *et al.*, 2013; Fu *et al.*, 2005). On the other hand, human sHSPs are stress-inducible and highly polydisperse. To illustrate, oligomers of HspB5 can vary from 10 to 50 subunits (Hochberg & Benesch, 2014; Collier & Benesch, 2020). Therefore, due to this plasticity, their structures are not fully characterized at high resolution; only structure is partially resolved with truncated protein. Vertebrate sHSPs are regulated not only by heat stress but also posttranslational modification (PTM), which is not common in sHSPs of plants, bacteria, archaea, and fungi (Garrido *et al.*, 2012; Collier & Benesch, 2020).

### **1.1.3 Recognition and Substrate Interactions of sHSPs**

sHSPs have ability to interact with unfolding protein substrates in a promiscuous manner. They have a high binding capacity towards a wide variety of substrates in vitro, in cell extracts, and in vivo (Mogk *et al.*, 2019). Even though there are several studies to identify substrate interactions in different sHSPs, it is still unclear how sHSPs recognize their substrates and which regions of sHSPs form contacts with the unfolded substrates. The emerging view indicated that substrate interactions occur by involvement of multiple sites in NTDs and ACDs, but participation of binding sites can be diverse depending on the substrate and structural state (Reinle *et al.*, 2021). Until now, the CTD was not found to directly involve in substrate interaction, but it has important roles in oligomerization and subunit exchange, so it has indirect affect on sHSPs chaperone activity (Janowska *et al.*, 2019).

Previous studies showed that the N-terminal domain could probably be the most important in recognizing and binding substrate proteins. It was documented that stress signals (*e.g.*, high temperature, N-terminal phosphorylation) destabilize the oligomeric structure by exposing hydrophobic interior surface including NTD that assists binding to substrate. An early study conducted on Human Hsp27 showed that substitution of aspartates for the native serines at sites 15, 78, and 82 mimicked the



phosphorylation of Hsp27, which resulted in enhanced the extent of binding to T4 lysozyme. Even binding to the substrate occurred to a larger extend in the double aspartate mutant, Hsp27-S78D/S82D, than single variant Hsp27-S15D. This research implies the importance of the N-terminus in substrate binding as well as how stress signal affect its mode of binding (Shashidharamurthy *et al.*, 2005).

In another research, an arm exchange of N-terminus showed the importance of this region in substrate binding and, chaperone activity. Basha *et al.*, (2006) found that exchanging N-terminal arm of two closely related dodecameric plant sHSPs, wheat Hsp16.9 and pea Hsp18.1, caused alterations in the chaperone function of these chimeric proteins such that chimeric sHSP with N-terminal of Hsp18.1 protected CS very effectively; on the other hand, proteins with Hsp16.9 N-terminus failed to protect CS. It was suggested that the higher effectiveness of Hsp18.1 against aggregation of model substrates can be related to the elevated hydrophobicity contributed by Phe residues at the N-terminal arm of Hsp18.1.

The involvement of the N-terminal region in chaperone activity was also deduced by truncation studies of sHSP from *Saccharomyces cerevisiae* (Hsp26). N-terminus truncated mutants showed decreased chaperone activity in suppressing the aggregation of the CS. This abolished chapeone activity was associated with the loss of the phenylalanine-rich region in the N-terminus (Stromer *et al.*, 2004; Haslbeck *et al.*, 2004). A similar report on Hsp16.3 of *Mycobacterium tuberculosis* also revealed that truncation within the N-terminal region affected the binding behavior of a hydrophobic probe, bis-ANS, and chaperone activity of the protein against DTT-induced aggregation of insulin adversely. These results strongly suggest that the N-terminal region of the Hsp16.3 protein involves in binding denaturing substrate proteins, and chaperone-like activity (Fu *et al.*, 2005).

Crosslinking experiments together with SDS-PAGE and immunoblotting analysis for PsHsp18.1 of pea also prove the role of the N-terminal region of sHSP in substrate binding. Single-site variants of PsHsp18.1 were generated by introducing the phenylalanine analog, Bpa, at specific positions throughout the protein that made crosslink with the substrate after heating. The results indicate that the PsHsp18.1 N-

terminal region is the most critical part for interacting with substrates; although multiple contact sites were identified throughout this sHSP (Jaya *et al.*, 2009). In addition, the extension of the N-terminus of Hsp16.5 of *Methanococcus jannaschii* by a 14 amino acid peptide, which is a sequence encoding human Hsp27 amino acids 57-70, yielded an ensemble of polydisperse oligomers with higher substrate-binding affinity (McHaourab *et al.*, 2012). Although the intrinsically disordered NTEs provide the majority of interaction sites for misfolded proteins, exposed hydrophobic sites in the ACD and the CTE also contact with substrates.

Amino acid substitutions, L9P, Q16R, and E25K, in the N-terminal region of cyanobacterium *Synechocystis* Hsp16.6 did not change the in-vitro chaperone activity of the mutant proteins significantly. At higher molar ratios of sHSP monomer to Luc (*i.e.*, 24/1), N-terminal mutants, L9P and E25K, showed better protection against heat-induced aggregation of Luc than WT, while the effect of Q16R was found to be similar to the WT. On the other hand, when the molar ratios of sHSP monomer to Luc decreased, L9P and E25K mutants behaved like WT, and Q16R mutant was found to be slightly less effective than WT (Giese *et al.*, 2005).

Similarly, replacing serine with alanine at the N-terminal domain of human  $\alpha\beta$  crystallin (S19, S45, S59) did not affect the chaperone activity of the mutant proteins (Muchowski *et al.*, 1999). In addition to such point mutations in the NTD, there are other reports that strengthen the idea of target protein binding via ACD of the sHSP. It was found from truncation studies that involve absence of the entire or parts of the N-terminal region (Pasta *et al.*, 2003; Yang *et al.*, 2005) or the C terminal region (Takemoto, 1994; Aquilina *et al.*, 2005) exhibited proper chaperone activity. A number of mutational studies were performed to enlighten the importance of the ACD of various sHSPs in substrate binding and chaperone activity. Introduction of charged residue, arginine at conserved residues found in core alpha crystallin domain of human  $\alpha\beta$  crystallin (F118, G141, T144, P160) resulted in slightly diminished chaperone activity relative to wild type for protecting the model substrates CS and ADH from thermal aggregation (Muchowski *et al.*, 1999).

Research conducted in the sHSP from pea, PsHsp18.1, showed substituting the

residues at the positions K72, K77, E79, D82, D83, and Q87 with an analog of phenylalanine Bpa made contacts with the model substrate Malate Dehydrogenase (MDH). The intensity of the crosslink was low in this region, even though increasing the surface hydrophobicity of the sHSP by Bpa could be thought to contribute to substrate interactions with forming hydrophobic surfaces. On the other hand, strong interactions in the alpha-crystallin domain of PsHsp18.1 occurred in or near the  $\beta$ 7 strand (*i.e.*, K112 and L114, which correspond to K84 and Y86 in *Tpv* sHSP14.3). This was interpreted as a potential substrate-binding site (Jaya *et al.*, 2009). Analysis of the crystal structure of wheat sHSP TaHsp16.9 revealed that the N-terminal part of the protein in 3-D structure was close to the  $\beta$ 7 strand of the ACD. Therefore, they suggested that strong interactions in this region of ACD could occur with the N terminal part, but no experimental data was documented (R. L. van Montfort *et al.*, 2001).

Substrate binding residues were identified by incorporating the Bpa in small heat shock protein from *E. coli*'s IbpB over a wide range of temperatures (*i.e.*, 20 to 50°C). Nineteen of eighty-one residues in the alpha-crystallin domain were replaced by Bpa. Three residues made crosslink at 30°C, but at elevated temperature, additional 11 residues participated in photocrosslinking (*e.g.*, Tyr45, Lys71, Lys78). It was stated that most of the substrate-binding residues consisted of polar amino acids, and they were both buried and exposed amino acids. Most of the residues capable of mediating substrate binding at low temperature were found in the N-terminus. This study showed substrate-binding points distributed over the structure of sHSP and activated at high temperatures (Fu *et al.*, 2013).

Mapping of the yeast sHSP and Malate Dehydrogenase (MDH) interaction sites was performed by chemical crosslinking of sHSP/MDH complexes and followed by identification via mass spectrometry. Results showed that Lys 45 of the NTD, Lys 151 of the ACD (loop connecting  $\beta$ 5- $\beta$ 7), and Lys195-Lys198 of CTD in yeast Hsp26 formed major crosslink sites to the MDH. Disuccinimidyl suberate (DSS) crosslinking pointed out that C-terminal extension of the model substrate, MDH, become surface-exposed upon heating; therefore, it provided the major substrate

interaction sites. It was suggested that MDH sites started to unfold after heating and were recognized by sHSPs immediately (Ungelenk *et al.*, 2016).

#### 1.1.4 Mechanism of sHSPs, Action and Their Role in the Proteostasis

Protein homeostasis is maintained by the regulation of several hundred proteins consisting of molecular chaperones and their regulators, which have critical roles in protein folding or refolding, and the ubiquitin-proteasome system (UPS) and autophagy system, which help removal of irreversibly aggregated proteins. A compromised proteostasis system was linked to several diseases in human (*e.g.*, neurodegeneration and dementia, type II diabetes, cystic fibrosis, cancer and heart issues) (Hartl *et al.*, 2011). Since small heat shock proteins prevent inappropriate interactions of the misfolded proteins with other macromolecules, they have central roles in the proteostasis network. When proteins are exposed to environmental stress or gene mutations or errors occur during protein synthesis steps, they start to unfold. sHSPs bind to substrate and make complex. This process is called "sequestration". Sequestered substrates can be released by the assistance of the Hsp70/100 "disaggregases". The substrates then, can be directed to refolding pathways with Hsp60/Hsp70/Hsp90 molecular chaperones or degraded by proteosomes or autophagy (Figure 1.3) (Reinle *et al.*, 2021).

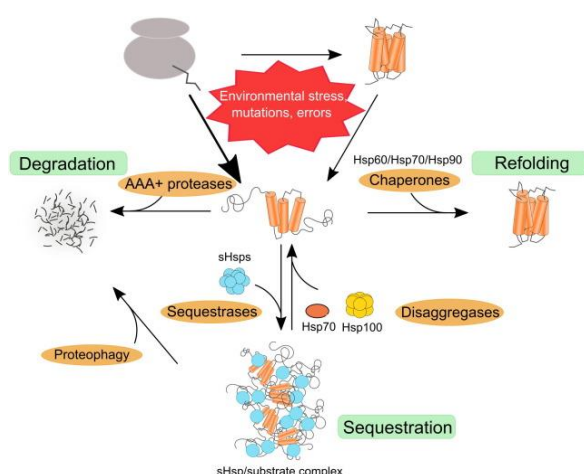
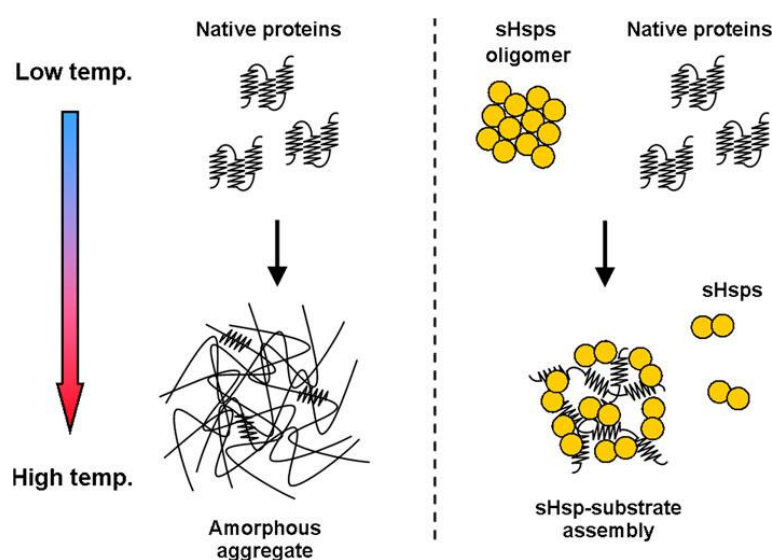


Figure 1.3. The organization of the proteostasis network (Reinle *et al.*, 2021).

Large oligomers of sHSPs are inactive, storage form of the sHSPs. It is believed that they are dissociated at high temperature into smaller species, mainly dimers which are the active form of the sHSPs (Mogk *et al.*, 2019). Intermediate sHSP and aggregation prone partially denatured polypeptide complexes may form huge assemblies consisted of both unfolded substrate and multiple sHSPs. By help of the sHSP outer shell, further aggregation of the cluster of folding intermediates are prevented (Figure 1.4) (Obuchowski *et al.*, 2021).



**Figure 1.4.** sHsps influence the substrate aggregation process (Obuchowski *et al.*, 2021).

### 1.1.5 Identification of the Critical Residues for Chaperone Activity on the ACD of sHSPs from Archaea

In small heat shock proteins, identification of the regions in the alpha crystallin domain for chaperone function and oligomerization has been assessed by using site-directed mutagenesis. In archaea, such studies are quite limited, and *Methanococcus jannaschii* has been the well representative model organism in these studies. Crystal structure analysis of MjHSP16.5 indicated extensive subunit-subunit contacts between  $\beta 2$  strand of one subunit and  $\beta 6$  strand in the other subunit (*e.g.*, Glu92–His 53, Glu 92–Lys 55, Glu 49–Arg 93, Asp51–Arg93). Also, extensive hydrophobic contacts at subunit interaction sites were found (*e.g.*, Phe42–Phe 42, Gly 127;

Gly62–Phe 42, Val 128; Ile 86–Ile 48; Ile 94–Ile 48). Furthermore, the putative substrate-binding region of ACD of Hsp18.1 as deduced by bis-ANS studies (Garrett J. Lee *et al.*, 1997) corresponds to the  $\beta$ 3- $\beta$ 4 loop of the MjHsp16.9. This loop was found to have conserved hydrophobic residues which are important in dimer interaction (K. K. Kim *et al.*, 1998).

When highly conserved residue R107 in the ACD of MjHSP16.5 was mutated into the glycine, crystal structure analysis did not show any significant alterations in the oligomeric state and overall structure of MjHSP16.5 WT and its R107G mutant. However, cryoelectron microscopy results revealed the structural changes at the dimer interface when the sample was heated at 60°C. The distance between dimer-dimer interface was increased from 12.4 nm to 13.6 nm after mutation. This conformational change was attributed to the loss of the intermolecular electrostatic interactions between R107 and E98 that might be critical for oligomerization (Quinlan *et al.*, 2013). Equivalent cataract causing R120G mutation in the CRYAB results in the loss of the two positive charge along the dimer interface and gain of two new interface ion pairs between His83 and Asp80. This interaction closes the groove by forming new salt bridges and consequently disturbs subunit exchange dynamics and impairs the chaperone activity (Bova *et al.*, 1999; Clark *et al.*, 2011). On the other hand, the equivalent interface in MjHSP16.5 R107G mutant became a more open state at 60°C, which could provide a possible explanation of its enhanced chaperone activity towards heat-induced aggregation of  $\beta$ L-crystallin (Quinlan *et al.*, 2013).

In another study, Xi *et al.*, 2014 searched the role of the non-conserved fragments and the conserved minimal Alpha Crystallin Domain of the Mj Hsp16.5 by generating two deletion mutants of the sHSP; N- and C-terminal deleted one ( $\Delta$ N $\Delta$ C), and mutant deleted N-terminal, C-terminal and  $\beta$ 6 dimerization loop ( $\Delta$ N $\Delta$ L $\Delta$ C). Results showed that mutant  $\Delta$ N $\Delta$ L $\Delta$ C had only minimal ACD domain of Mj Hsp16.5 was sufficient to protect aggregation of insulin B chain and the modified amyloidogenic peptide dansyl -SSTSAW at room temperature. Also, the chaperone activity of this mutant was higher than the WT. Bis-ANS experiment also

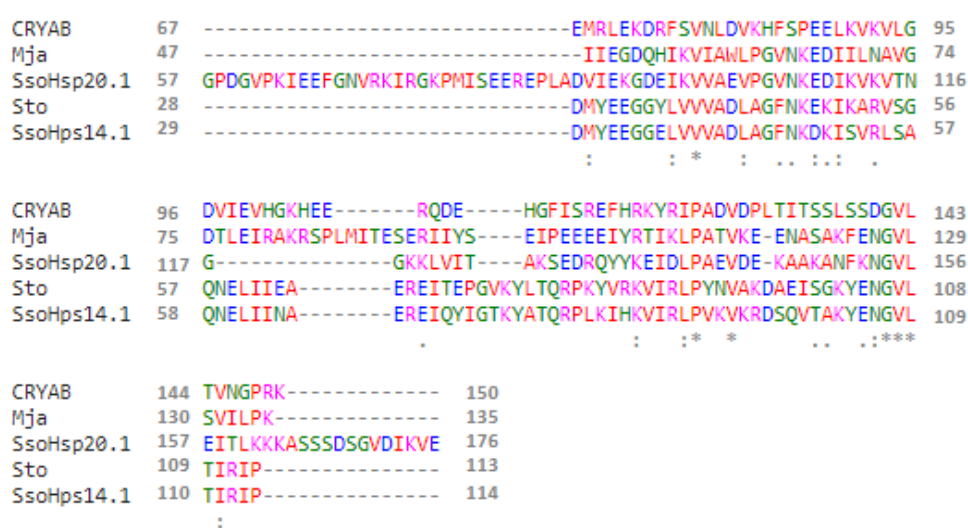
revealed that minimal ACD region exposed more hydrophobic surface for substrates than WT and  $\Delta N\Delta C$ . On the other hand, two deletion mutants failed to prevent CS thermal aggregation at 50°C, implying that N-terminal, C-terminal, and oligomerization were required to make complexes with this substrate to prevent further aggregation at higher temperatures. This study suggested that Mj Hsp16.5 utilized different mechanisms when exerting their chaperone-like activities for different substrates.

A structural study conducted on StHsp14.0 of the acidothermophilic archaeon *S.tokodaii* revealed that IXI/V motif and its binding cleft between the  $\beta 4$  and  $\beta 8$  strands are also the sites of specific interaction with peptide. The interaction between the IXI/V motif and the cleft was found weak. Ile120 and Ile122 in the IXI/V motif make hydrophobic interactions with the residues found in  $\beta 4$  strand (Ile50, Ala52, Val54), in  $\beta 8$  strand (Ile99), and in  $\beta 9$  strand (Leu108 and Ile 110). An exactly similar way of this binding mode is also seen in the MjHsp16.5 and TaHsp16.9. It has been suggested that those weak interactions were crucial for denatured substrate interactions and the formation of partially resolved oligomers that contribute to the activation of the StHsp14 (Hanazono *et al.*, 2012).

In another research, the resolved crystal structure of the ACD dimer of *Sulfolobus solfataricus* Hsp20.1 demonstrated that ionic interactions and interbackbone hydrogen bonds, including multiple charged or polar residues of ACD, were required for dimerization. For example, salt bridges between E101 in one monomer and R83 and R130 in other monomers involved in subunit- subunit interactions to make the dimer. Although these residues of Hsp20.1 were not highly conservative in non-metazoan sHSPs, other non-metazoan sHSP from *Sulfolobus tokodaii* Hsp14.0 had also charged residues to form salt bridges (*e.g.*, D41, R66 and R79) (L. Liu, Chen, Yang, & Wang, 2015)(Figure 1.5). Role of the predicted salt bridges and hydrogen core in the SsHsp14.1 was associated with the dynamic nature of the oligomers (dissociation/ association), which was proposed to their chaperone process (Wen *et al.*, 2010). A point mutation in the hydrophobic groove of *Sulfolobus solfataricus* P2 Hsp14.1 ( $\beta 4$ - $\beta 8$ ) to change a hydrophobic amino acid to a hydrophilic one, A102D,

results in loss of higher-order oligomers; instead, octamers were observed (Liu, Chen, Yang, Wang, *et al.*, 2015). This mutation significantly changed the NTD of the protein from straight helix conformation to the bent shape. Extensive hydrophobic interactions provided by the NTD were involved in the assembly of four SsHsp14 dimers of A102D to octamers.

Multiple sequence alignment of the ACD of the mentioned organisms are shown in the Figure 1.5.



**Figure 1.5. Multiple sequence alignment of ACD of the sHSPs sequences from different organisms by CLUSTAL O multiple sequence alignment server.** sp|P02511.2|CRYAB\_HUMAN, WP\_010869783.1: Mja: Hsp20/alpha crystallin family protein [*Methanocaldococcus jannaschii*], AAK42571.1: SsoHsp20.1: Small heat shock protein hsp20 family [*Saccharolobus solfataricus* P2], WP\_010979712.1: Sto: heat-shock protein Hsp20 [*Sulfolobus tokodaii*], WP\_009989320.1: SsoHsp14.1: Hsp20/alpha crystallin family protein [*Saccharolobus solfataricus*].

### 1.1.6 Identification of the Important Residues on the ACD of sHSPs from Prokaryotes

Critical residues for the function of ACD in prokaryotic sHSPs were investigated mostly in *Mycobacterium* sp. by mutagenesis. In an early study, Leu122 of Hsp16.3 from *Mycobacterium tuberculosis* was substituted by alanine or valine to understand the importance of this residue in structural stability and chaperone function. This residue was positioned on the highly conserved "GVLTVTV" motif of the core alpha crystallin domain. The results suggested this highly conserved Leu residue might be



important for in vivo and in vitro structural stability of Hsp16.3, and minor alterations on this residue could result in the loss of the chaperone-like activity (Mao & Chang, 2001). Similarly, an equivalent mutation in position Leu116 in HspH from *Bradyrhizobium japonicum* caused also structural and functional defects. The L116A mutant had reduced complex size (up to 110 kDa), and its activity was completely abolished against heat-induced precipitation of CS at 43°C (Lentze *et al.*, 2003a). Moreover, the corresponding residue is Leu129 of *M. jannaschii* Hsp16.5. In the 3D structure of Hsp16.5, Leu129 was positioned near the hydrophobic amino acids. It was proposed that the hydrophobic core might play a remarkable role in maintaining the oligomeric structure of sHSPs (Kim *et al.*, 1998; Mao & Chang, 2001).

In addition to these, Lentze *et al.*, 2003 investigated roles of the alpha crystallin domain in oligomerization and chaperone function of the soybean-symbiont *Bradyrhizobium japonicum* HspH by generating 12 different mutants. Among these, six of the mutants (*i.e.*, G74E, F94A, F94D, L100A, A109S, and L116A) had dramatically reduced or impaired chaperone function, and reduced the oligomer complex size. Residue F94 was replaced by either alanine or aspartic acid. Both mutations did not form even dimers (*e.g.*, Mwt 25 kDa), and they tended to precipitate. Substitution of G74 by alanine in HspH did not have a significant effect, but oligomerization and chaperone activity were diminished after G74E mutation. This study also pointed out that the  $\beta$ 5 and  $\beta$ 7 region of the ACD was important for the integrity of the overall structure. Mutations in this region R96A and R96E caused the formation of defective chaperones (Lentze *et al.*, 2003a).

A single mutation L66A of Hsp16.6 from Cyanobacterium *Synechocystis* sp. PCC 6803 impaired the chaperone activity and oligomerization (Giese & Vierling, 2002). The L66A mutant was unable to protect luciferase at a ratio of 1  $\mu$ M Luc to 24  $\mu$ M, mutant protein. Also, the oligomer of L66A was less stable mostly forming sHSP dimer or trimer with 40-50 kDa, indicated that loss of oligomerization could be a factor for the reduced chaperone activity. On the other hand, increasing hydrophobicity by intragenic suppressors favored oligomerization. Three changes at

Asp-80 (to Val, His, or Asn) and V108L reversed the deleterious effect of L66A. This result also indicating the importance of hydrophobic patch interactions in the alpha crystallin domain.

Another mutagenesis study conducted on the small heat shock protein from *Mycobacterium leprae* Hsp18 also showed that replacing serine by proline at the position of 52 perturbed both structural stability and chaperone activity, possibly due to decrease in the surface hydrophobicity of S52P (Nandi *et al.*, 2013). Furthermore, S52P mutation altered the secondary structure of the protein such that content of the beta-sheet decreased slightly (5-8 %) and random coil structure increased. Moreover, oligomeric size from 29 subunits in WT was reduced to 9 subunits in the S52P mutant. The mutant protein became more prone to tryptic digestion due to the decreased stability (Nandi *et al.*, 2013).

It is known that P/A-G doublet in the ACD of non-metazoan sHSPs has direct roles in the subunit-subunit interactions. As demonstrated by crystal structure analysis, Gly 62 of *M.jannaschii* was found in the PG motif and involved in intersubunit contacts (Kim *et al.*, 1998). Site-directed mutagenesis and chemical cross-linking analysis of *Mycobacterium tuberculosis* Hsp 16.3 at the position of G59 (G59W) in the Pro-Gly doublet demonstrated that Gly59 location is very important for the oligomeric assembly of the sHSP (Fu & Chang, 2006). Trp59 was found to be buried at room temperature, but after heating at 60°C, the residue became exposed. However, exposure of Trp to surface upon heating did not increase the chaperone-like activity of the G59W mutant, implying that PG doublet did not have a direct role in substrate binding.

In addition to these, crystal structure analysis of *Xanthomonas axonopodis* small heat shock protein XaHspA gave insight into the mechanism of the complex formation with non-native substrate proteins and the structural organization of these complexes (Hilario *et al.*, 2011). It was found that XaHspA forms closed or open hexamers of 36 mers in solution. Also, several key residues, mostly polar, charged which could play an essential role in XaHspA monomer-monomer interaction to form a dimer were identified.

Recently, the dodecameric structure of a small heat shock protein from *Mycobacterium marinum* M was resolved by crystallization. It has three small heat shock proteins (*i.e.*, M1, M2, and M3)(Bhandari *et al.*, 2019). The pentapeptide, "GRLLP" of M3 interacts with a hydrophobic cleft between  $\beta$ 2 and  $\beta$ 7 strands. Residues Val92, Leu93 from  $\beta$ 7 strand of one monomer, and Gln90 from  $\beta$ 7 strand of a neighboring monomer were involved in hydrogen bond interactions with this peptide. Other residues were also identified, which are providing hydrophobic surface and forming the buried surface area at the interface. Overall result from this study indicated the important role of hydrophobic interactions in mediating substrate recognition and binding.

#### **1.1.7 Identification of the Important Residues on the ACD of sHSPs from Eukaryotes**

There are several reports on the single point mutations within the core alpha-crystallin domain of human small heat shock protein and their relation to various genetic diseases in humans. Such studies provided valuable information to understand the molecular mechanisms of impairment of the functions of sHSPs due to mutations. The mutation of arginine by glycine at position 120 in the ACD of human  $\alpha$ B-crystallin has been implicated in desmin-related myopathy. It was shown that R120G mutation disrupts the pair of salt bridges which results in a protein without chaperone activity and coaggregates with substrates in chaperone assays. The equivalent arginine mutation (R116C) in  $\alpha$ A-crystallin (HSPB4) and R127W in Hsp 27 caused zonular central nuclear cataracts and Charcot–Marie–Tooth disease, respectively (Clark *et al.*, 2011). In human Hsp22, mutation of the homologous residue K141E was correlated to the development of distal hereditary motor neuropathy. UV circular dichroism spectroscopy analysis showed that the quantity of the beta-strand was decreased, while unstructured region increased as a result of this point mutation. Due to destabilization of its secondary structure, the mutant was more prone to proteolysis. Thus, it was proposed that reduced chaperone-like activity

of Hsp22 K141E mutant due to destabilization of its structure could contribute to the development of neuromuscular disease (M. V Kim *et al.*, 2006).

K137 of the Hsp22 is another conservative residue and located at the  $\beta$ 5- $\beta$ 7 loop of the ACD. The K137E mutant sHSP had larger oligomers than WT. The NTD of the Hsp22 protein was thought to be very flexible and involved in interaction with ACD through hydrogen bonds (Thériault *et al.*, 2004). For this reason, mutation in K141 or K137 could change these interactions which will affect the oligomerization. More drastic changes were recorded in the double mutant K137I41E such that beta content was reduced; whereas, turns and the unstructured regions were increased. These results might imply that mutations in the  $\beta$ 5- $\beta$ 7 loop destabilizes the structure of Hsp22 and the protein proteolysis increases (Kasakov *et al.*, 2007).

Another mutation study in human HspB1 is associated with the peripheral distal neuropathies, which was investigated by generating G84R and L99M mutations at the border of the NTD-ACD (Nefedova *et al.*, 2013). The results showed that mutants of HspB1 were interacted with HspB6 weakly, forming small hetero-oligomers (*e.g.*, 100–120 kDa), as compared to wild type protein. The mutants also formed less stable large homo-oligomers than wild-type HspB1 due to changes in the quaternary structure. In addition, the mutation reduced chaperone-like activity of the HspB1 proteins.

The Gly84 residue of HspB1 is located at the short  $\beta$ 2 strand, which provides flexibility and freedom of the movement of the N terminal domain, thus could be crucial for oligomer assembly. When this Gly residue was replaced by charged arginine, the freedom of the N-terminal movement was diminished that in turn affected the oligomerization properties of the mutant protein. Besides, the L99M (near the hydrophobic core of ACD) replacement destabilized the intersubunit interactions with Arg140. Therefore, mutations that cause alterations in the structure can affect the stability of the HspB1 and its interaction with other small heat shock proteins as well as client proteins (Nefedova *et al.*, 2013).

Recent study showed a different function of HspB1 protein rather than molecular chaperone. HspB1 neuropathy causing mutations, R127W, S135F, and P182L,

decreased the autophagic flux by affecting interaction with autophagy receptor SQSTM/p62 and impairing the phagore formation. This might be one of the mechanism that mutations in HspB1 causes peripheral neuropathy (Haidar *et al.*, 2019; Muranova *et al.*, 2020).

In human  $\alpha$ A crystallin, it is known that R116C mutation is associated with an autosomal dominant congenital cataract. The homologous mutation in the rat  $\alpha$ A crystallin could give information regarding the molecular basis of the cataract formation since rat  $\alpha$ A crystallin had 98 % homology with human  $\alpha$ A crystallin. The results showed that rat  $\alpha$ A crystallin R116C mutant formed 2-4 times larger oligomers than WT. It was attributed to the loss of the salt bridge, which could be required for the proper assembly of the oligomers. Furthermore, impairment of the salt bridges might be by changing the protein conformational states led to decreased the chaperone activity (Nilufer P. Shroff *et al.*, 2000). Consistent with these findings, Bera *et al.*, (2002) provided evidence that positive charge at position 116 was crucial for the structural and functional integrity of  $\alpha$ A crystallin (Bera *et al.*, 2002a).

Another identified cataract causing gene mutation was in human  $\alpha$ A crystallin is G98R, which had the characteristics of an increased oligomeric mass, reduced chaperone function, and loss of structural stability. Phadte *et al.*, (2019) suggested that the loss of function is due to the introduced charged residue in place of G98 within the  $\beta$ 5 strand of the ACD.

On the other hand, the mutation in the conserved arginine residue (R105G, R109G, and R110G) in the ACD of *Drosophila melanogaster* DmHsp22 did not significantly affect the structure and chaperone function. It was stated that the R109G mutation did not affect formation of oligomers and dimeric interaction (Dabbaghizadeh *et al.*, 2017). R109 is equivalent to the R140-HspB1, R116-HspB4 and R120-HspB5. Mutation of these residues in human sHSPs leads to development of myopathies and cataract as discussed above.

In addition to the importance of charged residues in the ACD of sHSPs, hydrophobic residues are also important for chaperone function. Hydrophobic sites are considered as the putative substrate binding sites.

Structural analysis of sHSP16.9 from wheat showed that Val 4 and Phe 10 in the N-terminal domain of the protein cover the equivalent hydrophobic patches formed by Trp 48 and Phe 110 on the alpha-crystallin domain. These hydrophobic residues, Trp 48 and Phe110, contributed to the formation of a hydrophobic surface and putative substrate-binding site that must become exposed during subunit exchange (R. L. van Montfort *et al.*, 2001).

Shroff *et al.*, (2001) showed that in rat small heat shock protein  $\alpha\beta$  crystallin, a hydrophobic residue methionine at the position 68, when replaced with a hydrophilic threonine, chaperone activity and thermal stability were reduced. Thus, reduced chaperone function of M68T mutant was explained by decreased hydrophobicity. Also, Santhoshkumar and Sharma (2006) revealed the importance of the hydrophobicity in the chaperone function of human CRYAB. The H83A mutation resulted in increase in bis-ANS binding and chaperone activity.

Ghosh & Clark (2005) by using protein pin array technology identified subunit-subunit interaction sites in the human  $\alpha\beta$  crystallin. It was suggested that point mutations within the interactive domains could alter the oligomerization and chaperone activity of the protein. The N78D and N146D mutations in 76-FSVNLDVK-83 and 141-GVLTVNGP-149 regions led to the formation larger oligomeric ensembles. In vitro chaperone activity assay using two interactive peptides 73-DRFSVNLDVKHFS-85 and 131-LTITSSLSDGV-141 showed that these peptides are effective chaperones in preventing the aggregation of client proteins. Therefore, these regions could be important for subunit assembly and chaperone function (Ghosh *et al.*, 2005).

### **1.1.8 Chaperone Activity of $\alpha$ -Crystallin Derived Mini Chaperones**

$\alpha$ -Crystallin is the most abundant protein in the lens. It is made up of two subunits,  $\alpha$  and  $\beta$  whose sequences show high homology to each other and to other small heat shock proteins. The  $\alpha$ - and  $\beta$ -crystallin exist as a polydisperse oligomers with an average molecular mass of 800 kDa.  $\alpha$ A-Crystallin is present mainly in the lens. On the other hand,  $\alpha$ B-crystallin is present in several other tissues, including, the retina,

lens, heart, and kidney. These peptides act as molecular chaperones and prevent the aggregation of unfolded proteins. This chaperone function is thought to be essential in maintaining lens clarity during aging (Nahomi *et al.*, 2013).

Short peptides of  $\alpha$ A- and  $\alpha$ B-crystallin that function as molecular chaperones similar to the parent molecules are called “mini-chaperones”. Also they are described as “ $\alpha$ A-mini chaperone”, “mini- $\alpha$ A peptide”, “mini- $\alpha$ B peptide” (Raju *et al.*, 2016). The amino acid sequence of the mini chaperones are highly conserved among several sHSPs. Structure analysis showed that  $\alpha$ A-mini chaperone region were found in the  $\beta$ 3 -  $\beta$ 4 region of human  $\alpha$ A crystallin (Kannan *et al.*, 2012).

The chaperone sites in  $\alpha$ -crystallin were first identified by cross-linking between mellitin and  $\alpha$ A-crystallin. It was found that mellitin and hydrophobic probe (bis - ANS) binding sites were common and overlapped. A 19 amino acid peptide sequence "DFVIFLDVKHFSPEDLTVK" of  $\alpha$ A-crystallin contained these binding sites. This minimum sequence within  $\alpha$ A crystallin displayed anti-aggregation property similar to the  $\alpha$ A crystallin in vitro (K K Sharma, 2001).

In the last 20 years, the experiments performed to compare mini chaperone's functional activity with that of native proteins showed that mini chaperones suppressed the aggregation of partially unfolded substrate proteins by heat, and dithiothreitol (DTT) treatment (Raju *et al.*, 2016). Moreover, the effect of mini-chaperone ("DFVIFLDVKHFSPEDLTVK") within  $\alpha$ A crystallin on the stability and function of the cataract causing mutant G98R was tested (Raju *et al.*, 2012). Results showed that the mini-chaperone supplementation decreased the aggregation of client protein significantly. Mini peptide also increased thermal stability of the mutant protein at 43°C. The same research group, Raju *et al.*, (2014) also characterized the chimeric mini-chaperone peptide, which was generated by the addition of the solvent-exposed C-terminal 164-173 region of  $\alpha$ A-crystallin to the mini-chaperone sequence "DFVIFLDVKHFSPEDLT". The results showed that the addition of a C-terminal sequence enhanced the solubility of the chaperone-client protein complex. Chaperone activity of the chimeric protein against denaturing substrates was similar to that of  $\alpha$ A-mini-chaperone alone. Chimeric protein in

contrast to mini chaperone did not form fibrils when incubated at 42°C with shaking. Therefore, this type of modified peptide was found to be promising for clinical uses. The mechanism underlying protection of substrates by amyloid fibrils was investigated by Fukuhara *et al.*, (2012). The mini-peptide  $\alpha$ A71-88 exhibited anti-aggregating function for model substrate protein, ADH. However, another amyloid peptide A $\beta$  25-35, which is the amyloid fibril of the cytotoxic fragment of amyloid  $\beta$  protein, promoted the substrate aggregation. Measurement of its zeta potential showed that amyloid fragment of A $\beta$ 25-35 is as a cation (*i.e.*, +9 mV). Therefore, its interaction with negatively charged ADH resulted in formation of the aggregates/coagulates. Also, variants of  $\alpha$ A71-88 fragment possessing a cationic surface lost the ability to protect ADH against aggregation. Therefore, protection mechanism of amyloid fibrils might depend on colloidal stability.

In addition to these, alpha crystallin derived mini-chaperone  $\alpha$ AC(70–88), inhibited the amyloid formation of A $\beta$ (1-40) peptide when it was incubated in the presence of 5-fold-excess mini-chaperone (Tanaka *et al.*, 2008).

Mini chaperones also function as anti-oxidants. Histidine was found to be the critical residue and its substitution with Ala (corresponding to residue 79 in native  $\alpha$ A crystallin) eliminates its redox-suppression activity (Raju *et al.*, 2011). Moreover, it is well known that chaperone peptides functions as antiapoptotic agents and prevents apoptosis-mediated cell death. Both mini- $\alpha$ A and mini- $\alpha$ B peptides have ability to protect human fetal retinal pigment epithelial cells (hfRPE), from oxidation-induced cell death by inhibiting caspase-3 activation (Nahomi *et al.*, 2013). Similarly, it was reported that mini  $\alpha$ A reduced NaIO<sub>3</sub>- induced RPE cell apoptosis by repressing autophagy and ER stress (Zhang *et al.*, 2015).

Several studies have demonstrated that the mini-peptide of HSPB4,  $\alpha$ A66-80 has aggregating properties. Recently published data showed the importance of the isomerization of Asp76 of HSPB4 peptide  $\alpha$ A66-80 that results in elderly cataract, in the assembly of amyloid-like fibrils of  $\alpha$ A mini-peptide by TEM image analysis and ThT fluorescence assay. On the other hand, H79A mutation increased hydrophobicity, and ThT assay showed that this mutant variant of  $\alpha$ A66-80 (L-iso



form) had a high propensity to form amyloid-like fibrils. Thus, these results emphasized the hydrophobicity, as well as iso-form residues, were required for the formation of amyloid-like structures (Magami *et al.*, 2021).

### **1.1.9 Interaction of sHSPs with Other Proteins and Their Therapeutics Importance**

$\alpha$ A-crystallin is one of the important proteins in the eye lens. Its high concentration in the eye ensures the protection of the lens epithelial cells by reducing cell stress and preventing apoptosis. In addition to the protective function,  $\alpha$ A-crystallin maintains lens transparency. If its chaperone-like activity is impaired, it results in the formation of amyloid fibrils leading to the development of cataracts. Cataracts is a serious eye disease implicated by partial or complete loss of vision. Therefore, understanding its molecular mechanisms and development of methods for effective therapy is the issue of current research (Tikhomirova *et al.*, 2017). Biotechnology companies have started to produce mini-chaperone peptides for therapeutic uses in cataracts and other eye diseases. Several strategies are still in progress to optimize mini-chaperone peptides, including the addition of non-natural D-amino acid and modification of selective residues (Raju *et al.*, 2016). For example, intraperitoneally injected acetylated peptides are expected to cross the aqueous blood barrier, thus can reduce pathogenic apoptosis, which is an integral component of several eye diseases such as age-related macular degeneration uveitis, glaucoma, and diabetic retinopathy. On the other hand, experimental animal models showed that  $\alpha$ B crystallin has an anti-apoptotic function, so it is effective against inflammatory neurological damage, encephalomyelitis, multiple sclerosis, and cardiac ischemia-reperfusion injury. Therefore, anti-apoptotic peptides can potentially be used as therapeutic agents (Nahomi *et al.*, 2013).

It has been known that alpha-crystallin regulates actin dynamics and stabilization of the cytoskeletal elements. Disruption of the cytoskeleton due to stress (*e.g.*, heat shock) and disaggregation of the actin fibers cause myofibrillar myopathy, cataract, and desmin-related cardiomyopathy. Thus, the protein pin array identified the

sequence 57-APSWFDTG-64 of  $\alpha$ B-crystallin that selectively binds to partially unfolded actin and protects it from heat-induced aggregation (Ghosh *et al.*, 2009; Singh *et al.*, 2007).

In addition to these, many neurodegenerative diseases, including Alzheimer Disease, Huntington's disease, Parkinson's Disease, are characterized by abnormal aggregation of misfolded proteins (*i.e.*, amyloid  $\beta$  ( $A\beta$ ) and tau in AD;  $\alpha$ -synuclein in PD; huntingtin in HD, and prion proteins, PrPSc. In this respect, molecular chaperones may play pivotal roles to relieve the adverse effect of accumulated misfolded proteins (Maiti *et al.*, 2014). Raman *et al.*, (2005) have reported that  $\alpha$ B-crystallin inhibited fibril formation by binding to amyloid fibrils and protected neuronal cells from cytotoxicity. Furthermore, polyglutamine (poly Q) expansion causes the formation of fibrillar protein aggregates and neuronal cell death in Huntington's Disease (HD), and the spinocerebellar ataxias (SCAs). Experimental data showed that sHSPs decreased the neural toxicity of these fibrils in HD by suppression of reactive oxygen species and stimulation of autophagy (Robertson *et al.*, 2010). Therefore, such studies and others provided new insights into understanding the role of sHSPs in protein conformational diseases.

## 1.2 Scope and Aim of the Study

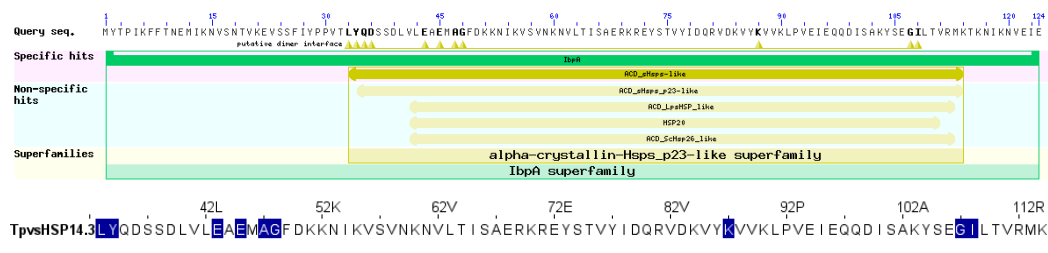
Although reports on sHSPs are ever increasing in recent years, molecular mechanisms of their structural and functional dynamics and substrate recognition have still remained open to questions. Due to their feature of polydispersity, crystallization studies are quite limited. Biochemical and structural studies provided some molecular level insights, but their mode of action on substrate specificity and protein interaction sites are not defined in detail. Most of the studies suggest that N-terminal domain of sHSPs plays major role in interaction with substrates. However, role of ACD of sHSPs in substrate binding remains to be further identified. Most of the research were conducted on human small heat shock proteins by site directed mutagenesis at the  $\beta$ 5- $\beta$ 7 loop of the ACD, which enlighten the residues important

for dimerization as well as their disease association (Kasakov *et al.*, 2007). In archaeal species, R107 residue of *M. jannaschii* sHSP, A102, K75, L109 and E33 residues of *S.solfataricus* sHSP were found to be important for its structural and functional integrity (L. Liu, Chen, Yang, Wang, *et al.*, 2015; Quinlan *et al.*, 2013; Wen *et al.*, 2010). The sequence of alpha crystallin domain of *Bradyrizobium japonicum* HspH was dissected by generating mutant variants (Lentze *et al.*, 2003b). Such limited studies targeting specific residues in the ACD although contributed valuable information to literature, are not sufficient to enlighten its role in substrate recognition and interaction.

The model organism, whose sHSP is being studied in this study is *Thermoplasma volcanium* (kingdom Euryarchaeota, order Thermoplasmatales, genus Thermoplasma). To date, two species, *Thermoplasma volcanium* and *Thermoplasma acidophilum* have been reported for this genus (Yasuda *et al.*, 1995). *T. volcanium* was first isolated from submarine and continental solfataras on Vulcano Island in Italy. The pH for growth is ranging between 1 and 4 and growth temperature is varying between 33°C and 67°C for this organism. However, the best growth is at pH 2.0 and at a temperature of 60°C (Auernik *et al.*, 2008). The cell of *T.volcanium* are enveloped by merely a plasma membrane and, therefore, the cells lack a definite shape. The shape/morphology of *T.volcanium* varies between pleomorphic to spherical, where the diameter of the sphere is between 0.6 –1.0 µm. Several flagella are present to assist its motility (Faguy *et al.*, 1996).

In this thesis project, by the aid of the site-directed mutagenesis method, specific residues in the core alpha-crystallin domain of *Tpv* sHSP 14.3 were replaced by opposite characteristics of amino acids to analyze their effects on structure and function of this protein. One of our particular goals is to determine critical residues, which are important in interacting with the model protein substrates and keeping them in a soluble/ partially folded state during heat stress. Additional goal is further improvement of substrate binding capacity of *Tpv* sHSP 14.3 by mutagenesis. Residues chosen for mutagenesis are predicted to be on the dimer interface as found

by the National Center for Biotechnology Information's Conserved Domain Database (CCD) search (Figure 1.6).



**Figure 1.6. Putative dimer interface points of *Tpv* sHSP 14.3.** Residues targeted for mutation in this thesis study are marked by blue color.

Among the targeted residues, L33, Y34, E43 and E45 are located at the beginning of the ACD, which is near to the dimerization loop  $\beta 6$  in the 3D structure. Two residues A47 and G48 are found in the loop that connects two antiparallel  $\beta 3$ - $\beta 4$  strands. MSA analysis showed that these residues are the components of the P/A-G motif, which is highly conserved in non-metazoan sHSPs (Figure 1.7). Previously, a single study in bacterial small heat shock protein, Hsp16.3 of *M.tuberculosis*, showed the importance of this motif in subunit interaction (Fu & Chang, 2006). Therefore, role of this highly conserved motif with respect to the structure and function of the sHSPs is needed to be further investigated and enlightened. Besides, another targeted residues, G107, and I108, are located in the highly conserved G-v-L motif (Figure 1.7). Previously, in this motif highly conserved leucine was exchanged by aspartic acid in small heat shock protein Hsp 14.1 of *S.solfataricus* (Wen *et al.*, 2010)(Figure 1.7). In this study, G residue is targeted for the mutagenesis and and characterization. Therefore, our study contributes novel information about the role of two motifs.

In addition, the mutation point, 87th residue was found in the  $\beta 7$  strand of the *Tpv* sHSP 14.3 protein. It has been reported for a non metazoan sHSP that  $\beta 7$  strand included potential substrate binding sites (Jaya *et al.*, 2009). Also, *Tpv* K87E equivalent mutation in human is associated with certain genetic disorders (*e.g.*, cataracts, myopathy). Homologous residues although were studied in human CRYAB and sHSP of hyperthermophilic archaeon *M. jannashii*, they are not characterized in other organism including archaea (Figure 1.7). Therefore, our study

will provide additional information about structural and functional importance of this conserved residue.

Double mutations were also generated (Y34FG48E, G48EI108K) in order to evaluate additive effects of some of the point mutations on the chaperone function of the *Tpv* sHSP 14.3 and oligomeric structure.

To assess the chaperone function of the WT *Tpv* sHSP 14.3 and its mutant variants, thermal aggregation measurements and chaperone activity assays were performed using two model substrate proteins: CS (citrate synthase) and Yeast Alcohol Dehydrogenase (ADH). 3-D model structure analysis (*i.e.*, -inter and -intra molecular interactions), contact map analysis, thermodynamic stability analysis and hydrophobic surface analysis of the WT *Tpv* sHSP14.3 and its ACD mutants were comparatively studied.

Final objective of this project is to design novel mini chaperone peptides derived from ACD of *Tpv* sHSP 14.3. The protection effects of these mini chaperones were investigated by using model substrate proteins (CS and ADH) against their thermal aggregation. Results of this research will be helpful for consideration of this novel mini-chaperones as alternative agents in chaperone therapy.

|       |  |     |
|-------|--|-----|
| Tvo   | -----M--YTPI--KFF-----TNEMIKNVSNTVKEVSSFLYPPVTL              | 34  |
| Sto   | -----M--YYL-----GKELQKRSEELSRGFYELVYPPVDMY-                  | 30  |
| Sso   | -----MM--NVI-----MREIGKKLDELRSREFYESVI PPIDMY-               | 31  |
| Mja   | -----MFGRDPFDSL--FERMFKEFFATPMTGTMIQ--SSTGIQISGKGFMPISII-    | 48  |
| CRYAB | MDIAIHHHPWIRRRPFPPHSPSRLFDQFFGEHLLSDFP----TSTSLSPFYLRPPSFLR  | 56  |
|       | : : : : : *  |     |
| Tvo   | -----QDSSDLVL  | 75  |
| Sto   | -----EEGGYLVVVADLAGFNKEKIKARVSGQNELI IEAEREITEP----GV        | 73  |
| Sso   | -----EEGGELVVVADLAGFNKDKISVRLSAQNELI INAEREIQYI----GT        | 74  |
| Mja   | -----EGDQHIKVIAWLPGVNKEDI ILNAVG-DTLEIRAKRSPLMITESERI        | 94  |
| CRYAB | APSWFDTGLSEMRLEKDRFSVNLVVKHFSPEELKVKVLG-DVIEVHGKHEERQD--EHGF | 113 |
|       | : : : . . . : : : . . .                                      |     |
| Tvo   | VY-IDQRVDKVVVVKLPVEIEQ-QDISAKYSE                             | 124 |
| Sto   | KY-LTQRPKYVRKVIRLPYNVAKDAEISGKYENGLTIRIPIAGTSV----IKIE----   | 123 |
| Sso   | KY-ATQRPLKIHKVIRLPVKVQRDSQVTAKYENGLTIRIPIVEGVS----IRIE----   | 124 |
| Mja   | IYSEIPEEEIYRTIKLPATVKE-ENASAKFENGLVSVILPKAESSIKK-GINIE----   | 147 |
| CRYAB | -----ISREFHRKYRIPADVPLTITSSLSDDGVLTVNGPRKQVSGPERTIPIITREEKP  | 167 |
|       | . : : * : : . : * : * : :                                    |     |
| Tvo   | -----  | 124 |
| Sto   | -----  | 123 |
| Sso   | -----  | 124 |
| Mja   | -----  | 147 |
| CRYAB | AVTAAPKK   | 175 |

**Figure 1.7. Multiple sequence alignment of *Tpv* sHSP14.3 among the sHSPs sequences from different organisms by CLUSTAL O multiple sequence alignment server.** The species names corresponding to each accession number are as given: CBY78065.1: Tvo: *Thermoplasma volcanium* GSS1, WP\_010979712.1: Sto: *Sulfolobus tokodaii*, WP\_009989320.1: Sso: *Saccharolobus solfataricus*, WP\_010869783.1: Mja: *Methanocaldococcus jannaschii*, sp|P02511.2|: CRYAB\_HUMAN. Mutation positions are marked by blue color.



## CHAPTER 2

### MATERIALS AND METHODS

#### 2.1 Materials

##### 2.1.1 Chemicals, Enzymes, and Kits

Ampicillin, agarose, glycerol, TRIS, magnesium chloride ( $\text{MgCl}_2$ ), potassium bicarbonate ( $\text{KHCO}_3$ ), coenzyme A, DTNB, oxaloacetic acid, pig heart citrate synthase, alcohol dehydrogenase,  $\beta$ -NAD, TEMED, acrylamide, bis-acrylamide, ammonium per sulfate, ethidium bromide, agar, bromophenol blue, coomassie brilliant blue R, Tricine, 6-Aminohexanoic acid, methanol, polyethylene glycol (PEG), acetic acid were purchased from Sigma Aldrich® (Missouri, USA).

Tryptone medium, sodium chloride ( $\text{NaCl}$ ), N-Z amine, imidazole and glucose were bought from Fluka Analytical, St. Gallen, Switzerland.

Yeast extract was bought from Difco (Detroit, USA).

DMSO, absolute ethanol, SDS, ammonia solution (25%), EDTA, sodium hydroxide ( $\text{NaOH}$ ), coomassie Brilliant Blue G-250, potassium dihydrogen phosphate ( $\text{KH}_2\text{PO}_4$ ), magnesium sulfate heptahydrate ( $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$ ), di-sodium hydrogen phosphate heptahydrate pure ( $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ ), 2mercaptoethanol, HEPES were purchased from Merck; Germany, Darmstadt.

Disodium hydrogen phosphate-2 hydrate ( $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ ) and sodium dihydrogen phosphate extra pure ( $\text{NaH}_2\text{PO}_4$ ) were obtained from Riedel-de Haën (Germany).

IPTG was bought from Fermentas (Waltham, Massachusetts, USA).

Ponceau S was purchased from Applichem, Darmstadt, Germany.

Restriction enzymes, *NdeI*, *BamHI*, *HindIII*, *ClaI*, *EcoRI*, *BclI*, *BglII*, *Sall*, *SacI* were purchased from ThermoFisher Scientific (Waltham, MA) and NEB (New England Biolabs, Ipswich, Massachusetts).

QIA Prep Spin Miniprep Kit and QIAquick Gel Extraction Kit were bought from QIAGEN Inc., Valencia, USA. DNA Ligation Kit was purchased from Novagen (Madison, WI).

The QuikChange Site-Directed Mutagenesis Kit was from Agilent Technologies Inc., USA and GeneArt Site Directed Mutagenesis Plus Kit was bought from ThermoFisher Scientific, Waltham, MA.

Gel Filtration High Molecular and Low Molecular Weight Calibration Kits were from GE Healthcare (USA).

HiTrap Q column was from Amersham Biosciences, Piscataway, NJ, U.S.A.

### **2.1.2 Buffer Solutions**

All buffer solutions and their components are listed in Appendix A. Buffers were prepared by using ultra pure water taken from Thermo Scientific Barnstead Smart2 Pure Water Purification System. Solutions were sterilized in an autoclave (Alp Co., Ltd.).

### **2.1.3 Molecular Size Markers and Plasmid Vectors**

PageRuler Prestained Protein Ladder for estimation of the molecular weight of protein samples on SDS-PAGE, and DNA ladders i.e., O'Gene ruler DNA Ladder mix, O' range ruler 200 bp DNA ladder, and Lambda DNA/*EcoRI* plus *HindIII* Marker were purchased from Thermo Scientific (Waltham, MA). Images of ladders are given in the Appendix B.

Expression vector pET21a (+) was bought from Novagen, Madison, WI. Vector map is given in the Appendix C.

pUC19 control plasmid was purchased from NEB (New England Biolabs, Ipswich, Massachusetts).



## **2.2 Strains and Growth Media**

*Thermoplasma volcanium* GSS1 strain was used in this thesis project as the model organism. BL21(DE3) *E.coli* Competent Cells and T7 Express *E.coli* cells were purchased from NEB (New England Biolabs, Ipswich, Massachusetts). NovaBlue Singles Competent cells were from Novagen, Madison, WI. XL1-Blue Supercompetent cells were supplied from Agilent's QuikChange Site-Directed Mutagenesis Kit. One Shot<sup>R</sup> MAX Efficiency<sup>R</sup> DH5 $\alpha$ <sup>TM</sup>-T1R competent *E. coli* cells were provided by GeneArt Site Directed Mutagenesis Plus Kit.

### **2.2.1 Growth Medium and Culture Conditions**

*Thermoplasma volcanium* GSS1 strain was grown in standard Volcanium Medium (pH 2.6), containing 0.5 % glucose and 0.1 % yeast extract at 60°C without shaking (Kocabiyik & Özel, 2007).

Recombinant WT and mutant strains of *E.coli* were grown on LB agar medium containing ampicillin (100  $\mu$ g/ml) at 37°C overnight. The cultures were transferred to fresh media by subculturing with 30 days intervals.

## **2.3 Methods**

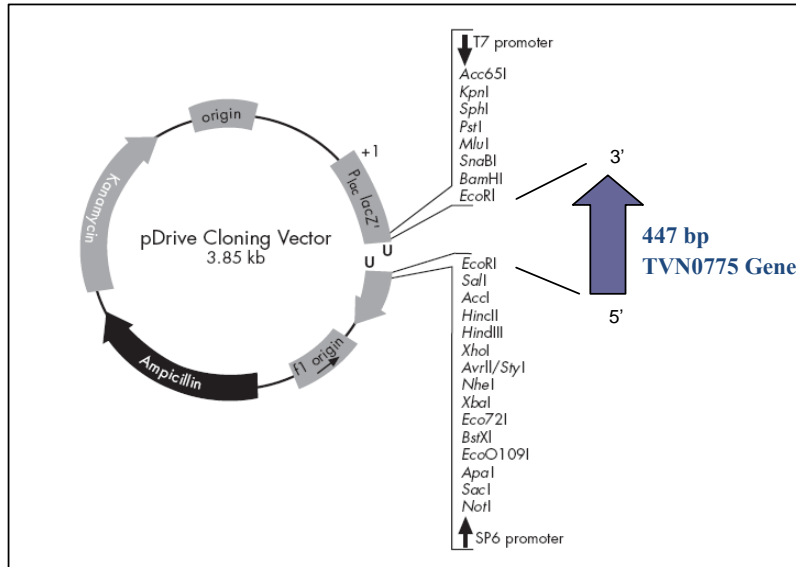
### **2.3.1 Expression of Recombinant *Tpv* sHSP14.3 Protein from its Own AUG**

In the first stage of this thesis study, expression of the recombinant *Tpv* sHSP14.3 Protein from its own AUG was attempted. For this purpose, pET21a expression vector system was used.

#### **2.3.1.1 Isolation of the Target DNA for Amplification**

The gene encoding the Hsp20/ $\alpha$  crystallin family protein in the *Thermoplasma volcanium* genome (locus name TVG\_RS04180, sequences 790978..791352 ) was

cloned into pDrive Cloning Vector (Qiagen) as described previously (Kocabiyık & Aygar, 2012) and named as *Tpv* sHSP 14.3. Recombinant plasmid diagram is shown in Figure 2.1.



**Figure 2.1. Recombinant plasmid diagram of cloned *tpv*-sHSP 14.3 gene in pDrive Cloning Vector**

This recombinant plasmid was isolated from *E.coli-tpv* 14.3 cells by following the instructions of the QIA prep Spin Miniprep Kit (QIAGEN Inc., Valencia, USA). Over-night culture of the *E.coli-tpv* 14.3 cells in 5 ml of liquid medium (LB) containing 100 µg/mL of ampicillin was grown at 37°C with vigorous shaking (181-182 rpm). The bacterial cells were harvested by centrifugation at 4600 rpm for 15 min in Thermo Scientific™ Labofuge™ 200 Centrifuge. The cell pellet was resuspended in 250 µL of Buffer P1 and transferred into a 1.5 mL of eppendorf tube. Then, 250 µL of Buffer P2 was added and mixed. The cell lysis was occurred after 5 min incubation. Next, 350 µL of neutralizing solution was added and mixed immediately. After centrifugation at 13000 rpm for 10 min, the supernatant was transferred into a spin column and centrifuged for 1 min at 13000 rpm. After washing first with 500 µL of PB and then with 750 µL of PE, the flow through was discarded. Finally, the plasmid DNA was eluted to a clean 1.5 ml eppendorf tube by adding 50 µL of EB and centrifugation. The cloned gene was excised from the purified plasmid

by restriction enzyme digestion using *EcoRI* restriction enzyme (ThermoFisher Scientific, Waltham, MA), and visualized by agarose gel electrophoresis.

### **2.3.1.2 Insert and Vector Preparation**

The *Tpv* sHSP 14.3 gene was amplified by PCR (Gene Cyclor, Techne Inc., NJ, USA) using forward primer (5'-T GAG CAT ATG TAT ACA CCC ATA AAG TTC TTT ACG-3') whose recognition site contains the ATG start codon and a reverse primer (5'-T GAG GGA TCC CAC CCA ATC ACA TCA AGC ATA C-3') including the *BamHI* recognition site, and Taq DNA polymerase enzyme (Fermentas, UK). The PCR amplicons were gel purified by using QIAquick Gel Extraction Kit (QIAGEN Inc., Valencia, USA) following the kit protocol. Clean PCR products were visualized after electrophoresis on 1% agarose gel with Bio-Print imaging system (France).

PCR amplified products and the expression vector pET21a (+) (Novagen, Madison, WI) were digested using enzymes *NdeI* and *BamHI* (New England Biolabs, Ipswich, Massachusetts). The digested insert and vector were then purified to clean up DNA using QIAquick Gel Extraction Kit (QIAGEN Inc., Valencia, USA). The samples then were loaded on an agarose gel to verify recovery.

### **2.3.1.3 Ligation and Transformation into Chemically Competent *E.coli* (BL21(DE3)) Cells**

The insert was ligated into the linearized pET21 vector by using a DNA Ligation Kit (Novagen, Madison, WI) at 4°C for overnight. Transformation was performed according to the manufacturer's instructions (pET System Manual) using BL21(DE3) competent *E.coli* cells (NEB). Before ligation, efficiency of the competent cells was checked by using control plasmid pUC19 (NEB) following the "High-efficiency transformation protocol" mentioned in the manual (catalog #C2527H) (New England Biolabs, Ipswich, Massachusetts). The transformants were then grown on LB agar

containing an ampicillin overnight at 37°C. The flow chart of the sub-cloning experiment is shown in the Figure 2.2.

#### **2.3.1.4 Screening for Recombinant Colonies by Plasmid Isolations and Restriction Enzyme Digestions**

The plasmid DNA was extracted from randomly selected transformant colonies using QIA Prep Spin Miniprep Kit (Qiagen) and checked to see if the ligation was successful by cutting with *Nde*I and *Bam*HI restriction enzymes. Further characterization of the putative recombinants was performed by using different combinations of restriction enzymes, which were chosen referring to restriction map of the recombinant plasmid to confirm the size of the insert.

#### **2.3.1.5 Confirmation of Cloning by DNA Sequencing**

To verify the recombinant colonies by sequencing of the cloned genes, high quality DNA with respect to purity and concentration is required. For this purpose, plasmid DNA of well characterized putative recombinant colonies was isolated using QIA prep Spin Miniprep Kit (Qiagen). Their purity and concentration were checked by Picodrop (Picopet 01, Picodrop Ltd. UK) and agarose gel electrophoresis. The selected plasmid DNA samples were sent for sequencing to Genscript (Piscataway, New Jersey, USA), and Oligomer (Ankara, Turkey). After sequence verification, plasmids of *Tpv* sHSP14.3 was transformed into NovaBlue Competent cells (Novagen, Madison, WI), which do not contain any source of T7 RNA polymerase. For this reason, these cells are ideal for maintenance of the recombinant plasmids without expression of the cloned gene.

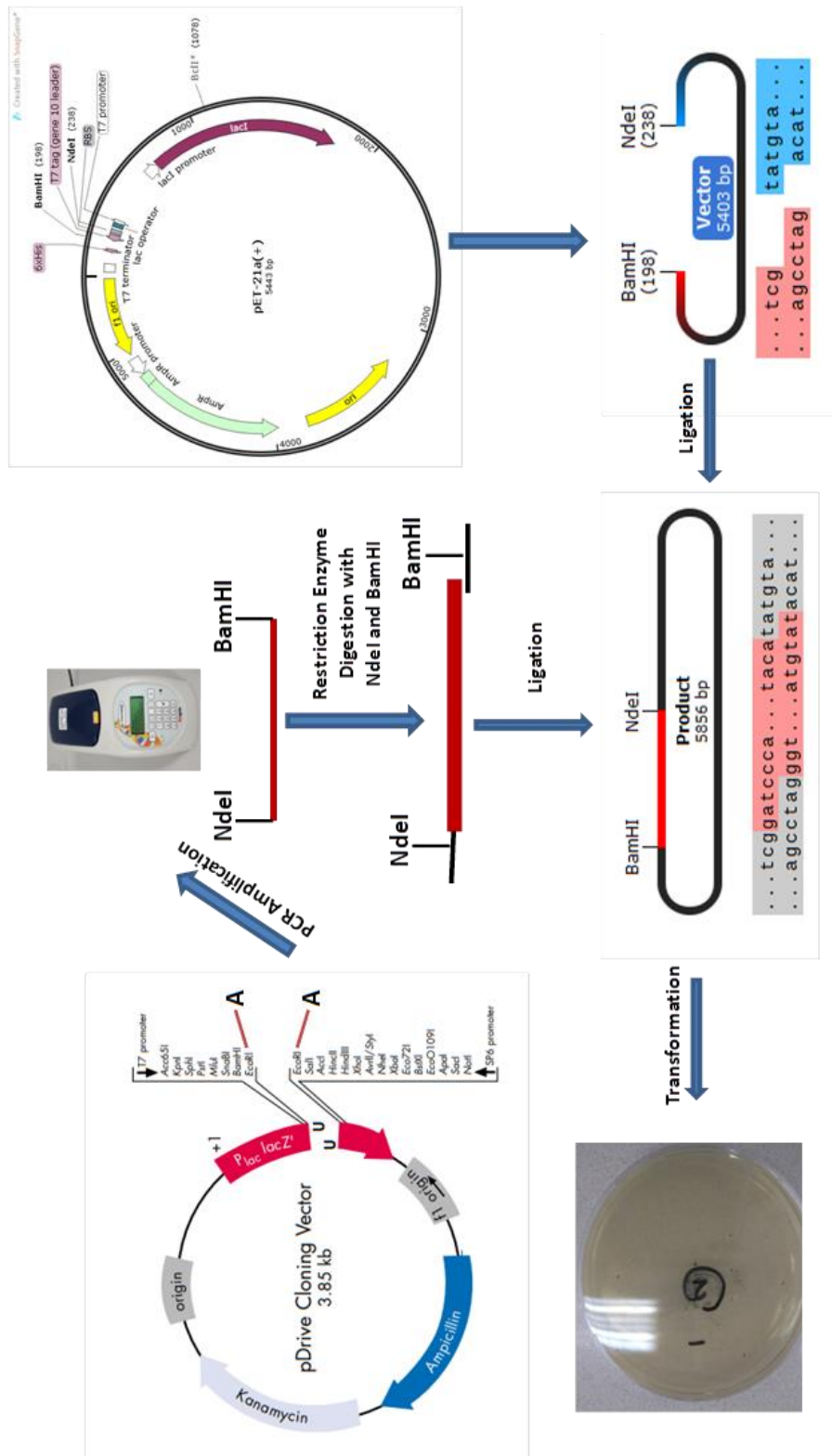


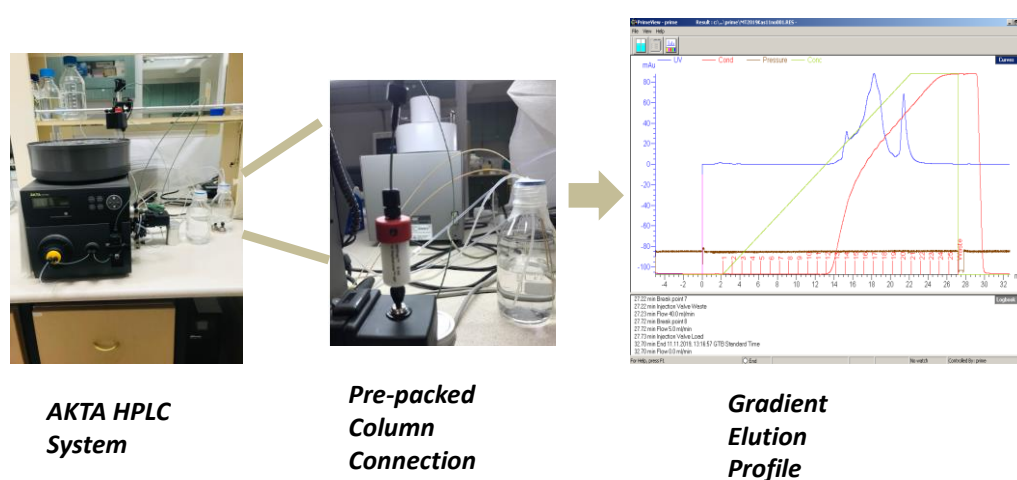
Figure 2.2. Flow chart of the sub-cloning experiment

### 2.3.2 Overexpression and Purification of the Wild Type *Tpv* sHSP 14.3

In the pET21 vector, the expression of the recombinant protein is induced by the supplementation of the growth medium with IPTG since the cloned gene is regulated by the T7 promoter and lac operator (See Figure 2.2). Expression of the cloned gene was achieved following the protocol given in the pET System Manual with some modifications. A single transformed BL21(DE3) colony, named as *E.coli* pET21\_tvshSP2 was inoculated into 10 mL LB medium supplemented with ampicillin (100 µg/mL) followed by shaking at 200 rpm overnight at 37°C. Then, 4 ml of overnight culture was transferred to 100 mL of LB medium containing ampicillin in a 500 ml Erlenmeyer Flask. The culture was grown at 37°C with vigorous shaking 300 rpm until the optical density at 600 nm ( $A_{600}$ ) reached 0.5-0.7 recorded by UV-Visible spectrophotometer (Schimadzu UV-1601A, Kyoto, Japan). Afterwards, expression of wild type *Tpv* sHSP 14.3 was induced by addition of isopropyl thio-β-D-galactoside (IPTG) to a final concentration of 1 mM. After induction, cells were harvested by centrifugation at 5000 x g for 20 min at 4°C (JOUAN SA, Herblain-France). The cell pellet was resuspended in 10 mL lysis buffer. Cells were broken by sonication (Sonics and Materials, CT, USA) followed by centrifugation at 20000 xg for 20 min at 4°C (Sigma 3K30 Centrifuge, Sigma Chemical Co., St. Louis, USA). The soluble fraction called cell-free extract was analyzed by SDS-PAGE. To study heat stability of the expressed protein, cell-free extract was heated at temperatures varying from 60°C to 80°C. The denatured protein was removed by centrifugation at 12000 xg for 1 hour at 4°C. The cleared cell lysate then was stored at -20°C until use.

In order to purify the wild type *Tpv* sHSP 14.3 by HPLC, cell culture volume was scaled up to 500 mL and the protocol explained above was followed. The heat-treated cell extract in lysis buffer was replaced by start buffer (20 mM Tris, pH 8.41) for anion exchange chromatography using VivaSpin 20 mL, 5K filter device (Sartorius, Göttingen, Germany). The WT sHSP 14.3 protein was filtered through a 0.45 µm filter (EMD Millipore, Burlington, Massachusetts, USA) and loaded to a HiTrap Q

column (Amersham Biosciences, Piscataway, NJ, U.S.A) equilibrated with the Start Buffer (20 mM Tris, pH 8.41). The bound protein was eluted at a flow rate of 5.0 mL/min in a linear gradient of NaCl (0-1 M) in start buffer and was fractionated by the AKTA prime HPLC system (Amersham Pharmacia Biotech). The eluted fractions were analyzed by running SDS/PAGE and OD<sub>280</sub> measurements. The samples containing the target protein was pooled, and concentrated using Vivaspin concentrator with a 5000 molecular weight cutoff (Sartorius, Göttingen, Germany). The HPLC system we used shown in the Figure 2.3.



**Figure 2.3.** HPLC purification system with a pre-packed anion exchange column.

### 2.3.3 Site Directed Mutagenesis

Site directed mutagenesis was carried out to make amino acid substitutions at the specified positions in the *Tpv* sHSP gene. The QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies Inc., USA) and GeneArt Site Directed Mutagenesis Plus Kit (ThermoFisher Scientific, Waltham, MA) were used to introduce the single and double mutations, respectively, according to the manufacturer's instructions.

### 2.3.3.1 Constructing Single Site Mutants

Nine recombinant single mutants which we designed, L33S, Y34F, E43V, E45G, A47D, G48E, G107A, G107D, and I108K were constructed by PCR-directed mutagenesis using Quick Change II Site-Directed Mutagenesis Kit. This kit is highly convenient that allows site-specific mutation in any double-stranded plasmid without requirement of specialized vectors, unique restrictions sites, multiple transformations or in-vitro methylation treatment steps. The mutagenesis experiment is consist of three major steps including mutant strand synthesis, *Dpn* I digestion of template and transformation. Recombinant plasmid pET21\_tvshSP2 was used as the source of cloned *Tpv* sHSP14.3 gene in the mutagenesis experiments. The plasmid DNA was isolated using the QIA Prep Spin Miniprep Kit (Qiagen). High quality plasmids with respect to purity and concentration as determined by Picodrop as well as restriction enzyme analysis were selected to be used in mutagenesis experiments. The mutagenic oligonucleotide primers were designed according to the Kit's instructions (Table 2.1). Primers were prepared by dilution to final concentration of 1 $\mu$ g/ $\mu$ L. The forward and reverse oligonucleotide primers, each complementary to opposite strands of the vector, were extended by PCR using PfuUltra HF DNA polymerase. The reaction mixture for thermal cycling included 5  $\mu$ L of 10 X reaction buffer, 5-50 ng of double stranded DNA template, 125 ng of forward, reverse primers, 1 $\mu$ L of dNTP mix and ddH<sub>2</sub>O in a final volume of 50  $\mu$ L. 1  $\mu$ L of PfuUltra HF DNA polymerase was added to this mixture. The thermal cycling parameters for the QuickChange Site-Directed Mutagenesis Method was followed (Table 2.2). Extension of the oligonucleotide primers generates a mutated plasmid DNA containing staggered nicks. The PCR product was digested with *Dpn* I endonuclease at 37°C for 1 hour to degradate the methylated parental DNA and to select mutation-containing synthesized DNA. The nicked vector DNA containing the desired mutations was then transformed into XL1-Blue supercompetent cells for nick repair. The transformants were grown on LB agar media containing ampicillin (100  $\mu$ g/ml). The plates were incubated overnight at 37°C and the colonies were counted in the



following day to calculate the transformation efficiency. Overview of the QuikChange II site-directed mutagenesis method is shown in Figure 2.4.

**Table 2.1. List of the mutagenic oligonucleotide primers**

| Mutation | Sequence of the Mutagenic Oligonucleotide Primers            |
|----------|--|
| L33S-f   | 5'-cca cca gtc acg <b>tea</b> tat caa gat agc-3'             |
| L33S-r   | 5'-ctatcttgatat <b>gac</b> gtgactgggtgg-3'                   |
| Y34F-f   | 5'-cca cca gtc acg tta <b>ttt</b> caa gat agc tct-3'         |
| Y34F-r   | 5'-agagctatctt <b>gaa</b> ataacgtgactgggtgg-3'               |
| E43V-f   | 5'-gat ctg gta ttg <b>gta</b> gca gaa atg gcc-3'             |
| E43V-r   | 5'-ggcatttctg <b>tacca</b> ataccagatc-3'                     |
| E45G-f   | 5'-gta ttg gaa gca <b>gga</b> atg gcc ggg ttt gac-3'         |
| E45G-r   | 5'-gtcaaacceggccatt <b>ctc</b> tgttccaatac-3'                |
| A47D-f   | 5'-ttg gaa gca gaa atg <b>gac</b> ggg ttt gac aag-3'         |
| A47D-r   | 5'-cttgcaaacce <b>gtc</b> catttctgttccaa-3'                  |
| G48E-f   | 5'-gaa gca gaa atg gcc <b>gag</b> ttt gac aag aaa aac-3'     |
| G48E-r   | 5'-gtttttctgtcaaa <b>ctc</b> ggccatttctgttc-3'               |
| G107A-f  | 5'-gct aag tat agt gaa <b>gcc</b> ata ctt aca gtt ag-3'      |
| G107A-r  | 5'-ctaactgtaagtat <b>ggc</b> ttcactatacttagc-3'              |
| G107D-f  | 5'-gct aag tat agt gaa <b>gac</b> ata ctt aca gtt ag-3'      |
| G107D-r  | 5'-ctaactgtaagtat <b>gtc</b> ttcactatacttagc-3'              |
| I108K-f  | 5'-gct aag tat agt gaa ggc <b>aaa</b> ctt aca gtt aga atg-3' |
| I108K-r  | 5'-cattctaactgtaag <b>ttt</b> gccttactatacttagc-3'           |

**Table 2.2. Cycling parameters for the QuickChange Site-Directed Mutagenesis Method**

| Segment | Cycles | Temperature | Time                           |
|---------|--------|-------------|--------------------------------|
| 1       | 1      | 95°C        | 30 seconds                     |
| 2       | 12-18  | 95°C        | 30 seconds                     |
|         |        | 55°C        | 1 minute                       |
|         |        | 68°C        | 1 minute/kb of plasmid length* |

\* For example, a 5-kb plasmid requires 5 minutes at 68°C per cycle.

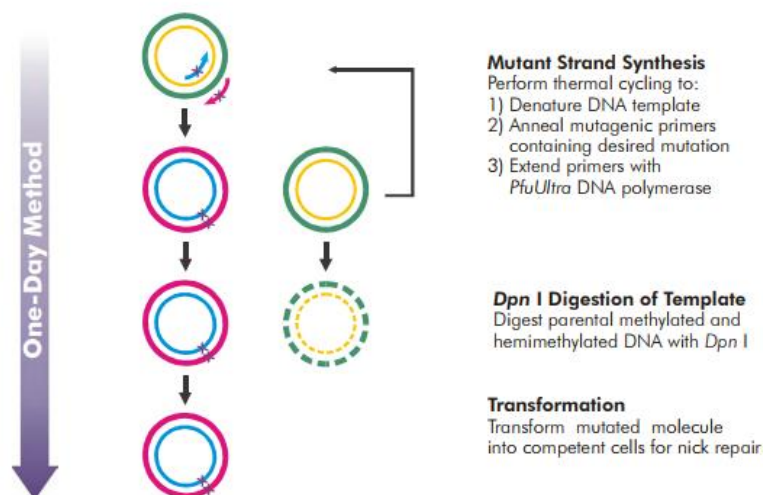
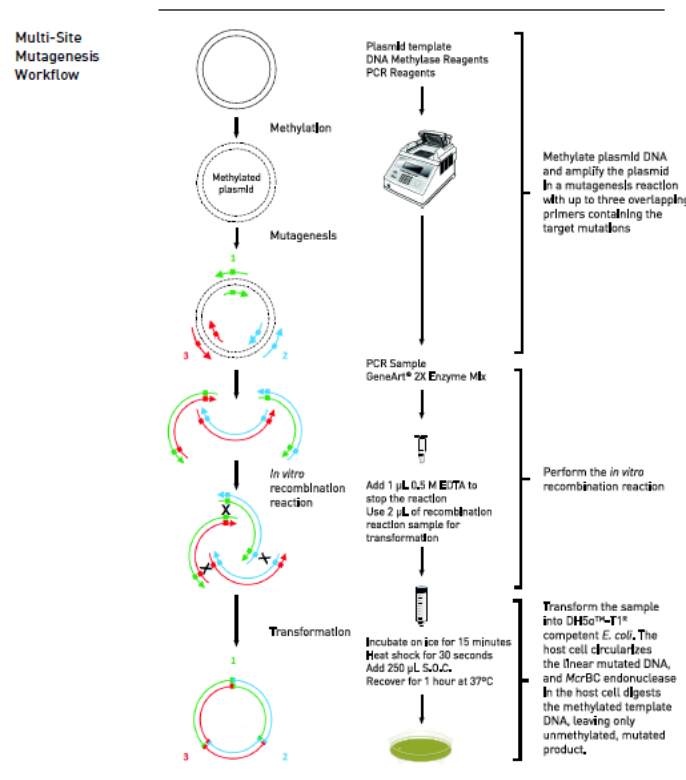


Figure 2.4. Overview of the QuikChange II site-directed mutagenesis method

### 2.3.3.2 Construction of Double Site Mutants

To generate double mutations at the specified positions in the *Tpv* sHSP14.3 gene (*i.e.*, G48E and I108K), site directed mutagenesis was carried out by using GeneArt Site Directed Mutagenesis Plus Kit (ThermoFisher Scientific, Waltham, MA). The mutagenic primers were designed according to the Kit's instructions. The master reaction mix was made up of reaction buffer, enhancer, source plasmid, DNA methylase, 25X SAM, Accuprime Pfx and water. This mixture was divided equally to two tubes, into which the mutagenic primer mixes were added. Each primer mix contained a forward or a reverse primer that corresponds to a separate mutation site, to prevent the primers from annealing to each other. Methylation of plasmid DNA and amplification of the plasmid occurred in a mutagenesis reaction with primers containing target mutation points. PCR was carried out in the thermal cycler (Gene Cyclor, Techne Inc., NJ, USA) considering the parameters suggested in the Kit protocol. After amplification, the PCR products were mixed together in the recombination reaction supplemented with Gene Art 2X enzyme Mix. Following EDTA addition to stop recombination reaction and then immediate transformation

into One Shot<sup>®</sup> MAX Efficiency<sup>®</sup> DH5 $\alpha$ <sup>™</sup>-T1R competent *E. coli* cells, mutagenesis procedure was completed. After spreading of the culture on LB agar media containing ampicillin, the plates were incubated overnight at 37°C. Flow chart for the mutagenesis is shown in Figure 2.5.



**Figure 2.5. Overview of the GeneArt site-directed mutagenesis method**

For verification of the mutations by sequencing (Genscript Biotech, Piscataway, New Jersey, USA), three putative colonies were selected randomly from the agar plates for each mutation. Their plasmids were isolated using QIA prep Spin Miniprep Kit (Qiagen). The purity and concentrations of the plasmid samples were checked by Picodrop. Also, these plasmids were double digested with *Bgl*III+*Hind*III (ThermoScientific) restriction enzymes to excise cloned genes.

### **2.3.4 Transformation into the Competent *E.coli* Cells for Protein Expression**

For high efficiency protein expression, mutant plasmids from XL-1 Blue cells and DH5 $\alpha$ <sup>TM</sup>-T1R cells used in mutagenesis were transformed into competent *E.coli* BL21(DE3) or T7 Express *E.coli* cells. In transformation experiments, out of three mutant strains of each mutant group, only one was selected for plasmid isolation to be used in transformation. Plasmid DNA was extracted using QIA prep Spin Miniprep Kit (Qiagen). Concentration and purity of the plasmids were determined by PicoDrop. Also, plasmids were checked by double digestion with *Bgl*III and *Hind*III restriction enzymes in order to confirm presence of the gene. Single mutants, L33S, Y34F, E43V, E45G, A47D, G48E, and I108K were transformed into BL21(DE3) competent *E. coli* cells, which were prepared according to the TSS protocol, described by (Chung *et al.*, 1989). Similarly, two more single mutants, G107A and G107D were transformed into NEB's *E.coli* BL21 competent cells (C2527H) according to the manufacturer's protocol (New England Biolabs, Ipswich, Massachusetts). In addition, T7 express competent cells (New England Biolabs, Ipswich, Massachusetts) which is useful for robust overexpression of recombinant proteins, were used in transformation of the mutant plasmids of G48E/I108K double mutation, G48E and E43V single mutations. The double mutant Y34F/G48E was obtained as a result of a Y34 to F substitution in the mutagenesis for planned G48E single mutation. This plasmid from this mutant was also transferred into T7 express competent cells. After transformation, the plasmid DNA was isolated from randomly selected two colonies of each mutant group and characterized by restriction digestion using *Xba*I and *Hind*III restriction enzymes.

### **2.3.5 Expression and Initial Purification of Mutant Proteins**

For all mutants, small-scale expression was performed to determine the optimum expression parameters following the method described above for WT *Tpv* sHSP 14.3 (See section 2.3.2).

IPTG final concentration for induction of the cultures of the mutants ( G48E, E43V, Y34FG48E and G48EI108K variants) in T7 Express *E.coli* competent cells was 0.4 mM and the lysis buffer of 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, (pH=8.0) was used for resuspension of the pellets. Heat treatments were performed between 60°C and 80°C as an initial step for the protein purification. Cell lysates before and after heat treatment were analyzed by 12% SDS–PAGE.

For large scale expression and purification of *Tpv* sHSP 14.3 mutant variants by HPLC, the optimized protocol for WT was followed (See section 2.3.2).

### **2.3.6 Chaperone Activity Assay with Pig Heart Citrate Synthase**

The ability of *Tpv* sHSP 14.3 wild type and ACD mutants to protect model substrate Pig Heart Citrate Synthase, CS (Sigma), from thermal inactivation was evaluated following the CS activity assay as described previously (Kocabıyık & Aygar, 2012). In this assay, CS was diluted in assay buffer (20 mM Tris, 1mM EDTA pH 8.0) to a final concentration of 0.0088 mg/mL (0.176 μM) and incubated in the presence or absence of the potent chaperone at 47°C for 10 min. The assay was performed at 35°C in a thermostatically controlled UV-Visible spectrophotometer (Schimadzu UV-1601A, Kyoto, Japan). The remaining activity was monitored by continuous measuring of the absorbance of the assay mixture at 412 nm. The slope of the initial linear increase in absorption was used to calculate the rate of reaction. Protection effect of *Tpv* sHSP 14.3 WT and ACD mutants was investigated at three different molar ratios; 1/1700, 1/850 and 1/500 for CS: sHSP molar ratio corresponding to w/w ratios of 1:500, 1:250 and 1:147, respectively when 220 μg/mL, 110 μg/mL and 64 μg/mL of chaperone and 0.44 μg/mL of CS was used. Each experiment was repeated at least three times.

### **2.3.7 Thermal Aggregation Assay using CS as Substrate**

Effect of the *Tpv* sHSP 14.3 wild type and its mutant variants on the thermal aggregation of CS from porcine heart (Sigma) was studied by measuring the increase

in light scattering at 320 nm with a spectrofluorometer (Multiskan Go, ThermoScientific, Waltham, MA). The buffer of the CS from Sigma was exchanged by assay buffer (50 mM HEPES-KOH, pH 7.56) using 5000 MWCO filter device (Ultrafree Low Binding Cellulose, EMD Millipore, Burlington, Massachusetts, USA) and stored at -20°C until use. Concentration of each variant protein as well as the CS was measured using Picodrop (Picopet 01, Picodrop Ltd.UK). CS was diluted with assay buffer to a final concentration of 1.4  $\mu$ M (70  $\mu$ g/mL) and heated at 45°C with 9.8  $\mu$ M (or 140  $\mu$ g/mL) and 49  $\mu$ M (or 700  $\mu$ g/mL) which corresponds to 1:2 and 1:10 w/w ratio of CS: sHSP, respectively, or without purified *Tpv* sHSP 14.3. Aggregation assay was performed in a final volume of 150  $\mu$ L per well of a microplate (Microplate Corning 3631, 96 well, Corning Incorporated Life Sciences, Acton, MA). Light scattering at 320 nm was monitored with continuous shaking at 1 min intervals for up to 2 hours. Assay buffer was used as blank. CS without *Tpv* sHSP 14.3 served as control. Each experiment was performed in triplicate. Results were interpreted as the percent aggregation where the aggregation of CS alone (without sHSP) was taken as 100%.

### **2.3.8      Chaperone Activity Assay with Alcohol Dehydrogenase**

The chaperone activities of wild type and *Tpv* ACD mutants were estimated using yeast alcohol dehydrogenase (ADH) (EC1.1.1.1, Sigma) as described by Kagi & Vallee (1960). Yeast ADH is a tetramer of 37.5 kDa subunits. Its temperature optimum is 30°C and the activity decreases sharply at temperature above 40°C (Son *et al.*, 2007). ADH catalyzes the conversion of ethanol to acetaldehyde with the help of  $\beta$ -NAD. ADH activity was monitored by recording the increase in amount of  $\beta$ -NADH<sup>+</sup>, which has a unique absorbance at 340 nm. The ADH (0.053  $\mu$ M) was heated in water bath for 20 min at 47°C in the presence at a molar ratio of 1:220 ADH/chaperone or w/w ratio of 1:90 ADH/chaperone or absence of WT or mutant *Tpv* sHSP 14.3 (12-13  $\mu$ M). After heat treatment, ADH activity assay was carried out in 0.3 mM sodium phosphate buffer pH 8.8, containing 7.5 mM  $\beta$ -nicotinamide adenine dinucleotide and 3.2% (v/v) ethanol in a final volume of 1.0 mL. The rate of

reduction of NAD<sup>+</sup> was monitored spectroscopically at 340 nm using a Shimadzu UV-1601A spectrophotometer attached to a thermostatic cell holder at 25°C. Spectrophotometric readings were taken every 7.5 seconds.

### **2.3.9 Thermal Aggregation Assay with ADH**

To analyze the effect of sHSP on the heat aggregation of ADH, *Tpv* sHSP14.3 WT and mutants were added at different concentrations to a reaction mixture in 50 mM sodium phosphate buffer containing 150 mM NaCl, pH 7.2. Aggregation kinetics was followed by measuring the light scattering at 340 nm in the absence or presence of the recombinant *Tpv* sHSP14.3 proteins at 43°C using a 96 well plate reader (Multiskan Go, ThermoScientific). The concentration of the sHSP tested were 22 μM, 30 μM and 44 μM (per monomer), which corresponds to a molar ratio of 1:5, 1:7 and 1:10 (ADH /chaperone) or w/w ratio of 1:2, 1:3 and 1:4 (ADH/ chaperone) respectively.

### **2.3.10 Native Polyacrylamide Gel Electrophoresis**

In order to assess the oligomeric states of the wild type and mutant *Tpv* sHSP 14.3 proteins, Polyacrylamide Gel Electrophoresis (PAGE) were performed under native conditions. Native PAGE was performed by preparing 4%-20% gradient gel. In this system, proteins are separated based on their charge to mass ratios. Running conditions were 100V; 30 mA for starting and then 80V; 30mA for almost 4 hours at 4°C. Protein standards were Ribonuclease A (13.7 kDa) Ovalbumin (44 kDa), Aldolase (158 kDa), Catalase (232 kDa) and Ferritin (440 kDa) (Amersham Biosciences, Piscataway, NJ, and GE Healthcare, USA).

### **2.3.11 Blue Native PAGE Analysis**

The Blue Native Polyacrylamide Gel Electrophoresis (BN-PAGE) is used for both soluble and membrane proteins to determine the molecular sizes of the protein complexes. It has advantages compared to the native page that uses an anionic dye;

Coomassie blue G-250 changes the isoelectric point of proteins to more negative values. Therefore, even basic characteristic of proteins become more negative and migrates to the anode regardless of their initial charges (Schägger, 2004).

In this study, BN-PAGE was used to determine molecular mass of the oligomeric states of *Tpv* sHSP 14.3 variants. 4%-13% gradient acrylamide gel was prepared as described by Wittig *et al.*, (2006). In order to observe the effect of heat on oligomeric distribution, sHSP samples were heat treated at 70°C for 10 min. Electrophoresis was performed in imidazole/tricine buffers by addition of 0.02% Coomassie Blue G-250 in cathode buffer at 4°C.

Ferritin (880 and 440 kDa), Catalase (232 kDa), Aldolase (232 kDa) and Ribonuclease A (13.7 kDa) were used as standards. The gel was covered with ~100 mL of the destain solution (25% of methanol and 10% of acetic acid) and background was destained with gentle agitation. The gel was visualized using ChemiDoc MP device (Bio-rad, Hercules, California, USA) and molecular weight analysis was performed with Image Lab Software (ver.5.2.1).

### **2.3.12 Designing Alpha Crystallin Derived Novel Mini-Chaperone Peptides**

Bis-ANS and mellitin binding sites of human  $\alpha$ A-crystallin or  $\alpha$ B-crystallin (Bhattacharyya *et al.*, 2006; Sharma, 2001), besides conserved hydrophobic groups from multiple sequence alignment were taken into consideration for designing *Tpv* sHSP 14.3 ACD based novel mini-chaperone peptides. The designed peptides with 8-amino acid, 11-amino acid, 18- amino acid and 20-amino acid that are residing in the  $\beta$ 3- $\beta$ 4,  $\beta$ 6- $\beta$ 7, and  $\beta$ 9 regions of ACD of *Tpv* sHSP14.3 protein were synthesized and purified by Genscript (Genscript Biotech, Piscataway, New Jersey, USA). Purity of peptides was >95 according to HPLC. Peptides are named as YIE (18 aa), GIK (8 aa), LVD (11 aa) and SDS (20 aa) depending on the first and last two amino acid residues in the sequence.



Function of the mini chaperone peptides to prevent protein aggregation was investigated by using the model substrate proteins, yeast ADH (EC1.1.1.1, Sigma Chemical Co., St. Louis, USA) and pig heart Citrate Synthase (EC 4.1.3.7, Sigma). Mini-chaperone peptides were dissolved in a suitable solution considering the company's peptide solubility test report and then their concentrations were re-checked by Picodrop.

The ADH aggregation assay procedure with some modifications was as described previously (K K Sharma, 2001). Briefly, 200-250 µg/mL of ADH was heated in 150 µL of 50 mM sodium phosphate buffer [100 mM NaCl] pH 7.04 in the absence or presence of the mini peptides at 43°C. Aggregation kinetics was recorded with pulsed shaking by measuring the light scattering at 340 nm using a 96 well plate reader (Multiskan Go, ThermoScientific).

In the heat aggregation assay when CS was used as a substrate, its aliquot of 50-60 µg/mL in 50 mM HEPES-KOH, (pH 7.5), in the absence or presence of mini-peptides was incubated at 45°C. The time course of CS aggregation was monitored by measuring light scattering at 320 nm using a 96 well plate reader (Multiskan Go, ThermoScientific, Waltham, MA).

### **2.3.13 Bioinformatics Analyses and Three-Dimensional Structure Modelling of *Tpv* sHSP14.3**

The multisequence alignments were performed by using Clustal W Program in the EMBL-EBI data base (<https://www.ebi.ac.uk/Tools/msa/clustalo/>), and results were viewed with Jalview (Waterhouse *et al.*, 2009). Jpred created in Jalview 2.11.1.3 was used for secondary structure prediction of *Tpv* sHSP 14.3. Also, PRALINE multiple sequence alignment tool was used for secondary structure prediction. The putative dimer interfaces were predicted by Conserved Domain Database (CCD) search of the National Center for Biotechnology Information. Isoelectric point (pI), and molecular weight (MW) of *Tpv* sHSP14.3 were calculated by the aid of ExPASy tool ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/)).

3-D model structures of the *Tpv* HSP 14.3 and its variants were predicted by using homology modelling (MODELLER 9.15 ver). Crystal structures of *S.tokodaii* (PDB entry 3AAC and 3VQM), *X.axonopodis* (PDB entry 3GLA) , *D.radiodurans* (PDB entry 4FEI) were used as templates. The model with the highest GA341 score was selected for dimer generation with Chimera Program and energy minimization was also achieved with same tool. Visualization and structure analyses of the models including structure superimpositions, 3-D hydrophobic surface analysis, residue-residue distance maps were obtained using UCSF Chimera package. Comparative analysis of the intra/inter molecular bonds after amino acid substitutions were performed by the BIOVIA Discovery Studio Visualizer (DSV) (Ver 4.5).

A web server, SCRATCH (available at <http://scratch.proteomics.ics.uci.edu>), was used to further define structural features of the *Tpv* sHSP 14.3 WT and its ACD mutants. Contacts between C $\alpha$  atoms of the amino acids within the protein were found by performing contact map analysis using SVMcon predictor. This analysis generates a 2-D matrix consisting of residue-residue interactions within a distance at 8 Å between C $\alpha$  atoms of amino acids.

MUpro web server was used for prediction of thermodynamic stability of *Tpv* sHSP 14.3 upon mutations. This program uses a new machine learning approach to predict stability changes for single amino acid substitutions by using the structure dependent and sequence dependent information. Accuracy of the program was found more than 84% (Jianlin Cheng *et al.*, 2006). Dynamut2 web server was also used for predicting the effects of single and double point mutations on protein stability and dynamics (a web server is available at <http://biosig.unimelb.edu.au/dynamut2>).

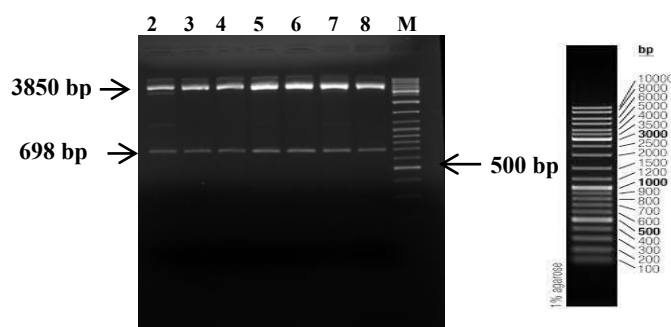
Zipper DB database (<https://services.mbi.ucla.edu/zipperdb/intro>) was used for the predictions of fibril forming segments within the sHSP proteins. Fibrillation propensity profile depends on the energy levels of hexapeptides whose energy threshold equal or below to -23 kcal /mol are associated with the high fibrillation propensity.

## CHAPTER 3

### RESULTS

#### 3.1 Sub-cloning of the *Tpv* sHSP14.3 Gene for Expression of the Recombinant Protein from its Own AUG

Expression of the recombinant *Tpv* sHSP 14.3 protein from its own AUG was achieved using the pET21a expression vector system and BL21 (DE3) competent *E.coli* cells. Previously, the gene encoding *Tpv* sHSP 14.3 (TVN0775) was cloned in our laboratory into pDrive Cloning Vector (Qiagen) together with its upstream and downstream sequences. This recombinant plasmid (pDrive-sHSP14.3) was isolated using the QIA prep Spin Miniprep Kit (Qiagen) and as a control cut with *Eco*RI restriction enzyme (ThermoScientific) to excise the cloned gene. The resulting fragments were analyzed by 1.5% agarose gel electrophoresis and the presence of the cloned gene was confirmed by size determination (Figure 3.1). pDrive-sHSP14.3 plasmid DNA was used as the source of *Tpv* sHSP14.3 gene in the sub-cloning experiments.



**Figure 3.1. Isolation of the target plasmid DNA.** 2 to 8 samples represent target DNA in pDrive Cloning Vector. Two correct DNA fragments appear on the gel. One belongs to the vector (3850 bp) and the other band 698 bp of *Tpv* sHSP14.3 with upstream and downstream sequences. M: O'Gene ruler DNA Ladder mix.

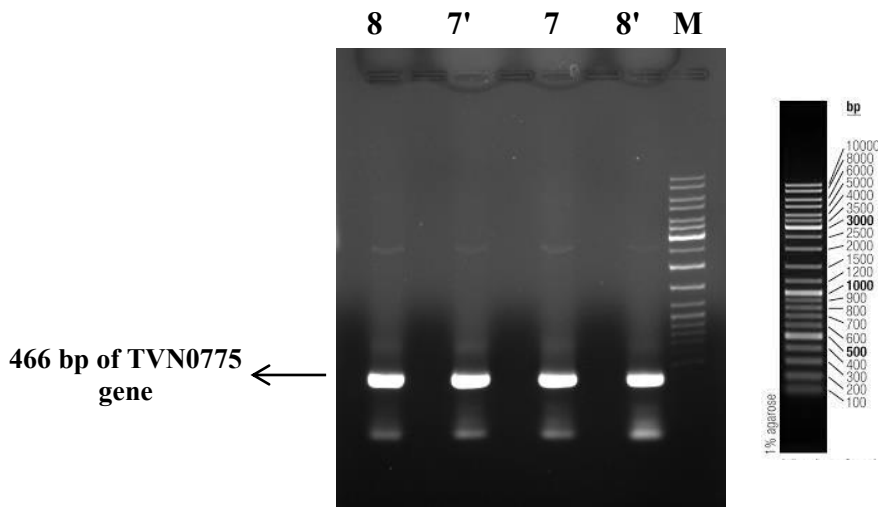
TVN0775 gene was amplified by PCR using the recombinant pDrive-sHSP14.3 vector as template DNA and forward and reverse amplification primers having *Nde*I

and *Bam*HI cut sites, respectively (Figure 3.2). Four different gene amplification results are shown in the Figure 3.3

```

CAAGCGGCAATAAGAAGGAGATAAGGTACAAGGAAGAGATTCGGCTTGCGATTT
CTAGAGAGATTGAAAAAAGGTTAACGGTGCTTTCGGTGAATTAATATCAGCTGA
GGACATTGCTAGGGATATGAGGATGAAGATCGATCCGAGTACAGTGGCTAGAAG
GCTATTTCACTCATCAAATGTAGCAGAAAAAGAAGTTAATTTTTAAATATATGCA
CGTATTTATTAAGT CATATG TATACACCCATAAAAGTTCTTTACG AATGAGATGAT
                  NdeI Forward Primer
AAAAAACGTATCGAATACTGTGAAAGAGGTCTCATCCTTTATATATCCACCAGTC
ACGTTATATCAAGATAGCTCTGATCTGGTATTGGAAGCAGAAATGGCCGGGTTTG
ACAAGAAAAACATAAAGGTCTCGGTAAATAAGAATGTACTCACTATAAAGTGCGG
AGAGAAAGAGAGAATACTCTACCGTATATATCGATCAGCGCGTTGACAAAGTGT
ATAAAGTAGTTAAGCTGCCCGTAGAGATTGAGCAGCAGGACATATCTGCTAAGTA
TAGTGAAGGCATACTTACAGTTAGAATGAAAACCAAGAACATAAAGAACGTAGA
AATAGAATAAAATCATTTTTTAATAATAATATATATGAAAAGTATTGCATCTATT
GCTATAA GTATGCTTGATGTGATTGGGTG GGATCC
          Reverse Primer BamHI
  
```

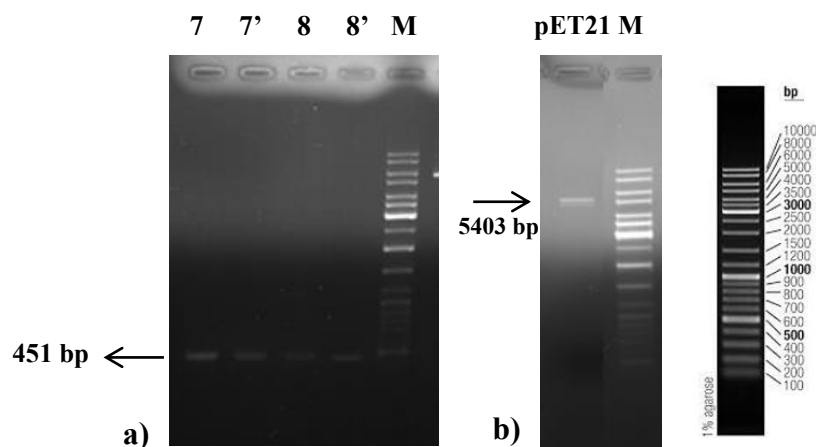
**Figure 3.2.** The sequence of the recombinant pDrive-sHSP14.3 as template DNA. Primers including restriction enzyme recognition sites are marked by green and purple color. TVN0755 gene is marked by yellow color.



**Figure 3.3.** Amplification of the TVN0775 gene by PCR. 7'-8' and 7-8 represent 2  $\mu$ L and 3  $\mu$ L of source DNA, respectively. M: O'Gene ruler DNA Ladder mix.

Then, PCR products cleaned by gel purification and cut with restriction enzymes *Nde*I and *Bam*HI and together with the pET21a expression vector. Purification of the digested vector and insert were carried out by using QIAquick Gel Extraction Kit

(Qiagen). After digesting, DNA samples were gel purified and loaded on an agarose gel to visualize the recovered fragments (Figure 3.4).



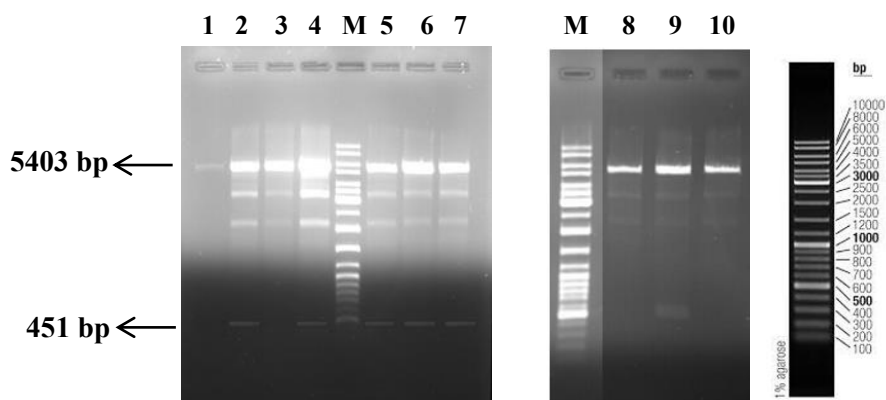
**Figure 3.4. Agarose gel images of double digested insert and vector after purification.** a) Double digested PCR fragments after purification was illustrated by arrow. 7 and 8' PCR products are cut with restriction enzymes *NdeI* and *BamHI* from BioLabs company. 7' and 8 are cut with the restriction enzymes from ThermoScientific Company. b) Double digested pET21 vector after purification is seen in the picture. M: O'Gene ruler DNA Ladder mix.

After ligation of the cloned gene fragment into pET21a vector, the ligation products were transferred into BL21(DE3) competent cells by transformation. Before transformation, the efficiency of the competence of the cells were checked by using control plasmid pUC19 (NEB). Although expected transformation efficiency was  $1-5 \times 10^7$  cfu/ $\mu$ g, we found that it varied between  $1 \times 10^6$  and  $2 \times 10^6$  cfu/ $\mu$ g. The transformants were grown on LB agar containing ampicillin. Transformation efficiency with different ligation mixes was found ranging between  $2 \times 10^2$  and  $2 \times 10^4$  cfu/ $\mu$ g.

### 3.1.1 Screening of Transformant Colonies for Recombinant Clones by Plasmid Isolations and Restriction Enzyme Digestions

The plasmid DNA is extracted from randomly picked ten transformant colonies and checked to see if the ligation was successful by cutting with *NdeI* and *BamHI* restriction enzymes. Out of 10 isolated plasmids, 6 of them were seemed to be

recombinants, because double digestion yielded two bands: one is the vector DNA (5403 bp) and the second band is 451 bp cloned *Tpv* sHSP 14.3 gene. The analysis of the gel images showed that plasmid samples of 2, 4, 5, 6, 7, and 9 were putative recombinant plasmids (Figure 3.5).

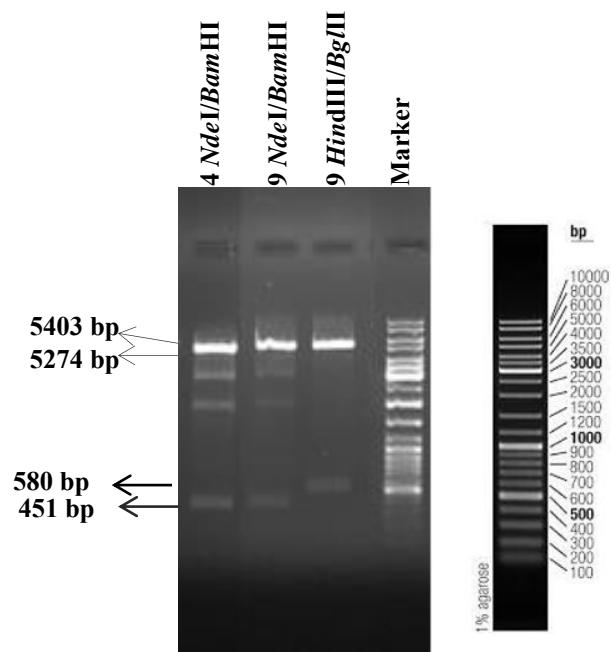


**Figure 3.5. Screening of the colonies by double digestion of plasmids samples of 1 to 10.** Plasmids are cut with restriction enzymes *NdeI* and *BamHI*. Two correct DNA fragments appear on the gel. One belongs to the vector backbone (5403 bp) and the other band 451 bp of the cloned *Tpv* sHSP14.3 gene. M: O'Gene ruler DNA Ladder mix.

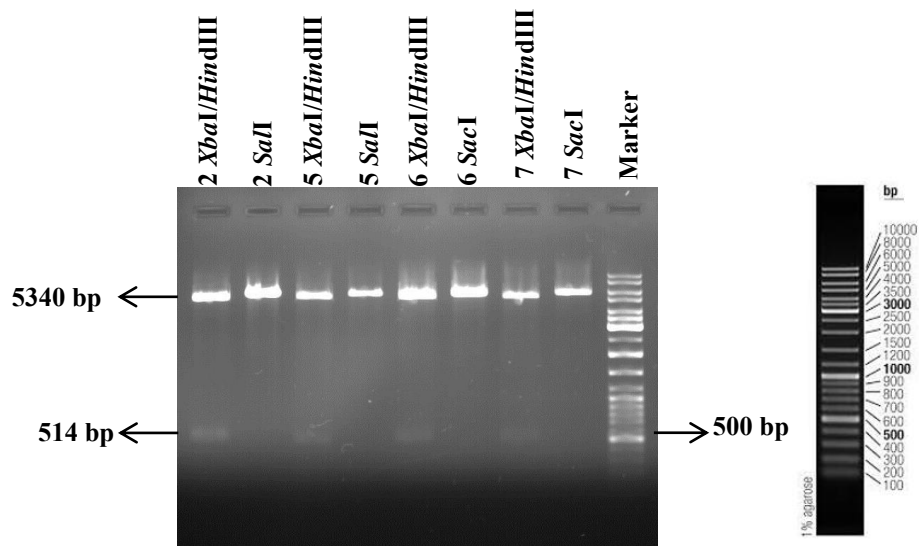
The putative recombinant colonies were selected for further characterization by digestion with different combinations of restriction enzymes. Restriction enzymes were chosen referring to the restriction map of the recombinant plasmid. *HindIII*+*BglII* digestion of the plasmid sample 9 yielded correct size DNA fragments at a position of 5274 bp and 580 bp. The plasmid samples 4 and 9, when cut with *NdeI*+*BamHI* restriction enzymes, a 451 bp of insert DNA was excised. This digestion confirmed that plasmids of samples 4 and 9 are most probably true recombinant plasmids (Figure 3.6).

For verification of the recombinant plasmid samples of 2, 5, 6 and 7, they were double digested with *XbaI* and *HindIII* restriction enzymes. The correct size DNA fragments at a position of 5340 bp and 514 bp were seen on 1% agarose gel. Restriction enzymes, *SalI* and *SacI*, are noncutter for TVN0755 gene, but they cut pET21 vector at a single position. Digestion of these four plasmids DNA with *SalI* or *SacI* restriction enzymes linearized them and yielded fragments at expected

positions (5854 bp) (Figure 3.7). Therefore, these plasmids were proved to be true recombinant plasmids, as well.



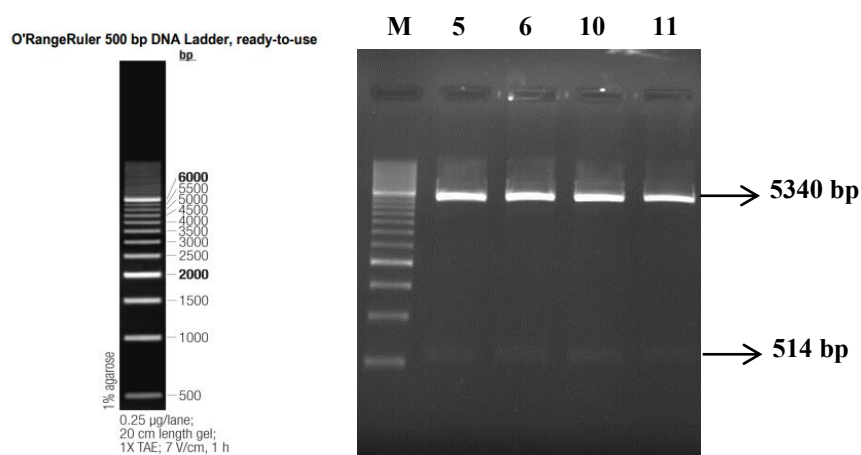
**Figure 3.6. Screening of the colonies of the plasmid samples 4 and 9 by restriction enzyme digestions.** Plasmid samples 4 and 9 are double digested with *NdeI*+*Bam*HI restriction enzymes, yielding a 451 bp of insert and 5403 bp of vector backbone. *Hind*III+*Bg*III digestion of the plasmid sample 9 yields correct size DNA fragments at a position of 5274 bp (vector backbone) and 580 bp (insert). Marker: O'Gene ruler DNA Ladder mix.



**Figure 3.7. Screening of the colonies of the plasmid samples 2, 5, 6 and 7 by restriction enzyme digestions.** Double digestion of the plasmid samples with *XbaI*/*Hind*III yield fragments at a position of 5340 bp (vector backbone) and 514 bp (insert). *Sal*I or *Sac*I linearized the recombinant plasmid and single band corresponds to the 5854 bp. Marker: O'Gene ruler DNA Ladder mix.

### 3.1.2 Verification of Restriction Mapping Results by DNA Sequencing

To verify the recombinant clones by sequencing, among the positive clones named *E.coli* pET21\_tvshsp9 (with plasmid sample 9) was selected and its plasmid DNA was prepared to be sent for sequencing. Also, pET21\_tvshsp9 recombinant plasmid was transferred into NovaBlue Competent cells which contains no source of T7 RNA polymerase, thus it is ideal for the long-term storage of the recombinant plasmids under nonexpression conditions. After transformation, plasmids were checked by double digestion with *Xba*I +*Hind*III restriction enzymes in order to confirm presence of the gene (Figure 3.8).



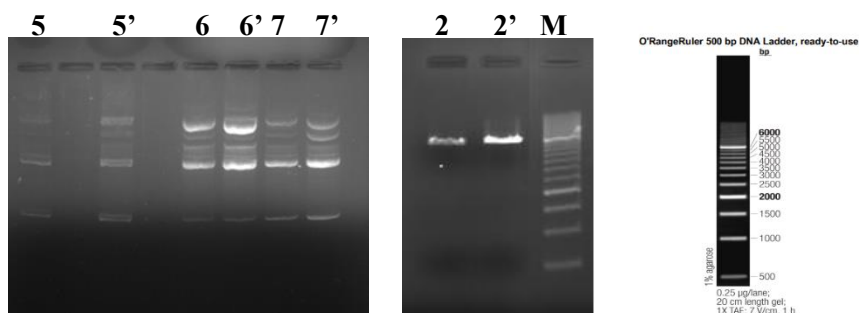
**Figure 3.8. Agarose gel image *E.coli* pET21\_tvshsp9 samples.** Isolated plasmid DNAs from *E.coli* pET21\_tvshsp9 colony 5, 6, 10 and 11 are digested with *Xba*I+*Hind*III restriction enzymes. The correct size DNA fragments at a position of 5340 bp and 514 bp are seen on 1% agarose gel. Marker: O' RangeRuler 500 bp DNA Ladder.

However, sequence analysis showed that there was a missense mutation by replacement of Lysine by Glutamic acid at position 87. Multiple sequence alignment and also chromatogram images were shown in Appendix D.

Therefore, plasmids of other positive clones (pET21\_tvshsp2, pET21\_tvshsp5, pET21\_tvshsp6, pET21\_tvshsp7) were isolated and their purity and concentration were checked by Picodrop and agarose gel electrophoresis as a preparation for sequencing. Two replicates of undigested plasmid preparations of samples 5, 6, and 7 were run on 1% agarose gel (Figure 3.9). The plasmid sample 2 was single digested

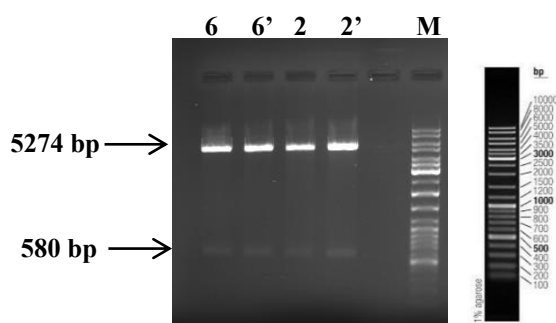


with *SalI* restriction enzyme. Based on gel images as well as pico drop results, samples with a high concentration of DNA were sent for sequence analysis. Sequencing results well confirmed that all four clones are true recombinants with no changes in the original sequence of TVN0775 (Appendix E).



**Figure 3.9. DNA quality determination for sequencing by agarose gel electrophoresis.** Plasmid DNA of pET21-tvsHSP2 is single cut with *SalI* restriction enzyme. Expected band size appears as 5854 bp. Other colonies (5, 6 and 7) are run without restriction enzyme digestion. Each plasmid samples are isolated twice and replicates are marked by prime symbol ('). Colonies 2', 5', 6', and 7' are selected for sequence analysis. Marker: O'RangeRuler 500 bp DNA Ladder, ready-to-use.

After verification of the constructs by DNA sequencing, colony number 2 and 6 were introduced into NovaBlue Competent Cells (Novagen) by transformation in order to stable maintenance of the plasmid under nonexpressing conditions. After transformation, plasmids were checked by double digestion with *BglII* + *HindIII* (ThermoScientific) restriction enzymes in order to confirm the presence of the gene. The expected DNA fragments at positions of 5274 bp and 580 bp were observed on 1% agarose gel (Figure 3.10).



**Figure 3.10. Agarose gel image of recombinants in NovaBlue Competent cells.** Plasmids are digested with *BglII*+*HindIII* restriction enzymes. Two correct DNA fragments appear on the gel. One belongs to the vector backbone (5274 bp) and the other band 580 bp. 6 and 6' are the plasmids of pET21-tvsHSP6; 2 and 2' are the plasmids of pET21-tvsHSP2. Each plasmid samples are isolated twice and replicates are marked by prime symbol ('). M: O'Gene Ruler DNA Ladder Mix (ThermoScientific).

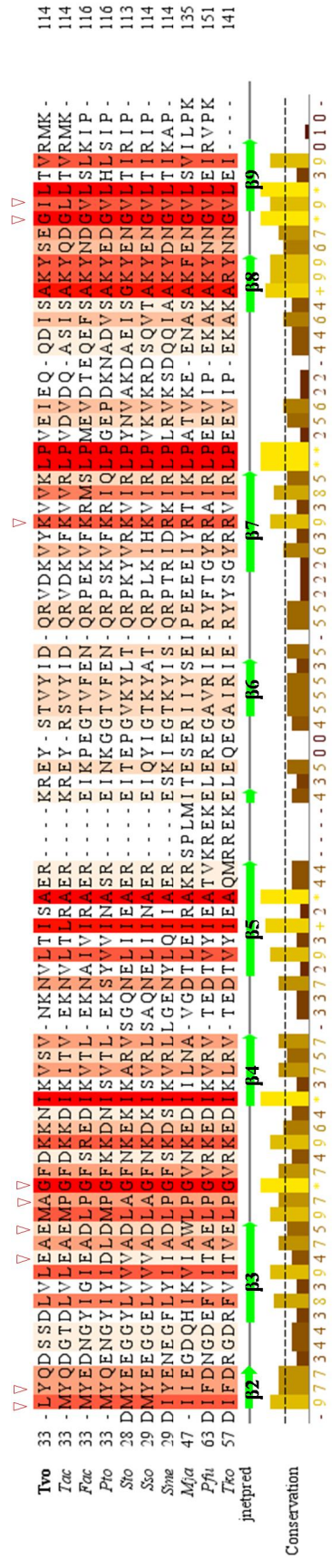
### 3.2 Selection of the Amino Acids to be Targeted in Mutagenesis

In this study, the ACD of the *Tpv* sHSP14.3 is the region of focus, which is a characteristic structural signature of the sHSPs. The residues in the core domain targeted for mutagenesis were L33, Y34, E43, E45, A47, G48, G107 and I108. They were located at the putative dimer interface as predicted by the National Center for Biotechnology Information's Conserved Domain Database (CCD) search. Double mutations were also generated by combining two single mutations (Y34FG48E, G48EI108K) in order to evaluate additive effects of those mutations as well as on the structure and function of the *Tpv* sHSP 14.3.

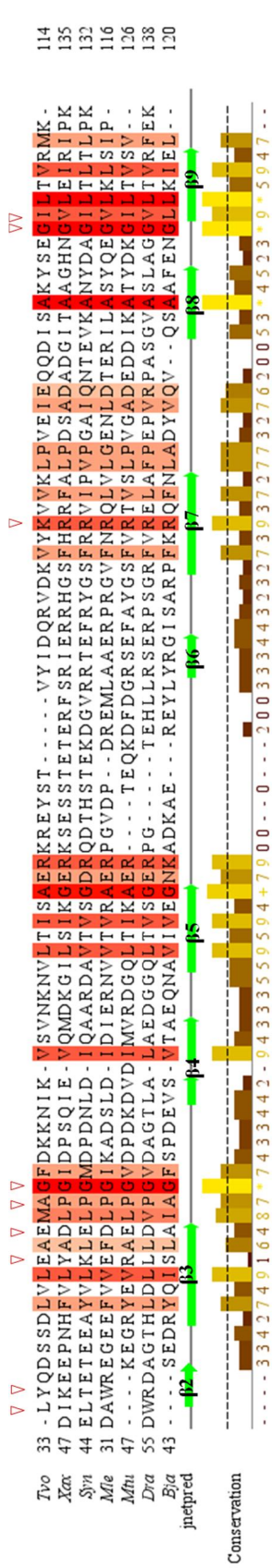
The selected residues for mutagenesis were found to be highly conserved among the sHSPs from archaea, eubacteria and eukarya as identified by multiple sequence alignments (Figure 3.11-3.13). The multiple sequence alignments (MSAs) were performed by Clustal W program and results were viewed with Jalview tool (Waterhouse *et al.*, 2009). A summary of the equivalent residues at the targeted positions as revealed by the multiple sequence alignment analysis is given in the Table 3.1.

**Table 3.1. List of chosen residues and their equivalents as determined by MSA in the ACDs of the sHSP from top ten blast hit species archaea, eubacteria, and eukarya sHSPs.**

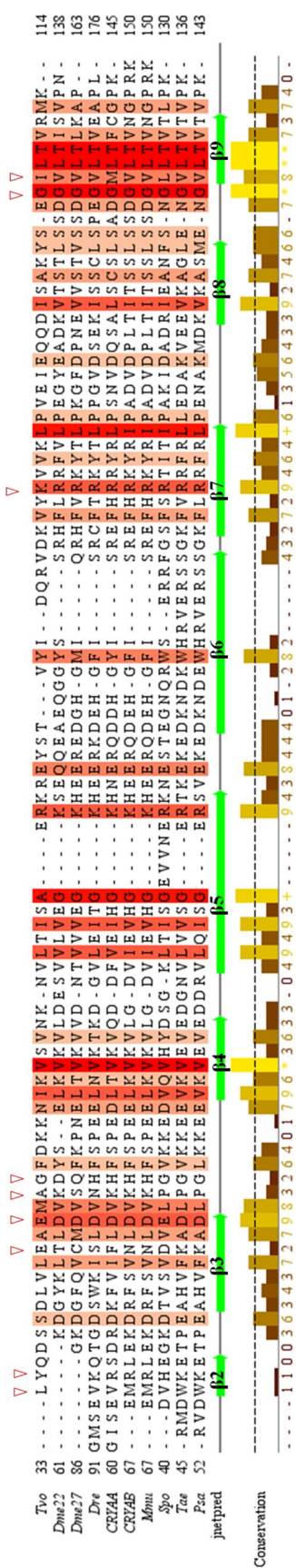
| Positions in <i>Tpv</i> sHSP14.3 | Conservation among archaea            | Conservation among eubacteria | Conservation among eukaryotes |
|----------------------------------|---------------------------------------|-------------------------------|-------------------------------|
| <b>L33</b>                       | Methionine, Valine, Isoleucine        | Valine, or isoleucine         | Leucine, or isoleucine        |
| <b>Y34</b>                       | Tyrosine                              | Hydrophobic aa dominant       | Histidine, Tyrosine           |
| <b>E43</b>                       | Glutamic acid, Valine, and Isoleucine | Glutamic acid, Arginine       | Glutamic acid and Threonine   |
| <b>E45</b>                       | Glutamic Acid or Aspartic Acid        | Glutamic Acid                 | Glutamic Acid                 |
| <b>A47</b>                       | Alanine or Proline                    | Proline                       | Proline                       |
| <b>G48</b>                       | Glycine                               | Glycine                       | Glycine                       |
| <b>K87</b>                       | Lysine or Arginine                    | Lysine or Arginine            | Lysine or Arginine            |
| <b>G107</b>                      | Glycine                               | Glycine                       | Glycine                       |
| <b>I108</b>                      | Valine                                | Isoleucine or Valine          | Valine                        |



**Figure 3.11. Multiple sequence alignment of Alpha Crystallin Domain of Tpy sHSP 14.3 with the sHSPs from archaea.** The species names corresponding to each accession number are as given: **CBY78065.1:Tvo:** *Thermoplasma volcanium* GSSI, **CAC11993.1:Tac:** *Thermoplasma acidophilum*, **ARD84221.1:Fac:** *Ferroplasma acidophilum*, **AAT43324:Pto:** *Picrophilus torridus*, **WP\_054838418.1:Sme:** *Sulfuracidifex metallicus*, **WP\_010979712.1:Sto:** *Sulfolobus tokodaii*, **WP\_009989320.1:Sso:** *Saccharolobus solfataricus*, **WP\_010869783.1:Mja:** *Methanocaldococcus jannaschii*, **AAF71367.1:Pfu:** *Pyrococcus furiosus*, **WP\_048053707.1:Tko:** *Thermococcus kodakarensis*. Mutation points are marked by triangle. Conservation of the amino acids is shown by numerical index from 0 to 11. Scores are shown under the histogram. Complete conservation is associated with a "+". Highly conserved residues are shaded by dark red.



**Figure 3.12. Multiple sequence alignment of Alpha Crystallin Domain of Tpv sHSP 14.3 with the sHSPs from eubacteria.** The species names corresponding to each accession number are as given: **CBY78065.1: Tvo:** *Thermoplasma volcanium* GSS1, **KPL49747.1: Xax:** *Xanthomonas axonopodis*, **WP\_010871624.1: Syn:** unclassified *Synechocystis*, **WP\_010908539.1: Mie:** *Mycobacterium leprae*, **WP\_031682063.1: Mtu:** *Mycobacterium tuberculosis*, **NP\_295414.1: Dra:** *Deinococcus radiodurans* R1, **CAA09014.1: Bja:** *Bradyrhizobium japonicum*. Mutation points are marked by triangle. Conservation of the amino acids is shown by numerical index from 0 to 11. Scores are shown under the histogram. Complete conservation is associated with a "+". Highly conserved residues are shaded by dark red.



**Figure 3.13. Multiple sequence alignment of Alpha Crystallin Domain of Tpv sHSP 14.3 with the sHSPs from eukaryotic organisms.** The species names corresponding to each accession number are as given: **CBY78065.1:** Tvo: *Thermoplasma volcanium* GSS1, sp|P02515.4| Dme22: *Drosophila melanogaster*, AA28638.1; Dme27: *Drosophila melanogaster*, AAU25839.1; Dre: *Danio rerio*, P02489.2; CRYAA\_HUMAN, P02511.2; CRYAB\_HUMAN, NP\_001276711.1; Mmu: *Mus musculus*, NP\_596091.1; Spo: *Schizosaccharomyces pombe*, Q41560.1; Tae: *Triticum aestivum*, P19243.1; Psa: *Pisum sativum*. Mutation points are marked by triangle. Conservation of the amino acids is shown by numerical index from 0 to 11. Scores are shown under the histogram. Complete conservation is associated with a "+". Highly conserved residues are shaded by dark red.

Comparison of the ACD sequences of *Tpv* sHSP14.3 with that of archaea sHSPs showed that there is a high homology between *Thermoplama acidophilum* and *Thermoplasma volcanium* with a percent identity of 73.17 %. The other high close homologs of *Tpv* sHSP14.3 are HSP20 family proteins of *Ferroplasma acidophilum*, *Picrophilus oshimae*, *Picrophilus torridous* and *Sulfolobulus tokadaii* with the homology scores of 39.02 %, 36.59 %, 36.59 %, and 40.24 %, respectively. On the other hand, percent identities were low for hyperthermophilic archaea species (*i.e.*, *Methanococcus jannaschii* 24.69 % and *Pyrococcus furious* 29.63%).

Homology scores between the ACD of *Tpv* sHSP14.3 and that of eubacteria species were found low as compared to the archaeal species. The close homologs of *Tpv* sHSP14.3 are small heat shock proteins of *Xanthomonas axonopodis* (30.49 %), *Mycobacterium tuberculosis* (32.43 %) and *Synechocytisis* (26.83 %). Small heat shock proteins from *Mycobacterium leprae*, *Deinococcus radiodurans* and *Bradyrhizobium japonicum* have homology scores 19.75 %, 20.51 % and 14.67 % respectively. Furthermore, human small heat shock protein CRYAA and CRYAB are not close homolog of *Tpv* sHSP 14.3 with homology scores 18.18 % and 15.58 %. On the other hand, small heat shock proteins of *Triticum aestivum* and *Pisum sativum* have slightly high homology with *Tpv* sHSP14.3 as compared to human sHSPs (*i.e.*, 23.08 % and 26.92%, respectively). Among these organisms, the most distant homolog of *Tpv* sHSP14.3 was HspH from bacteria *Bradyrhizobium japonicum* (14.67 %) indicating that sHSPs are variable in sequences.

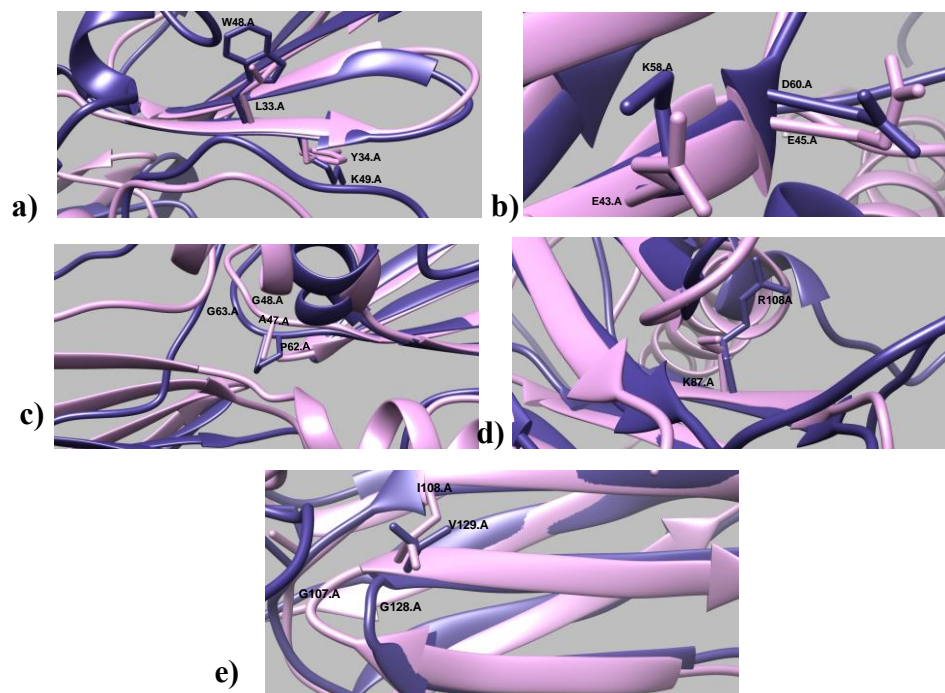
MSA of ACDs displays highly conserved residues and motifs in sHSPs of different organisms (Figure 3.11-3.13). These residues are mostly found along the  $\beta$ 7- $\beta$ 9 zone. Within this region, there is disease associated R residue in  $\beta$ 7 strand. In *Tpv* sHSP14.3, the equivalent residue is K87 in the *Tpv* sHSP14.3 (Figure 3.13). In the L89 loop and beginning of the  $\beta$ 9 strand, the conserved motif is N/D-G-hydrophobic-L. Two residues G107 and I108, which are targeted for mutation in our study are within this conserved sequence motif. In the  $\beta$ 2- $\beta$ 5 zone mainly in  $\beta$ 4 and  $\beta$ 5 strands, there are hydrophobic amino acids (*e.g.*, L and G/A in the  $\beta$ 5). Two mutation points, L33 and Y34 are on the  $\beta$ 2, while another target position E43 is located in the  $\beta$ 3

strand. Another substituted residue G48 belongs to P-G doublet, which is highly conserved in non-animal sHSP and found in dimer interface of HSP16.5 of *Mycobacterium tuberculosis* (Fu & Chang, 2006) (Figure 3.12). In some archaeal sHSP including *Tpv* sHSP14.3, P-G doublet is replaced by A-G doublet (Figure 3.11). The A47 residue of this doublet also targeted for point mutation.

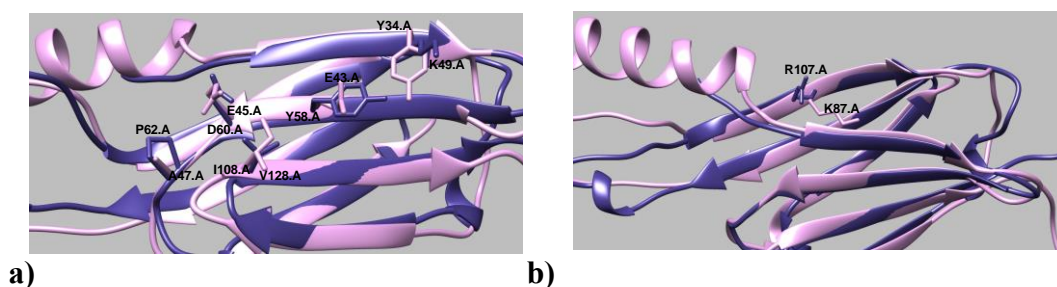
### **3.2.1 Structure Superimpositions of *Tpv* sHSP 14.3 with Well Known Small Heat Shock Proteins**

The 3-D structure analyses of the sHSPs from other organisms also predicted critical roles for the residues we focus on in this research. Structural analysis of the crystal structure of the plant sHSP (*i.e.*, sHSP16.9 of wheat), showed that Trp 48 and Phe 110 on the alpha crystallin domain might be important for substrate binding (R. L. M. Van Montfort *et al.*, 2001). The residue Trp 48 was equivalent to L33 in *Tpv* sHSP14.3 as we found by MSA and 3D structure superimposition (Figure 3.14). Other, predicted dimer interface points of the sHSP16.9, K49, K58, D60, P62, G63, R108, G128 and V129 were found topologically equivalent to the selected residues, Y34, E43, E45, A47, G48, K87, G107, and I108, for the mutagenesis study of *Tpv* sHSP14.3.

Another example of the crystal structure study is the *Xanthomonas* small heat shock protein XaHspA. The results gave insight about the mechanism of the complex formation with nonnative substrate proteins and structural organization of these complexes (Hilario *et al.*, 2011). In addition, key residues for dimerization were suggested as R45, V46 (main chain), D47, K49, E50, E51, Y58, D60, P62, I64 (main chain), D65, R84, E91, R94, F95, R97, E99, R100, R101, R107, N126, and V128. The residues Y34, E43, E45, A47, K87 and I108 that we targeted for mutation in the *Tpv* sHSP14.3, were found to have equivalents of K49, Y58, D60, P62, R107, and V128 residues, respectively in the subunit contacts of the *Xanthomonas axonopodis* sHSP XaHspA when their 3-D structures were superimposed (Figure 3.15).



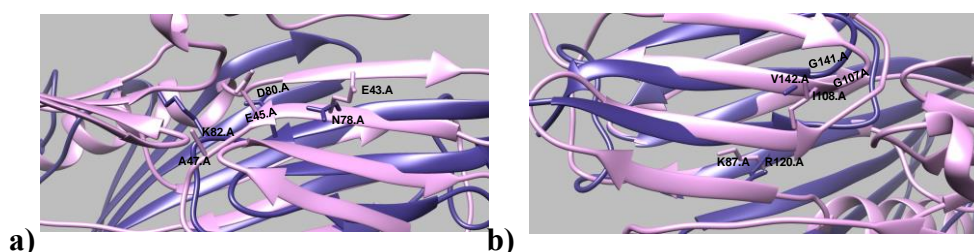
**Figure 3.14. Superimposition of the wheat sHSP, HSP16.9 (blue, PDB:1gme) and *Thermoplasma volcanium* sHSP14.3 (pink).** The positions of the residues in the putative dimer interface of *Tpv* sHSP14.3 that were substituted by site directed mutagenesis shows topological equivalence to the subunit contacts of the wheat sHSP16.9. **a)** The residues of L33 and Y34 are found to equivalent to the W48 and K49 of the wheat sHSP16.9, respectively. **b)** The residues of E43 and E45 are found to equivalent to the K58 and D60 of the wheat sHSP16.9, respectively. **c)** The residues of A47 and G48 are found to equivalent to the P62 and G63 of the wheat sHSP16.9, respectively. **d)** The residue of K87 is found to equivalent to the R108 of the wheat sHSP16.9, respectively. **e)** The residues of G107 and I108 are found to equivalent to the G128 and V129 of the wheat sHSP16.9, respectively.



**Figure 3.15. Superimposition of the dimeric *Xanthomonas axonopodis* sHSP XaHspA, (blue, PDB: 3GUF) and *Thermoplasma volcanium* sHSP14.3 (pink).** The positions of the residues in the putative dimer interface of *Tpv* sHSP14.3 that were substituted by site directed mutagenesis show topological equivalence to the subunit contacts of the *Xanthomonas* sHSP XaHspA. **a)** The residues of Y34, E43, E45, A47, I108 are found to equivalent to the K49, Y58, D60, P62, V128 of the *Xanthomonas* sHSP XaHspA, respectively. **b)** The residue of K87 is found to equivalent to the R107 of the *Xanthomonas* sHSP XaHspA.



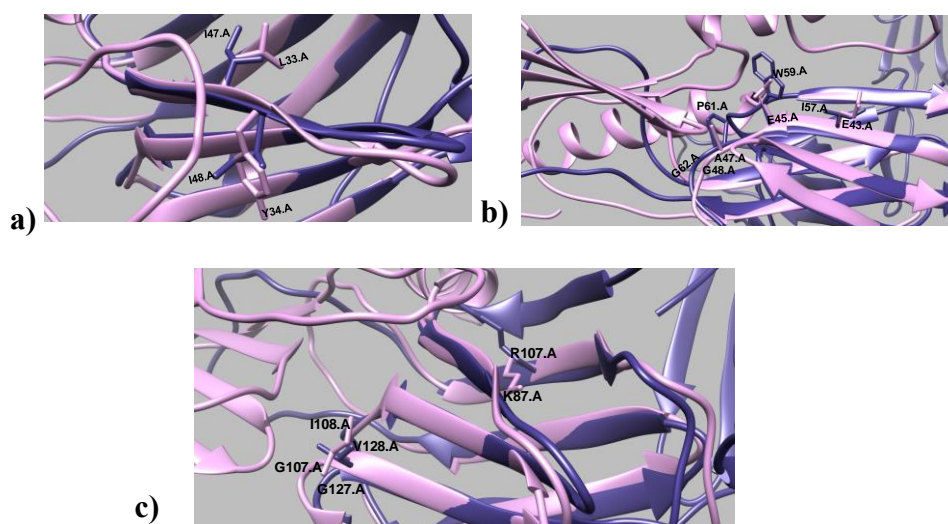
In another research, the subunit-subunit interaction sites were identified by protein pin array technology for the human small heat shock protein alphaB crystallin (CRYAB) (Ghosh & Clark, 2005). According to these results, N78D mutation in interactive sequence 76-FSVNLDVK-83 resulted in the formation of larger oligomeric assemblies. The equivalent residue for *T.volcanium* is E43 as found by MSA and by superimposition of its model 3D structure with that of  $\alpha\beta$  crystallin (Figure 3.16). The pin array analysis also identified exposed residues containing side chains that are capable of forming intermolecular interactions in  $\alpha\beta$  crystallin. The key residue of human CRYAB predicted for dimerization were N78, D80, K82, G141 and V142, having topological equivalents of E43, E45, A47, G107 and I108 in the *Tpv* sHSP14.3, respectively, as we found by structure superimposition (Figure 3.16).



**Figure 3.16. Superimposition of the dimeric human CRYAB (blue, PDB:2klr) and *Thermoplasma volcanium* sHSP14.3 (Pink).** The positions of the residues in the putative dimer interface of *Tpv* sHSP14.3 that were substituted by site directed mutagenesis show topological equivalence to the subunit contacts of the human CRYAB. **a)** The residues of E43, E45, and A47 are found to equivalent to the N78, D80 and K82 of the human CRYAB, respectively. **b)** The residues of K87, G107, and I108 are found to equivalent to the R120, G141 and V142 of the human CRYAB, respectively.

The crystal structure analysis of the HSP16.5 from the archaean *Methanococcus jannaschii* showed that residues I47, I48, I57, W59, P61, G62, R107, G127 and V128 could be involved in the intermolecular interactions (Kim *et al.*, 1998). 3-D structure comparison of the *Tpv* sHSP14.3 with the HSP16.5 by superimposition was performed to identify the equivalent residues of the subunit-subunit contacts. The topological equivalent residues showed conservation of the physico-chemical properties of the amino acids such that highly hydrophobic L33, A47, G48, G107, and I108 were corresponding to the hydrophobic residues I47, P61, G62, G127 and V128 of the MjHSP16.5, respectively (Figure 3.17). Also, well studied residue of

R107 of MjHSP16.5, mutations of its equivalent residues are associated with certain genetic disorders in human (R120G in  $\alpha$ B-crystallin, R116C in  $\alpha$ A-crystallin (HSPB4) and R127W in Hsp 27 in desmin-related myopathy, central nuclear cataracts and Charcot-Marie-Tooth disease) (Clark *et al.*, 2011), is corresponding to the K87 of *Tpv* sHSP14.3 in the 3D- superimpositions in Figure 3.17.

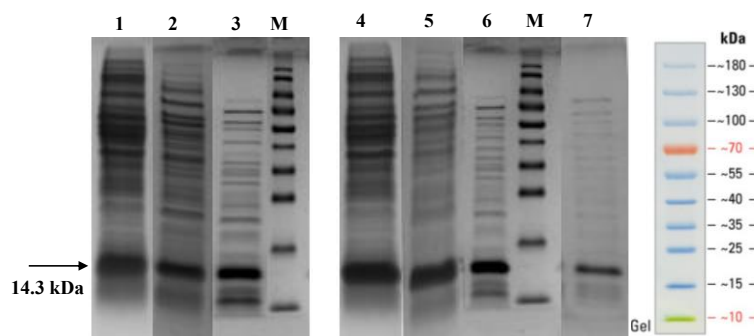


**Figure 3.17. Superimposition of the *Methanococcus jannaschii* HSP16.5 (blue, PDB:1shs) and *Thermoplasma volcanium* sHSP14.3 (Pink).** The positions of the residues in the putative dimer interface of *Tpv* sHSP14.3 that were substituted by site directed mutagenesis show topological equivalence to the subunit contacts of the *Methanococcus jannaschii* HSP16.5. **a)** The residues of L33, and Y34 are found to equivalent to the I47 and I48 of the Mja HSP16.5, respectively. **b)** The residues of E43, E45, A47 and G48 are found to equivalent to the I57, W59, P61, and G62 of the Mja HSP16.5, respectively. **c)** The residues of K87, G107 and I108 are found to equivalent to the R107, G127 and V128 of the Mja HSP16.5, respectively.

### 3.3 Overexpression of the Wild Type *Tpv* sHSP 14.3 and Heat Treatment

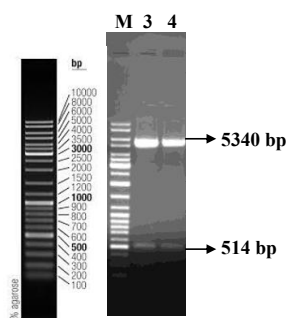
*E.coli* pET21\_tvshSP2 and *E.coli* pET21\_tvshSP6 clones were selected for expression using pET System. Small scale cell lysates were analyzed by SDS-PAGE. Results showed that recombinant *Tpv* sHSP 14.3 from both clones were successfully expressed. The *Tpv* sHSP14.3 protein bands were observed at a position of 14.3 kDa as expected (Figure 3.18). In order to eliminate the unstable host cell proteins, cell lysates were heat treated at 60°C for 15 min followed by centrifugation (Figure 3.18, Lane 2 and Lane 5). This treatment was not enough to denature most

of the *E.coli* proteins. When the temperature was increased to the 70°C after 30 min, still significant amount of the proteins of cell extract remained stable (Lane 3 and Lane 6). Heat treatment at 80°C for 30 min was sufficient to denature majority of the heat labile *E.coli* proteins (Figure 3.18, Lane 7).



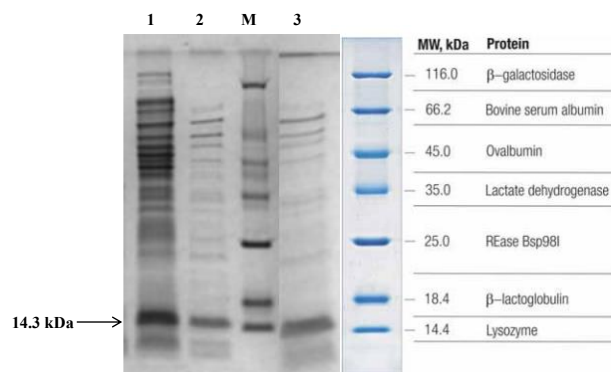
**Figure 3.18. SDS PAGE gel for the expression and heat stability of recombinant *E.coli* pET21\_tvshSP2 and *E.coli* pET21\_tvshSP6 clones.** Lane 1: Cell free extract of the *Tpv* sHSP14.3 colony 6 before heat treatment (BH), Lane 2-3: Heat treatment at 60°C, 70°C, respectively. Lane 4: Cell free extract of the *Tpv* sHSP14.3 colony 2 before heat treatment (BH), Lane 5-7: Heat treatment at 60°C, 70°C and 80°C, respectively. Marker: SM Page ruler prestained protein ladder.

Plasmid from wild type recombinant clone *E.coli* pET21\_tvshSP2 was used for transformation into the T7 Express *E.coli* cells suitable for high efficiency transformation and expression. Transformation was successful with an efficiency of  $2 \times 10^6$  cfu/ $\mu$ g plasmid DNA. The plasmids from transformed T7 competent *E.coli* cells were characterized by restriction digestion using *Xba*I and *Hind*III restriction enzymes. Double digestion yielded correct size DNA fragments at a position of 5340 bp and 514 bp (Figure 3.19).



**Figure 3.19. Screening of the colonies from transformed T7 competent *E.coli* cells by double digestion of plasmids.** Plasmids are digested with *Xba*I+*Hind*III restriction enzymes. 3: T7 Express *E.coli* pET21-tvshSP4. 4: T7 Express *E.coli* pET21-tvshSP5. Two correct DNA fragments appear on the gel: The vector backbone (5340 bp) and the other band 514 bp gene fragments. M: O'Gene Ruler DNA Ladder Mix (ThermoScientific).

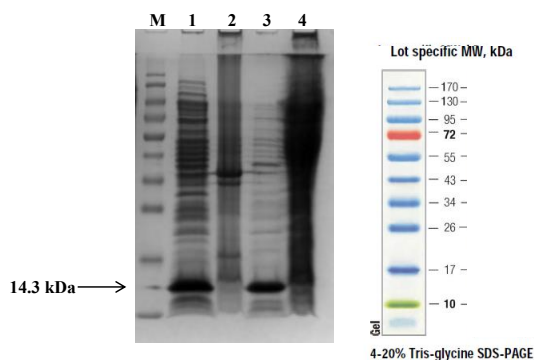
We prepared the cell lysates of T7 Express *E.coli* pET21-tvsHSP4 clones. Expression level of *Tpv* sHSP14.3 protein in the T7 Express *E.coli* cells were found less than that of BL21 competent *E.coli* cells (Figure 3.20).



**Figure 3.20. SDS PAGE gel for the expression and heat stability of recombinant T7 Express *E.coli* pET21-tvsHSP4 clones.** Lane 1: Cell free extract of the *Tpv* sHSP14.3 colony 4 before heat treatment (BH), Lane 2-3: Heat treatment at 65°C for 15 min, 70°C for 30 min, respectively. Marker: Unstained Protein Molecular Weight Marker (SM0431).

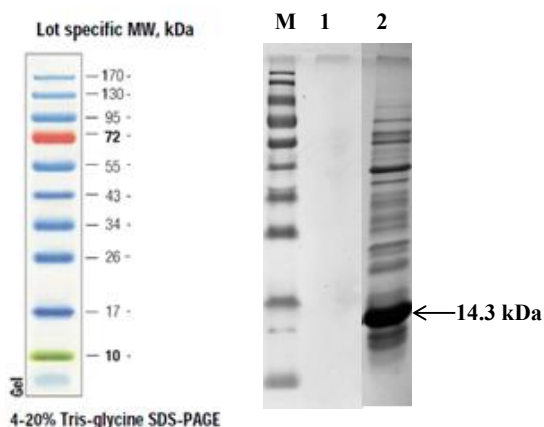
### 3.3.1 Large Scale Cell Lysate Preparation for *Tpv* sHSP14.3 Protein Purification

Cell extracts of *E.coli* pET21-tvsHSP2 cells were prepared in order to purify and express wild type *Tpv* sHSP 14.3 protein. The cell free extract was heat treated at 70°C for 30 min, and soluble and insoluble fractions of the cell lysate was analyzed by SDS-PAGE (Figure 3.21). Quite strong band of the WT protein was observed after heat treatment in the soluble fraction (Figure 3.21, Lane 3).



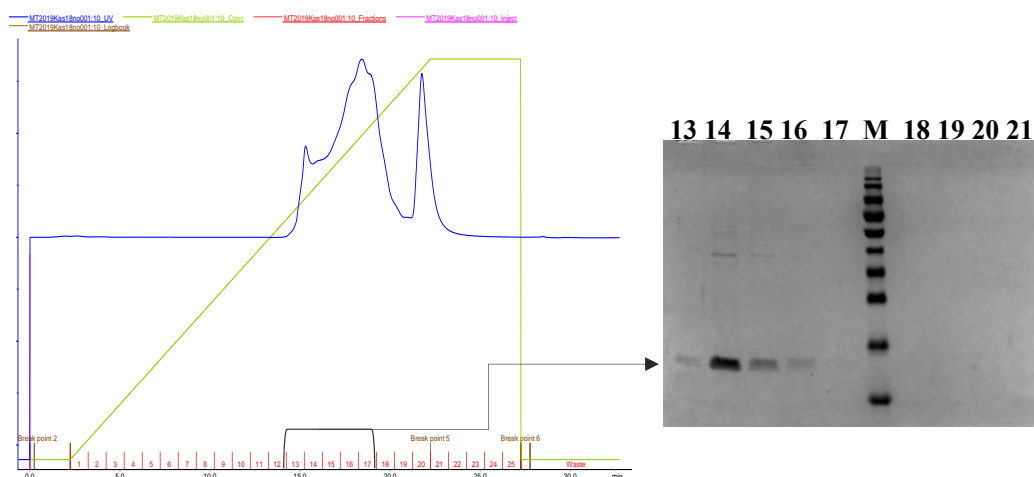
**Figure 3.21. SDS PAGE analysis of the prepared cell lysate of WT *Tpv* sHSP14.3 protein.** Lane 1: Supernatant of the WT before heat treatment. Lane 2: Pellet of WT colony 2 before heat treatment. Lane 3: Supernatant of the WT after heat treatment at 70°C. Lane 4: Pellet of WT colony 2 after heat treatment at 70°C. Marker: PageRuler Pre-stained Protein Ladder (SM0671).

Before anion exchange chromatography, buffer exchange of the sample was performed using VivaSpin 20 mL, 5K (Sartorius). *Tpv* sHSP14.3 protein in lysis buffer was replaced by Start buffer of Anion IEC (20 mM Tris Buffer, pH 8.6) and concentrated. High recovery of protein was obtained as deduced by SDS PAGE analysis (Figure 3.22).

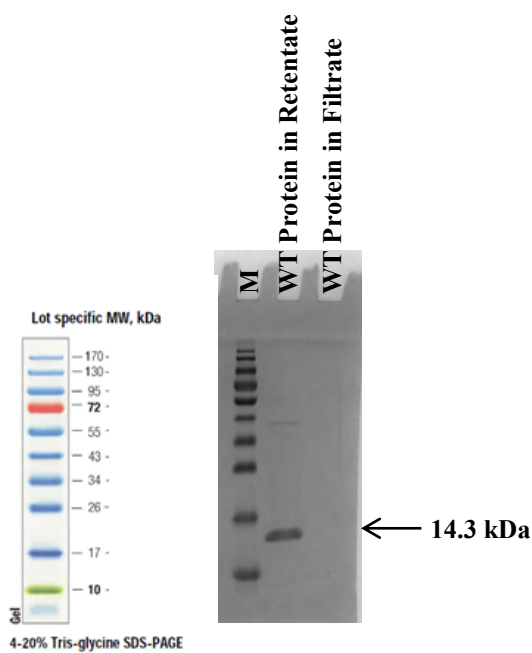


**Figure 3.22. SDS PAGE analysis of the buffer exchanged sample.** WT *Tpv* sHSP14.3 in lysis buffer was replaced by Start Buffer for Anion IEC. **Lane 1:** Filtrate, **Lane 2:** Retentate, Marker: PageRuler™ Prestained Protein Ladder (10-170 kDa).

WT *Tpv* sHSP 14.3 protein was applied to a HiTrap Q column (Amersham Biosciences, Piscataway, NJ, U.S.A) equilibrated with the Start Buffer and the bound proteins were eluted with a linear gradient of 0-1 M NaCl in the Start Buffer. One of the representative results of the HPLC purification of the wild type *Tpv* sHSP14.3 was shown in Figure 3.23. The fractions eluted by NaCl gradient were analyzed by SDS-PAGE (Figure 3.23), as well as, by the OD<sub>280</sub> measurements by pico drop. According to the UV absorbance profile, at the peak level (87 mAU), 0.5 mg/mL protein was obtained. Fractions containing *Tpv* sHSP14.3 were pooled, concentrated by ultrafiltration and buffer was exchanged by 50 mM Sodium Phosphate +150 mM NaCl, pH=7.1 (Figure 3.24).



**Figure 3.23. Chromatogram of the wild type *Tpv* sHSP14.3 purification and its SDS PAGE analysis.** Fractions containing WT *Tpv* sHSP14.3 eluted in the linear gradient of 0-1 M NaCl in start buffer. The *Tpv* sHSP14.3 is eluted at 0.6-0.7 M NaCl, corresponding to the fractions #13 to #17.

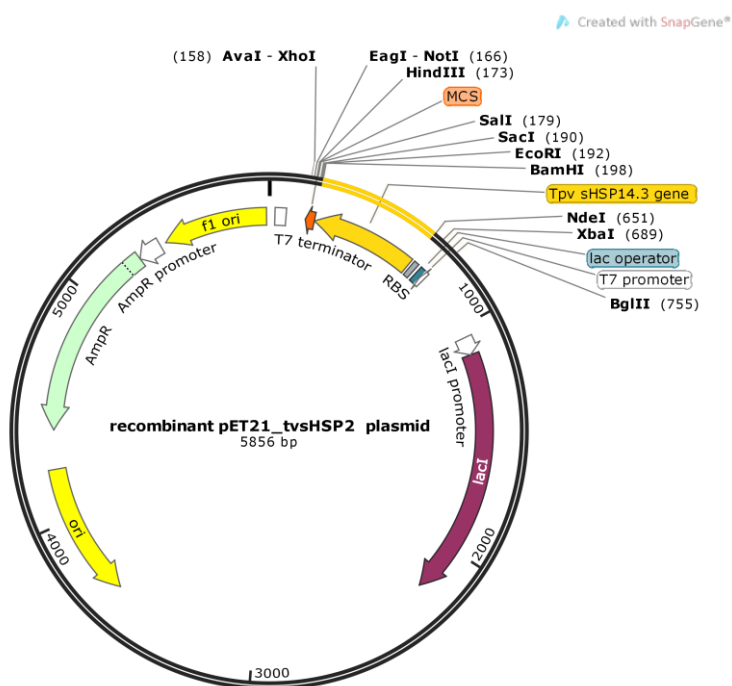


**Figure 3.24. SDS PAGE analysis of the concentrated fractions of WT *Tpv* sHSP14.3 protein after anion exchange.** Fractions of #13 to #17 from anion exchange chromatography contains the WT *Tpv* sHSP14.3 protein and these fractions are pooled and concentrated by ultrafiltration. Marker: PageRuler™ Prestained Protein Ladder (10-170 kDa)

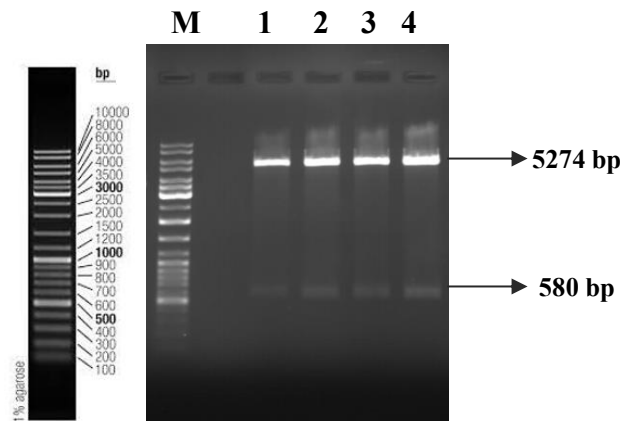
### 3.4 Site Directed Mutagenesis

#### 3.4.1 Preparation of plasmids

Recombinant plasmid pET21\_tvshSP2 was used as the source of TVN0775 gene to be used as template in the mutagenesis experiments. The diagram of this recombinant plasmid is shown in the Figure 3.25. It was isolated using the QIA Prep Spin Miniprep Kit (Qiagen). The purity and concentration were checked by Picodrop measurements. The presence of the gene was checked by double digestion with *BglIII*+*HindIII* (ThermoScientific) restriction enzymes (Figure 3.26).



**Figure 3.25. The vector map of the recombinant pET21\_tvshSP2 plasmid.** Multiple cloning site starts from the *XhoI* (158 bp) restriction enzyme cut site and ends at *NdeI* restriction enzyme cut site (651bp). *Tpv sHSP14.3* gene is inserted between *NdeI* and *BamHI* restriction enzyme sites.



**Figure 3.26. Agarose gel image of plasmids for mutagenesis.** Four recombinant plasmid (1-4) pET21\_tvshSP2 are isolated and digested with *Bgl*III+*Hind*III restriction enzymes. Two correct DNA fragments appear on the gel: The vector backbone (5274 bp) and the other band 580 bp gene fragments. M: O'Gene Ruler DNA Ladder Mix (ThermoScientific).

### 3.4.2 Generating single site mutants

Nine point mutations which we designed, (L33S, Y34F, E43V, E45G, A47D, G48E, G107A, G107D, and I108K) were generated by PCR-directed mutagenesis using Quick Change II Site-Directed Mutagenesis Kit as previously described in the Materials and Methods part. The selected positions and amino acid substitutions are shown in the DNA and amino acid sequence of the *Tpv* sHSP14.3 (Figure 3.27).

```

atgtatacaccataaagtctttacgaatgagatgataaaaaacgtatcgaatactgtg
1  M Y T P I K F F T N E M I K N V S N T V 20
aaagaggctcctatcctttatataatccaccagtcacgttatatcaagatagctctgatctg
21  K E V S S F I Y P P V T L Y Q D S S D L 40
gtattggaagcagaaatggccgggtttgacaagaaaaacataaaggctcggtaaataag
41  V L E A E M A G F D K K N I K V S V N K 60
aatgtactcactataaagtgcggagagaaagagagaatactctaccgtatataatcgatcag
61  N V L T I S A E R K R E Y S T V Y I D Q 80
cgcggtgacaaaagtgtataaagttagttagctgcccgtagagattgagcagcaggacata
81  R V D K V Y K V V K L P V E I E Q Q D I 100
tctgctaagtatagtgaggcacttacagttagaatgaaaaccaagaacataaagaac
101 S A K Y S E G I L T V R M K T K N I K N 120
gtagaaatagaataa
121 V E I E - 124

```

**Figure 3.27. Single mutation positions in the small heat shock sequence of the *Tpv* sHSP14.3.** Mutation points are marked by red color.

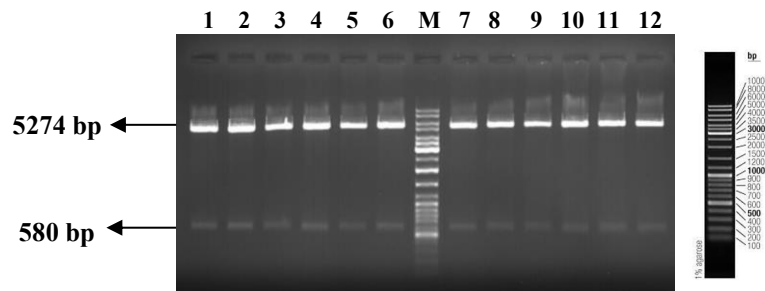


The in-vitro constructed plasmid DNA with the desired mutation were transferred into XL1-Blue supercompetent cells. The transformation efficiency varied between  $2 \times 10^4$  and  $2 \times 10^6$  cfu/ $\mu$ g DNA as shown in the Table 3.2.

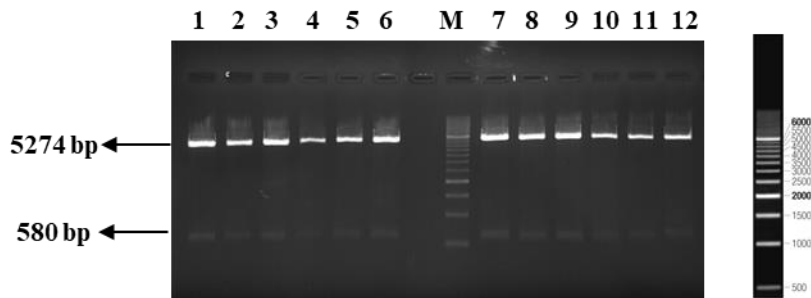
**Table 3.2. Transformation results of single mutants**

| <b>Mutation Group</b> | <b>Transformation efficiency (cfu/<math>\mu</math>g DNA)</b> |
|-----------------------|--|
| L33S                  | $4 \times 10^5$  |
| Y34F                  | $5 \times 10^4$  |
| E43V                  | $3.4 \times 10^5$  |
| E45G                  | $7 \times 10^5$  |
| A47D                  | $7 \times 10^5$  |
| G48E                  | $2 \times 10^6$  |
| G107A                 | $2.4 \times 10^5$  |
| G107D                 | $2 \times 10^5$  |
| I108K                 | $2 \times 10^4$  |

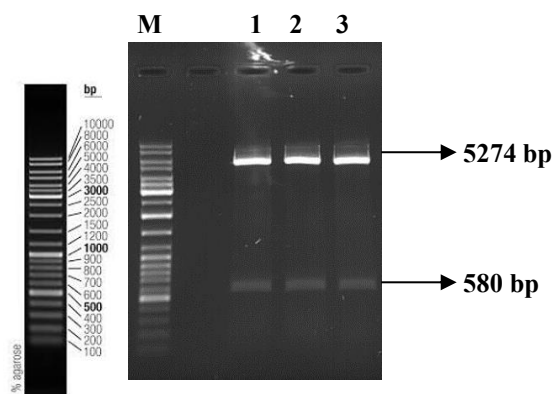
For each mutation, three putative mutant colonies were selected and their plasmids were isolated to verify mutations by sequencing. The plasmids were double digested with *Bgl*III +*Hind*III restriction enzymes to excise the cloned genes. Agarose gel electrophoresis images of digestion products are shown in Figure 3.28-3.30. Sequencing data were analyzed by performing multiple sequence alignment and also chromatogram files were checked (Figure 3.31-3.32). Overall sequencing results are given in Table 3.3. In general, sequencing results well confirmed that intended mutations were achieved at targeted positions. Out of three putative G107D mutants, two of them had another point mutation (E45G) by error. The G107D1 was used in the further experiments in this study. Also, one of the G48E putative mutants, G48E1, has a second site mutation, Y34F, besides G48E mutation. This second site error was confirmed twice by sequencing and although initially it was not planned, included in this study, as well.



**Figure 3.28. Agarose gel image of 1st set putative mutant plasmids.** Plasmids are digested with *Bgl*II+*Hind*III restriction enzymes. Two correct DNA fragments appear on the gel: The vector backbone (5274 bp) and the other band 580 bp gene fragments. M: O'Gene Ruler DNA Ladder Mix (ThermoScientific). 1 to 3: E45G; 4 to 6: G48E; 7 to 9: G107A; 10 to 12: I108K.



**Figure 3.29. Agarose gel image of 2nd set putative mutant plasmids.** Plasmids are digested with *Bgl*II+*Hind*III restriction enzymes. Two correct DNA fragments appear on the gel: The vector backbone (5274 bp) and the other band 580 bp gene fragments. M: O'Gene Ruler 500 bp DNA Ladder (ThermoScientific). 1 to 3: L33S; 4 to 6: Y34F; 7 to 9: E43V, 10 to 12: A47D.



**Figure 3.30. Agarose gel image of 3rd set of putative mutant plasmids.** Plasmids are digested with *Bgl*II+*Hind*III restriction enzymes. Two correct DNA fragments appear on the gel: The vector backbone (5274 bp) and the other band 580 bp gene fragments M: O'Gene Ruler DNA Ladder Mix (ThermoScientific). 1 to 3: G107D.

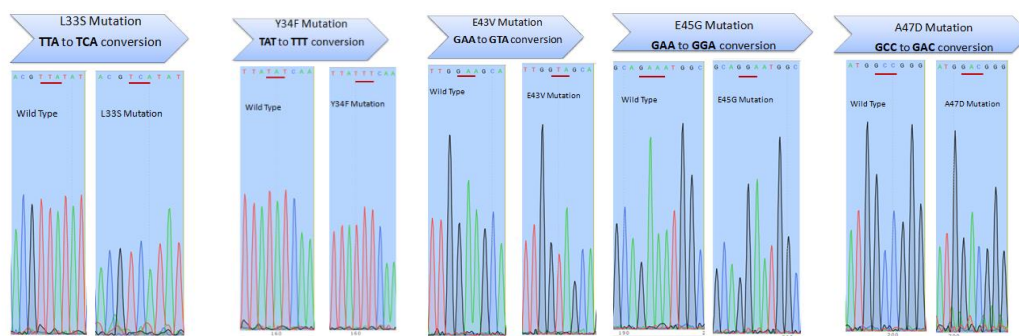


Figure 3.31. Chromatogram images of L33S, Y34F, E43V, E45G, and A47D single mutations.

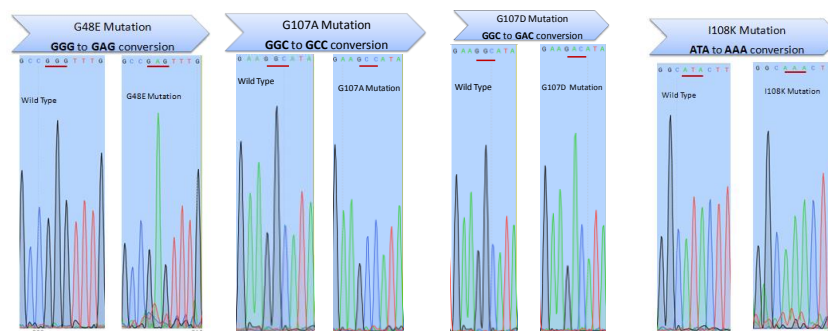


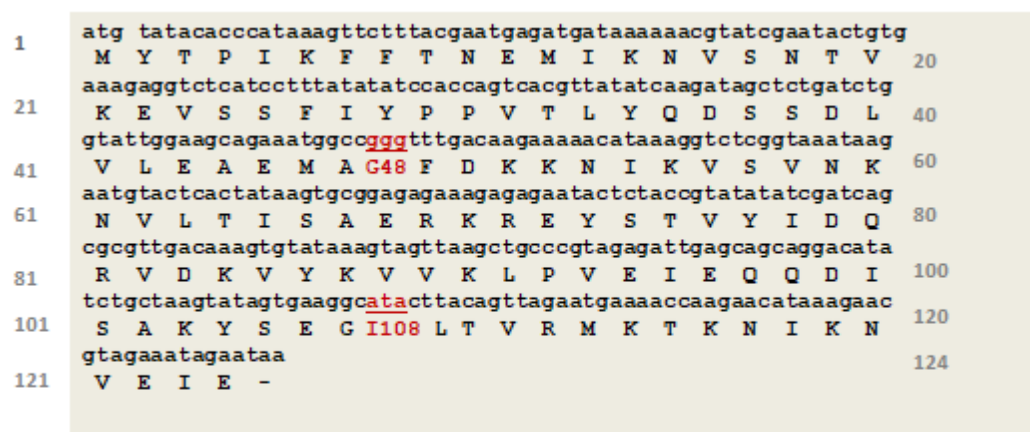
Figure 3.32. Chromatogram images of G48E, G107A, G107D and I108K single mutations.

Table 3.3. Summary of the sequencing results of the single amino acid substitutions

| Expected mutation | Sequence Result with Forward Primer | Sequence Result with Reverse primer | Label for mutation |
|-------------------|-------------------------------------|-------------------------------------|--------------------|
| <b>L33S</b>       | expected                            | expected                            | L33S1              |
|                   | Reading is not good                 | expected                            | L33S2              |
|                   | expected                            | expected                            | L33S3              |
| <b>Y34F</b>       | expected                            | expected                            | Y34F1              |
|                   | expected                            | expected                            | Y34F2              |
|                   | expected                            | expected                            | Y34F3              |
| <b>E43V</b>       | expected                            | expected                            | E43V1              |
|                   | expected                            | expected                            | E43V2              |
|                   | expected                            | expected                            | E43V3              |
| <b>E45G</b>       | expected                            | expected                            | E45G1              |
|                   | expected                            | expected                            | E45G2              |
|                   | expected                            | expected                            | E45G3              |
| <b>A47D</b>       | expected                            | expected                            | A47D1              |
|                   | expected                            | expected                            | A47D2              |
|                   | expected                            | expected                            | A47D3              |
| <b>G48E</b>       | Second site Error                   | Second site Error                   | Y34F /G48E1        |
|                   | expected                            | expected                            | G48E2              |
|                   | expected                            | expected                            | G48E3              |
| <b>G107A</b>      | expected                            | expected                            | G107A1             |
|                   | expected                            | expected                            | G107A2             |
|                   | expected                            | expected                            | G107A3             |
| <b>G107D</b>      | expected                            | expected                            | 107D1              |
|                   | Second site error E45G              | Second site error E45G              | 107D2              |
|                   | Second site error E45G              | Second site error E45G              | 107D3              |
| <b>I108K</b>      | expected                            | expected                            | I108K1             |
|                   | expected                            | expected                            | I108K2             |
|                   | expected                            | expected                            | I108K3             |

### 3.4.3 Generating Double Site Mutants

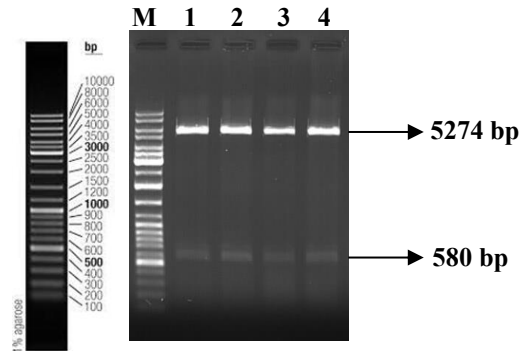
Double mutations G48E and I108K was generated in the *Tpv* sHSP gene by site directed mutagenesis using GeneArt Site Directed Mutagenesis Plus Kit as described in the Materials and Methods. The positions of the mutations in the DNA and amino acid sequence of the *Tpv* sHSP 14.3 are shown in the Figure 3.33.



**Figure 3.33. Double mutation positions in the small heat shock sequence of the *Tpv* sHSP14.3.** Mutation points are marked by red color.

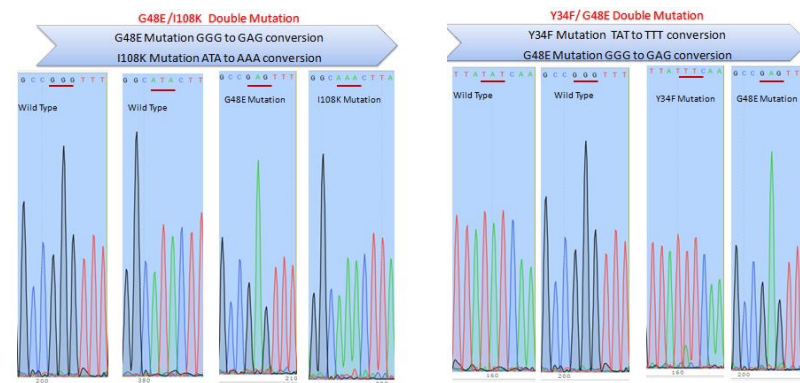
Recombinant plasmid pET21\_tvsHSP2 was used as the source of cloned *Tpv* sHSP14.3 gene. After the mutagenic PCR amplification, the PCR products were mixed together in the recombination reaction and then transformed into One Shot<sup>®</sup> MAX Efficiency<sup>®</sup> DH5 $\alpha$ <sup>™</sup>-T1R competent *E. coli* cells. Calculated transformation efficiencies were  $1 \times 10^6$  cfu/ $\mu$ g.

The mutations were verified by sequencing in the three putative mutant colonies that were randomly selected. Together with the double mutant G48E I108K, Y34FG48E mutant formed by a second site mutation during G48 to E mutagenesis, also included in our research for further characterization. Agarose gel electrophoresis images of their plasmids after double digestions with *Bgl*III+*Hind*III are shown in Figure 3.34.



**Figure 3.34. Agarose gel image of putative double mutant plasmids.** Plasmids are digested with *Bgl*II+*Hind*III restriction enzymes. Two correct DNA fragments appear on the gel: The vector backbone (5274 bp) and the other band 580 bp gene fragments M: O'Gene Ruler DNA Ladder Mix (ThermoScientific). 1 to 3: G48E I108K; 4: Y34FG48E.

Sequencing data with two primers were analyzed by performing multiple sequence alignment, as well as the chromatograms (Figure 3.35).

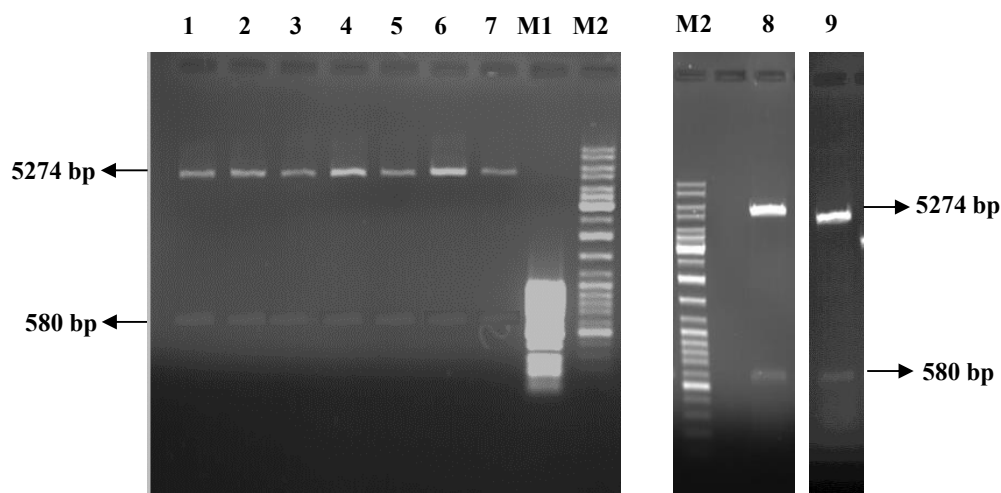


**Figure 3.35. Chromatogram image of the G48E/I108K and Y34FG48E double mutations.**

### 3.5 Transformation into the Competent BL21(DE3) *E.coli* for Protein Expression

In site directed mutagenesis, *Tpv* sHSP ACD mutant plasmids were transformed into XL1-Blue supercompetent cells which do not carry T7 RNA polymerase. Since *Tpv* genes cloned in pET21 vector are under the control of a T7 promoter, transformation to BL21(DE3) competent cell is essential for expression. For this purpose, out of

three mutant strains for each mutant group, only one was selected for plasmid isolation and transformation. Plasmid DNA after double digestion with *Bgl*III and *Hind*III restriction enzymes were analyzed by agarose gel electrophoresis (Figure 3.36).



**Figure 3.36. Agarose gel image of plasmids DNA of *Tpv* sHSP14.3 single mutants.** Plasmids are digested with *Bgl*III+*Hind*III restriction enzymes. Two correct DNA fragments appear on the gel: The vector backbone (5274 bp) and the other band 580 bp gene fragments. **1:** L33S3(1-5-2); **2:** Y34F3 (2-5-2); **3:** E43V3(3-5-1); **4:** E45G3(1-5-5); **5:** A47D1(4-1-2); **6:** G48E3(2-5-10); **7:** I108K1(4-1), **8:** G107A3, **9:** G107D1. M1: O'GeneRuler 50 bp DNA Ladder (ThermoScientific). M2: O' Gene Ruler DNA Ladder Mix (ThermoScientific).

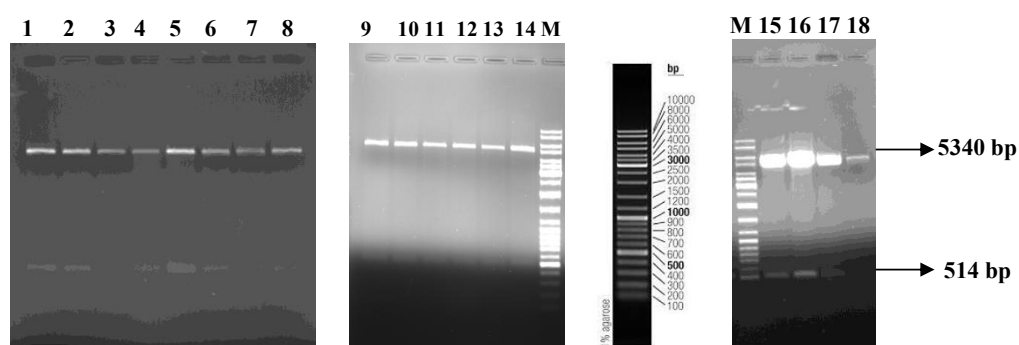
Single mutants, L33S, Y34F, E43V, E45G, A47D, G48E, and I108K were transformed into BL21(DE3) competent *E. coli* cells as described by Chung *et al.*, 1989. Two single mutants, G107A and G107D were transformed into *E. coli* BL21 competent cells (C2527H) according to the manufacturer's protocol (New England Biolabs).

Transformation efficiency of these competent cells that were checked using the control plasmid pUC19 (NEB) was varied between  $2 \times 10^6$  and  $7 \times 10^6$  cfu/ $\mu$ g. The efficiency of transformation using mutant plasmids were listed in Table 3.4.

**Table 3.4. Transformation efficiency with *Tpv* sHSP mutant plasmids.** Sample names chosen for further experiment given in the bracket.

| Sample                  | TE (cfu/ $\mu$ g)                 |
|-------------------------|-----------------------------------|
| <b>L33S3(1-5-2/1)</b>   | $5 \times 10^3$ - $1 \times 10^4$ |
| <b>Y34F3 (2-5-2/2)</b>  | $8 \times 10^3$ - $1 \times 10^4$ |
| <b>E43V3(3-5-1/3)</b>   | $6 \times 10^3$ - $1 \times 10^4$ |
| <b>E45G3 (1-5-5/4)</b>  | $4 \times 10^3$ - $2 \times 10^4$ |
| <b>A47D1(4-1-2/5)</b>   | $2 \times 10^3$                   |
| <b>G48E3 (2-5-10/6)</b> | $7 \times 10^3$ - $1 \times 10^4$ |
| <b>I108K1(4-1/7)</b>    | $4$ - $6 \times 10^3$             |
| <b>G107A3(3-4-11)</b>   | $9 \times 10^4$ - $2 \times 10^5$ |
| <b>G107D1 (2-6)</b>     | $10^5$ - $2 \times 10^5$          |

The plasmids from randomly selected two colonies of each mutant were characterized by restriction digestion using *Xba*I and *Hind*III restriction enzymes. Double digestion yielded correct size DNA fragments at a position of 5340 bp and 514 bp which were visualized by agarose gel electrophoresis (Figure 3.37).

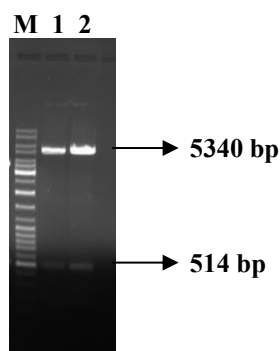


**Figure 3.37. Screening of the colonies by double digestion of plasmids 1 to 18.** Plasmids are digested with *Xba*I+*Hind*III restriction enzymes. Two correct DNA fragments appear on the gel: The vector backbone (5340 bp) and the other band 514 bp gene fragments. **1:** L33S3(1-5-2/1-3); **2:** L33S3(1-5-2/1-8); **3:** Y34F3 (2-5-2/2-1); **4:** Y34F3 (2-5-2/2-6); **5:** E43V3(3-5-1/3-4); **6:** E43V3(3-5-1/3-5); **7:** E45G3(1-5-5/4-4); **8:** I108K1(4-1/7-8); **9:** E45G3(1-5-5/4-6); **10:** A47D1(4-1-2/5-2); **11:** A47D1(4-1-2/5-3); **12:** G48E3(2-5-10/6-1); **13:** G48E3(2-5-10/6-5); **14:** I108K1(4-1/7-4); **15:** G107A3 (3-4-11/3\_4); **16:** G107A3 (3-4-11/3\_8); **17:** G107D1(2-6/4\_4); **18:** G107D1 (2-6/4\_6), M: DNA Ladder Mix (ThermoScientific).

### 3.6 Transformation into the T7 Express *E.coli* cells for protein expression

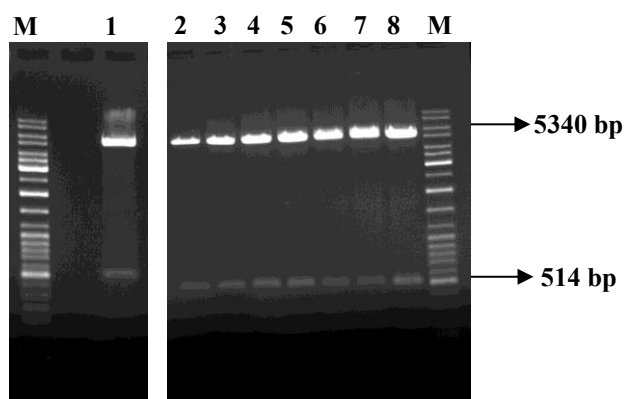
Mutant plasmids G48EI108K, Y34FG48E of double mutations, and of G48E3 and E43V3 single mutants, were transformed into T7 express competent cells, which

yield high efficiency transformation and protein expression. Transformation efficiency of the T7 cells that we have calculated according to manufacturer's instructions was  $3 \times 10^8$  cfu/ $\mu$ g. Plasmid DNA isolated to be used in the transformation after double digestion with *Xba*I and *Hind*III visualized by agarose gel electrophoresis (Figure 3.37 & 3.38).



**Figure 3.38. Agarose gel image of isolated plasmid DNA of *Tpv* sHSP14.3 double mutants.** Plasmids are digested with *Xba*I+*Hind*III restriction enzymes. Two correct DNA fragments appear on the gel. The vector backbone (5340 bp) and the other band 514 bp gene fragments. **1:** G48EI108K; **2:** Y34FG48E. M: DNA Ladder Mix (ThermoScientific).

Transformation performed according to the manufacturer's protocol yielded high efficiency ranging between  $2 \times 10^7$  to  $6 \times 10^8$  cfu/ $\mu$ g. The plasmid DNA from randomly selected two colonies of each mutant group were extracted and were double digested with *Xba*I/*Hind*III enzymes (Figure 3.39).



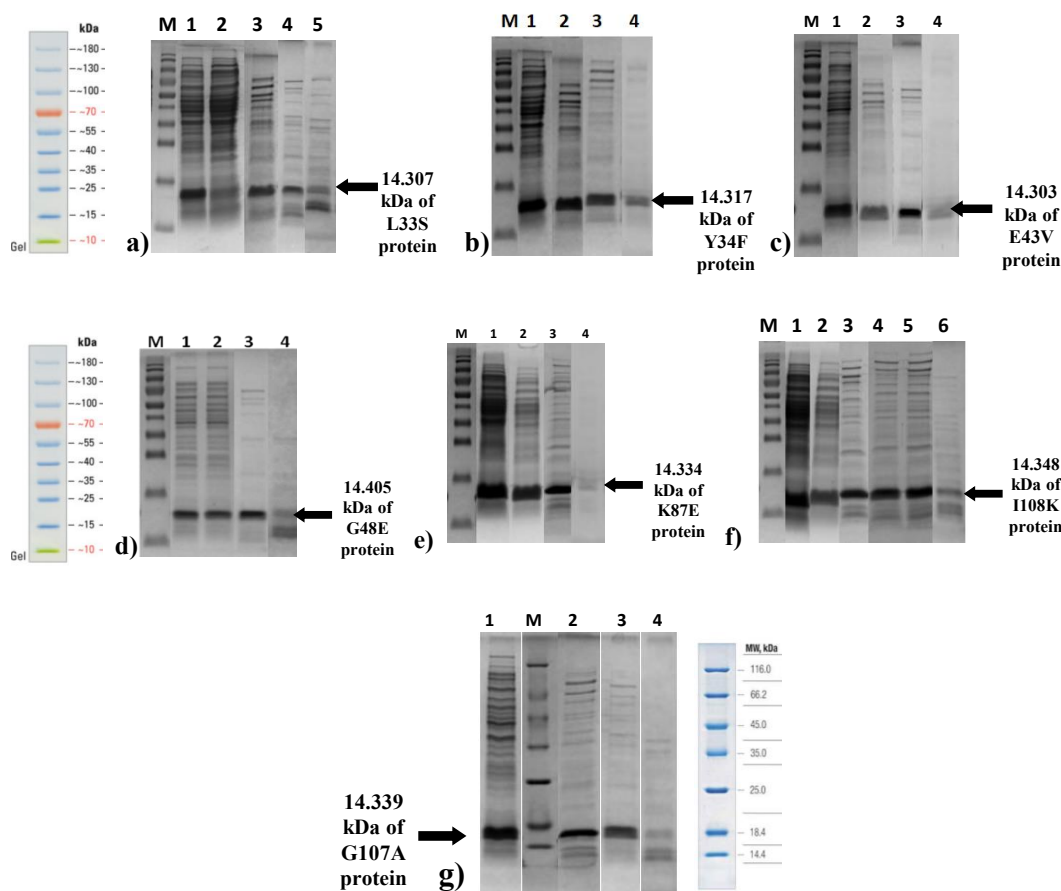
**Figure 3.39. Screening of the colonies by double digestion of plasmids 1 to 8.** Plasmids are digested with *Xba*I+*Hind*III restriction enzymes. Two correct DNA fragments appear on the gel. The vector backbone (5340 bp) and the other band 514 bp gene fragments. **1 and 2:** E43V3; **3 and 4:** G48E3; **5 and 6:** Y34FG48E3; **7 and 8:** G48EI108K1, M: DNA Ladder Mix (ThermoScientific)



### 3.7 Expression and Initial Purification of Mutant Proteins

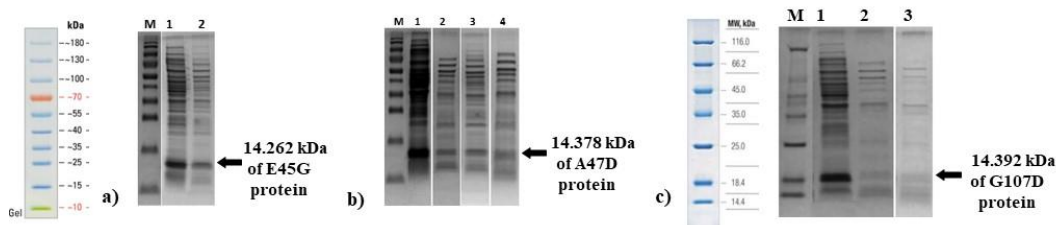
Small-scale cultures of mutant *E.coli* BL21 and T7 cells were prepared for initial expression screening and to determine the optimum expression parameters. All recombinant mutant proteins in the soluble fraction of the cell lysates were found to be expressed at high levels, as revealed by SDS PAGE images of the protein bands (14.3 kDa) (Figure 3.40). Absence or weak band observed in the cell lysate of the negative control (without *Tpv* sHSP14.3 plasmid) indicating the effectiveness of mutant protein's expression after induction in BL21(DE3) cells (Figure 3.40-a, Lane 2).

To get rid of heat-labile proteins of *E.coli* cells overexpressing ACD mutant proteins, cell lysates were subjected to heat treatments at temperatures between 60°C and 80°C. Overexpression of the single mutant L33S in BL21 showed good expression (Figure 3.40-a). Cell lysate of L33S mutant when heat-treated at 60°C for 15 min, significant amount of host proteins remained undenatured (Figure 3.40-a, Lane 3). However, majority of the *E.coli* originated proteins were eliminated after heating at 70°C for 30 min or at 80°C for 15 min, while L33S mutant sHSP protein remained stable (Figure 3.40-a, Lane 4-5). Similarly, *E.coli* BL21/Y34F, *E.coli* T7 Express/E43V and *E.coli* T7 Express/G48E mutant variants were not affected from heat treatment up to 70°C, as deduced from band intensities on SDS polyacrylamide gel (Figure 3.40-b-c-d, Lane 3). However, heating at 80°C resulted in loss of the *Tpv* sHSP protein stability to the extent that vary according to the mutant strain. All of these single point mutant proteins of sHSP were resistant to high temperature up to 70°C. In contrast to wild type, their heat-stability reduced upon heating at 80°C (15 min). While *E.coli* BL21/I108K and *E.coli* BL21/Y34F sHSP proteins were the least effected mutant proteins, *E.coli* BL21/K87E mutant sHSP was completely denatured by heating at 80°C (Figure 3.40-e, Lane 4).



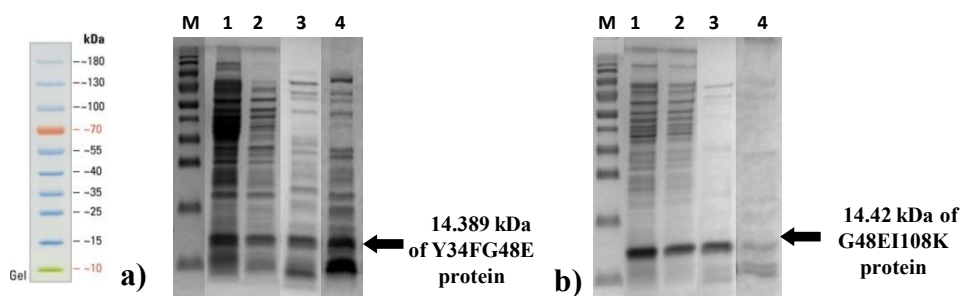
**Figure 3.40. SDS PAGE analysis of *Tpv* sHSP14.3 L33S, Y34F, E43V, G48E, K87E, I108K and G107A single mutants before and after heat treatment at 60-80°C.** **a)** SDS PAGE analysis of L33S mutant sHSP before and after heat treatment. M: SM Page ruler prestained protein ladder, **Lane 1:** Cell free extract of the *Tpv* sHSP14.3 L33S before heat treatment (BH), **Lane 2:** Negative Control BL21, **Lane 3-5:** Heat treatment of L33S mutant protein at 60°C, 70°C, and 80°C, respectively. **b)** SDS PAGE analysis of Y34F mutant sHSP before and after heat treatment. M: SM Page ruler prestained protein ladder. **Lane 1:** Cell free extract of the *Tpv* sHSP14.3 Y34F before heat treatment (BH), **Lane 2-4:** Heat treatment of Y34F mutant protein at 60°C, 70°C, and 80°C, respectively. **c)** SDS PAGE analysis of E43V mutant sHSP before and after heat treatment. M: SM Page ruler prestained protein ladder. **Lane 1:** Cell free extract of the *Tpv* sHSP14.3 E43V before heat treatment (BH), **Lane 2-4:** Heat treatment of E43V mutant protein at 60°C, 70°C, and 80°C, respectively. **d)** SDS PAGE analysis of G48E mutant sHSP before and after heat treatment. M: SM Page ruler prestained protein ladder. **Lane 1:** Cell free extract of the *Tpv* sHSP14.3 G48E before heat treatment (BH), **Lane 2-4:** Heat treatment of G48E mutant protein at 60°C, 70°C, and 80°C, respectively. **e)** SDS PAGE analysis of K87E mutant sHSP before and after heat treatment. M: SM Page ruler prestained protein ladder. **Lane 1:** Cell free extract of the *Tpv* sHSP14.3 K87E before heat treatment (BH), **Lane 2-4:** Heat treatment of K87E mutant protein at 60°C, 70°C, and 80°C, respectively. **f)** SDS PAGE analysis of I108K mutant sHSP before and after heat treatment. M: SM Page ruler prestained protein ladder. **Lane 1:** Cell free extract of the *Tpv* sHSP14.3 I108K before heat treatment (BH), **Lane 2-6:** Heat treatment of I108K mutant protein at 60°C, 65°C, 70°C, 68°C and 80°C, respectively. **g)** SDS PAGE analysis of G107A mutant sHSP before and after heat treatment at 60-80°C. M: The Unstained Protein Molecular Weight Marker (SM0431, Fermentas). **Lane 1:** Cell free extract of the *Tpv* sHSP14.3 G107A before heat treatment (BH), **Lane 2-4:** Heat treatment of G107A mutant protein at 65°C, 70°C, and 80°C, respectively.

G107D, E45G and A47D mutant proteins although well expressed in *E. coli* BL21 cells, they had very low heat stability as compared to the WT and other mutant variants. After heating at or above  $\geq 60$  °C (for 15 min), protein amount was considerably lost (Figure 3.41).



**Figure 3.41. SDS PAGE analysis of *Tpv* sHSP14.3 E45G, A47D and G107D mutants before and after heat treatment at 60-80°C.** a) SDS PAGE analysis of E45G mutant sHSP before and after heat treatment. M: SM Page ruler prestained protein ladder, **Lane 1:** Cell free extract of the *Tpv* sHSP14.3 E45G before heat treatment (BH), **Lane 2:** Heat treatment of E45G mutant protein at 60°C. b) SDS PAGE analysis of A47D before and after heat treatment at 60°C. M: SM Page ruler prestained protein ladder, **Lane 1:** Cell free extract of the *Tpv* sHSP14.3 A47D before heat treatment (BH), **Lane 2-4:** Three independent heat treatment of A47D mutant protein at 60°C. c) SDS PAGE analysis of G107D mutant sHSP before and after heat treatment at 60-80°C. M: The Unstained Protein Molecular Weight Marker (SM0431, Fermentas). **Lane 1:** Cell free extract of the *Tpv* sHSP14.3 G107D before heat treatment (BH), **Lane 2-3** Heat treatment of G107D mutant protein at 60°C and, 70°C, respectively.

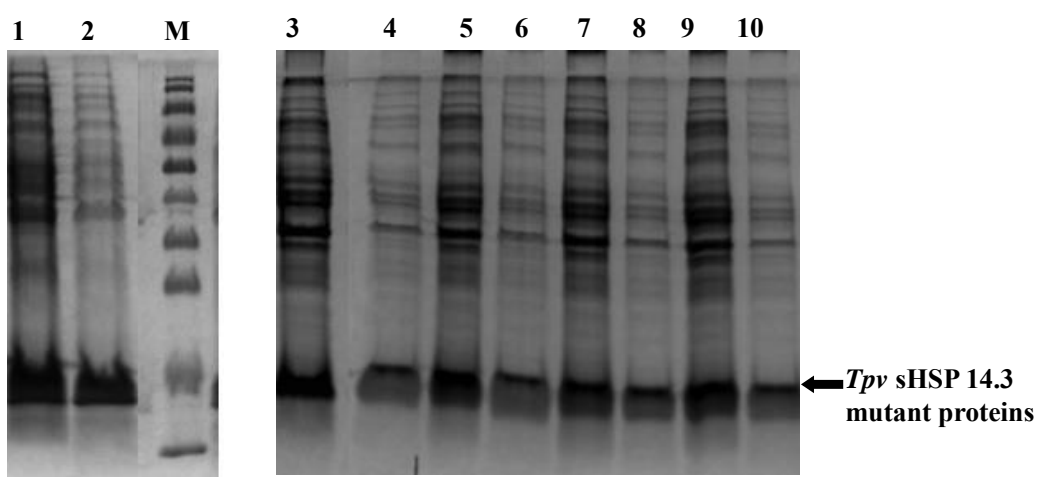
Double mutant plasmid constructs, Y34FG48E and G48EI108K, which were transformed into the *E. coli* T7 Express cells expressed the mutant proteins at high levels. Both mutant proteins remained stable after heating at 70°C (Figure 3.42). Although Y34FG48E double mutant protein amount almost did not change like its single mutant variants after heating at 80°C (15 min), there has been significant loss in the amount of G48EI108K mutant sHSP at that high temperature.



**Figure 3.42. SDS PAGE analysis of *Tpv* sHSP14.3 double mutant proteins before and after heat treatment at 60-80°C.** a) Y34FG48E *Tpv*-sHSP14.3 b) G48EI108K *Tpv*-sHSP14.3, M: SM Page ruler prestained protein ladder. **Lane 1:** Cell free extract of the *Tpv* sHSP14.3 before heat treatment (BH), **Lane 2-4:** Heat treatment at 60°C, 70°C, and 80°C, respectively.

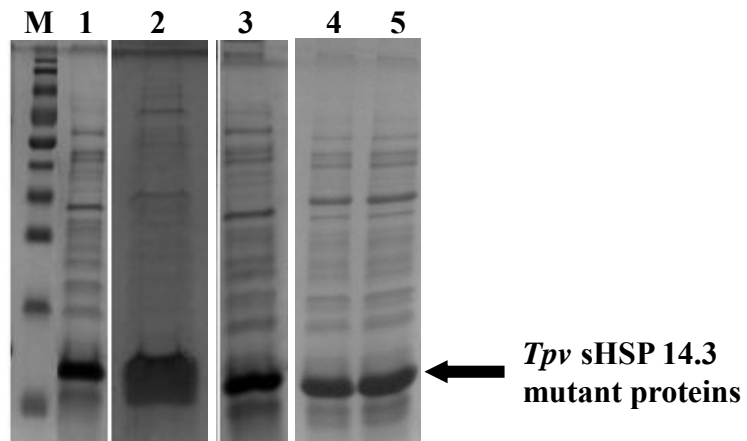
### 3.8 HPLC Purification of the *Tpv* sHSP14.3 ACD Mutants

For large scale purification of mutant sHSP protein samples by HPLC chromatography, the amount of cell cultures was increased 10x fold (1lt) to prepare cell free extracts. Clarified cell lysates were heated at 70°C for 30 min in order to eliminate heat labile *E.coli* proteins. Images of lysates before and after heat treatment obtained by SDS-PAGE for the selected mutant proteins are shown in Figure 3.43.

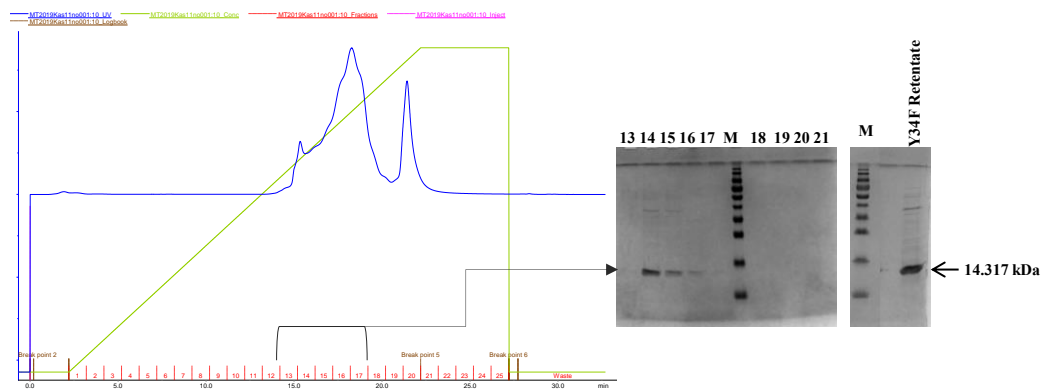


**Figure 3.43. SDS PAGE analysis of the prepared five cell lysates.** Supernatant of the before heat treatment and after heat treatment at 70°C. **Lane 1-2:** Y34F (14.317 kDa), **Lane 3-4:** E43V (14.303 kDa), **Lane 5-6:** G48E (14.405 kDa), **Lane 7-8:** K87E (14.334 kDa), **Lane 9-10:** I108K (14.348 kDa), respectively. **Marker:** PageRuler Pre-stained Protein Ladder (SM0671).

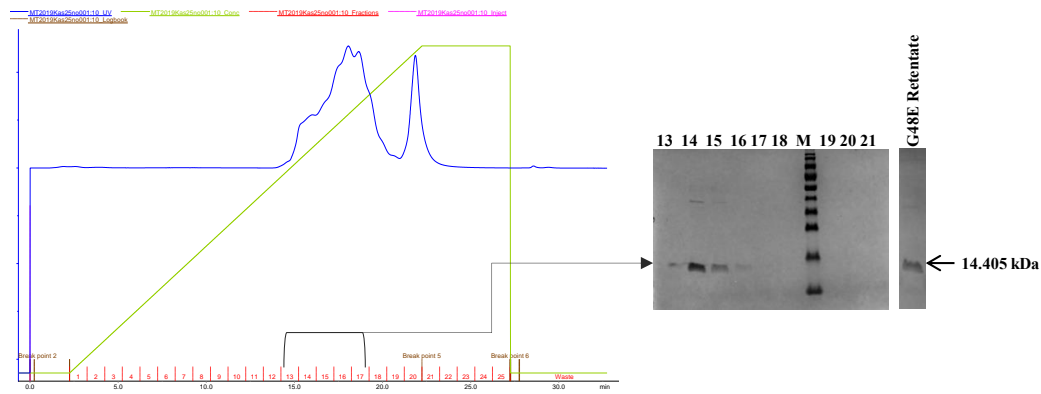
Before purification, cell free extracts in lysis buffer were subjected to buffer exchange by start buffer of anion exchange chromatography. High recovery of the protein was obtained in the retentate after ultrafiltration, while protein was not found in the filtrate, as deduced by SDS PAGE analysis (Figure 3.44). A representative anion exchange chromatography process for *E.coli* BL21/Y34F, G48E and I108K is shown below (Figure 3.45-3.47). In the mutant protein purification process, 15-20 mg of the protein was loaded on the HiTrap Q column, and eluted. When the fractions eluted with NaCl gradient were run on a SDS gel, the protein bands observed in the fractions from no.13 to no.17. According to the UV absorbance profile, the general trend was that at the peak level (85 mAU), 0.5-0.6 mg/mL protein was purified. These fractions were combined and after buffer exchange, concentrated protein was obtained.



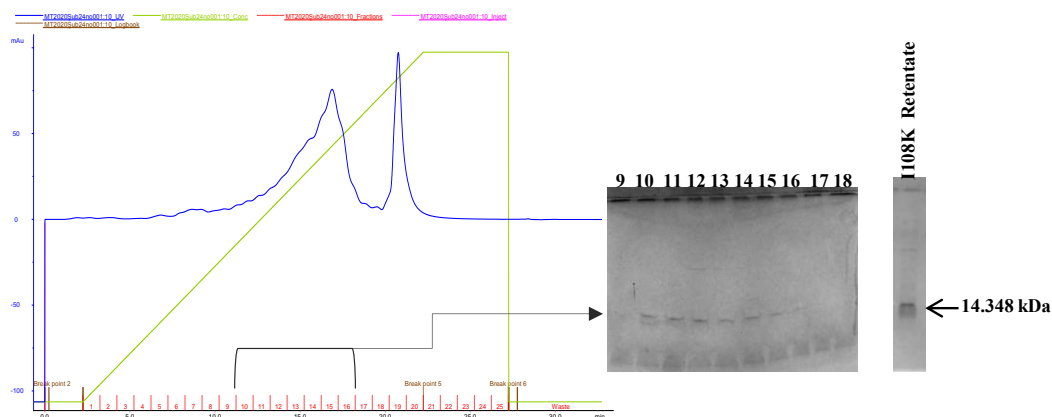
**Figure 3.44. SDS PAGE analysis of the buffer exchanged samples.** *Tpv* sHSP14.3 ACD mutant proteins in lysis buffer are replaced by Start Buffer for Anion IEC. **Lane 1- 5:** Retentate of the Y34F (14.317 kDa), E43V (14.303 kDa), G48E (14.405 kDa), K87E (14.334 kDa) and I108K (14.348 kDa) mutant proteins, respectively. Marker: PageRuler™ Prestained Protein Ladder (10-170 kDa).



**Figure 3.45. Chromatogram of the *Tpv* sHSP14.3 Y34F mutant protein purification and its SDS-PAGE analysis.** Fractions containing Y34F *Tpv* sHSP14.3 elute in the linear gradient of 0-1 M NaCl in start buffer. The Y34F *Tpv* sHSP14.3 protein obtained at 0.6-0.8 M NaCl, corresponding to the fractions 13 to 17. These fractions are pooled and concentrated by ultrafiltration. Marker: PageRuler™ Prestained Protein Ladder (10-170 kDa).



**Figure 3.46. Chromatogram of the *Tpv* sHSP14.3 G48E mutant protein purification and its SDS-PAGE analysis.** Fractions containing G48E *Tpv* sHSP14.3 elute in the linear gradient of 0-1 M NaCl in start buffer. The G48E *Tpv* sHSP14.3 protein obtained at 0.6-0.8 M NaCl, corresponding to the fractions 13 to 16. These fractions are pooled and concentrated by ultrafiltration. Marker: PageRuler™ Prestained Protein Ladder (10-170 kDa).

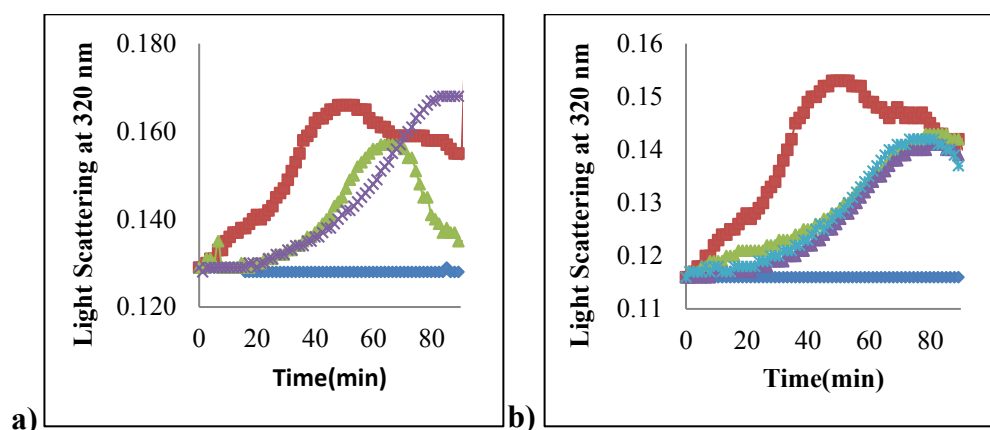


**Figure 3.47. Chromatogram of the *Tpv* sHSP14.3 I108K mutant protein purification and its SDS-PAGE analysis.** Fractions containing I108K *Tpv* sHSP14.3 elute in the linear gradient of 0-1 M NaCl in start buffer. The I108K *Tpv* sHSP14.3 protein obtained at 0.5-0.7 M NaCl, corresponding to the fractions 10 to 15. These fractions are pooled and concentrated by ultrafiltration. Marker: PageRuler™ Prestained Protein Ladder (10-170 kDa).

### 3.9 Thermal Aggregation Measurements with Pig Heart Citrate Synthase

Before performing thermal aggregation measurements, to find out optimum concentration of CS to be able to follow its aggregation process, we tested different amounts of the enzyme (*i.e.*, 10.5  $\mu\text{g}$ -35  $\mu\text{g}$ ). Based on the time course aggregation profile of CS, 10.5  $\mu\text{g}$  of the enzyme increased light scattering intensity more than the other two concentrations (21 and 35  $\mu\text{g}$ ) (Figure 3.48-a). Therefore, this concentration of the CS was used in all thermal aggregation assays.

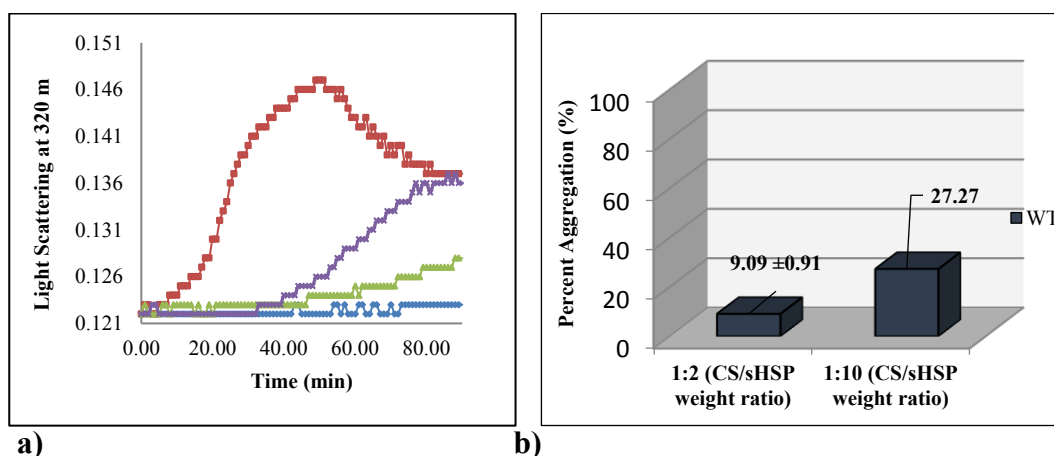
Thermal aggregation of CS was performed by measuring light scattering of Pig heart citrate synthase (1.4  $\mu\text{M}$ , 70  $\mu\text{g}/\text{mL}$ ) alone and together with *Tpv* sHSP 14.3 wild type or and its variants (at 9.8  $\mu\text{M}$  or 140  $\mu\text{g}/\text{mL}$  and 49  $\mu\text{M}$  or 700  $\mu\text{g}/\text{mL}$  concentration) at 45°C, as described in the Materials and Methods. Therefore, the effect of the *Tpv* sHSP14.3 on the prevention of the thermal aggregation of CS was assessed at 1:2 CS/sHSP w/w ratio and 1:10 CS/sHSP w/w ratio. As a control, aggregation of CS was considered 100 %, when incubated alone at 45°C, as deduced from the slope of its light scattering plot.



**Figure 3.48. Optimization of the CS (a) and *Tpv* sHSP14.3 (b) concentration for the thermal aggregation measurements. a)** (—◆—) Blank, (—■—) 10.5  $\mu\text{g}$  of CS, (—▲—) 21  $\mu\text{g}$  of CS, (—×—) 35  $\mu\text{g}$  of CS. **b)** (—◆—) Blank, (—■—) CS alone (10.5  $\mu\text{g}$ ), (—▲—) CS+Chaperone (10.3  $\mu\text{g}$ ), (—▲—) (CS+Chaperone (20.6  $\mu\text{g}$ ), (—\*—) CS+Chaperone (51.5  $\mu\text{g}$ ).

Addition of the *Tpv* sHSP 14.3 WT prevented the aggregation of CS and reduced light scattering by 91% and 73%, respectively at 1:2 CS/sHSP w/w and 1:10

CS/sHSP w/w ratio (Figure 3.49). Therefore, suppression of aggregation by WT *Tpv* sHSP14.3 is more effective at low molar ratio of CS/*Tpv* sHSP14.3.



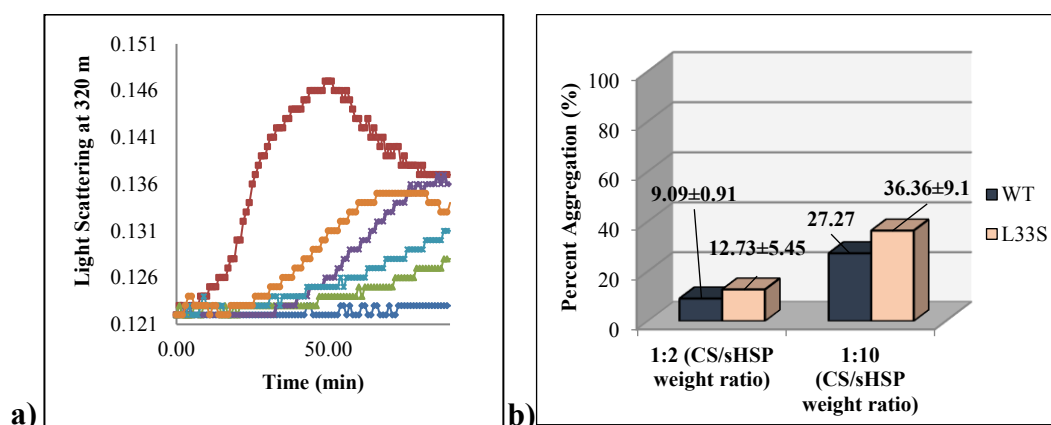
**Figure 3.49.** Effect of the *Tpv* sHSP 14.3 WT on the thermal aggregation of the CS. **a)** CS aggregation was monitored by measuring light scattering at 320 nm at 45°C. (—◆—) Blank, (—■—) CS alone, (—▲—) CS + WT (1:2 w/w ratio), (—×—) CS + WT (1:10 w/w ratio). **b)** Percent aggregation when the aggregation of CS alone (Control) was taken as 100 % (100±4.55). Each experiment was repeated at least three times.

### 3.9.1 Effect of the *Tpv* sHSP14.3 Mutants with Altered Hydrophobicity on CS Thermal Aggregation

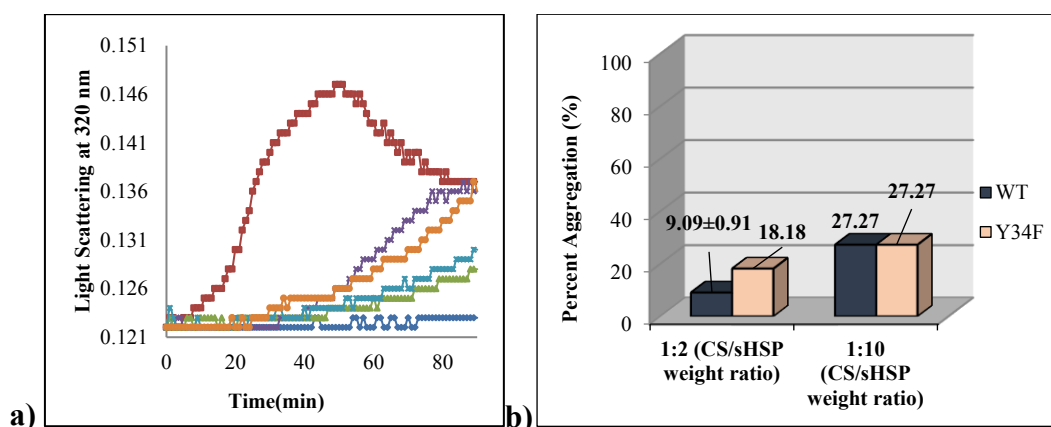
Three *Tpv* sHSP14.3 ACD mutants, L33S, Y34F and G107A, changed the hydrophobicity of the protein. CS thermal aggregation measurements showed that *Tpv* sHSP 14.3 L33S mutant, constructed by replacing hydrophobic amino acid with polar amino acid at  $\beta$ 2 sheet, showed similar chaperone activity as compared to the wild type at the 1:2 CS/sHSP w/w ratio. However, when 5-fold excess *Tpv* sHSP 14.3 L33S protein was used, protection ability of this mutant also decreased, which showed 9 % more aggregation than that of the WT sHSP (Figure 3.50). Similarly, aggregation of the CS was found to be well protected when hydrophobicity was increased at position 34 on  $\beta$ 2 sheet. Y34F mutant showed appreciable chaperone activity in reducing light scattering of CS to 18.2 % at CS/sHSP ratio of 1:2 w/w ratio, but its activity was still less than the WT. When the concentration of the Y34F mutant was increased by 5-fold, extend of the protection (*i.e.*, ~ 73 %) against thermal aggregation became equal to the that of WT sHSP at CS/chaperone ratio of



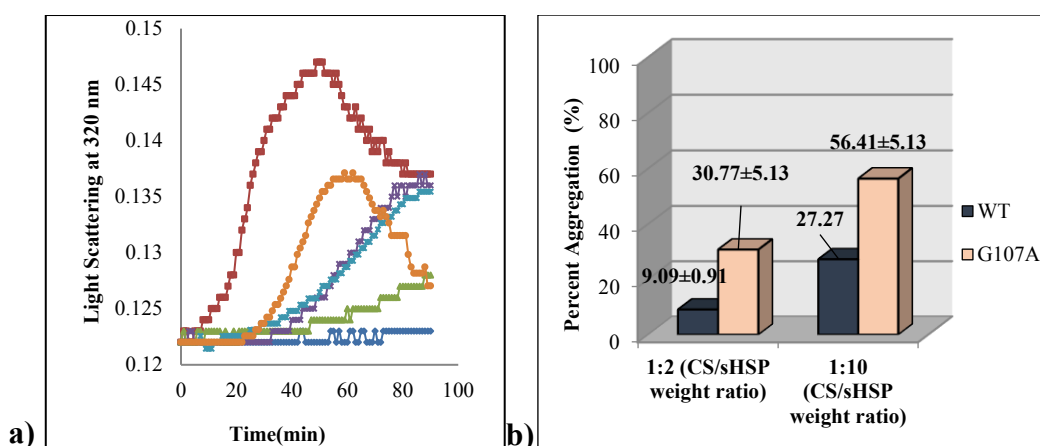
1:10 w/w ratio (Figure 3.51). On the other hand, decrease in the chaperone activity of *Tpv* sHSP14.3 was observed by G107 (residue positioned on the short loop connecting the  $\beta$ 8- $\beta$ 9 sheets) to A substitution. CS thermal aggregation measurements showed that G107A mutant protected the CS against the heat induced denaturation more effectively at 1:2 CS/sHSP w/w ratio (~ 70%) than 1:10 CS/sHSP w/w ratio (44 %). At both ratios, protection ability of G107A mutant was found less than the WT (Figure 3.52).



**Figure 3.50. Effect of the *Tpv* sHSP 14.3 L33S mutant on the thermal aggregation of the CS.** a) CS aggregation was monitored by measuring light scattering at 320 nm at 45°C. (—◆—) Blank, (—■—) CS alone, (—▲—) CS + WT (1:2 w/w ratio), (—×—) CS + WT (1:10 w/w ratio), (—\*—) CS + L33S (1:2 w/w ratio), (—○—) CS + L33S (1:10 w/w ratio). b) Percent aggregation when the aggregation of CS alone (Control) was taken as 100 % (100±4.55). Each experiment was repeated at least three times.



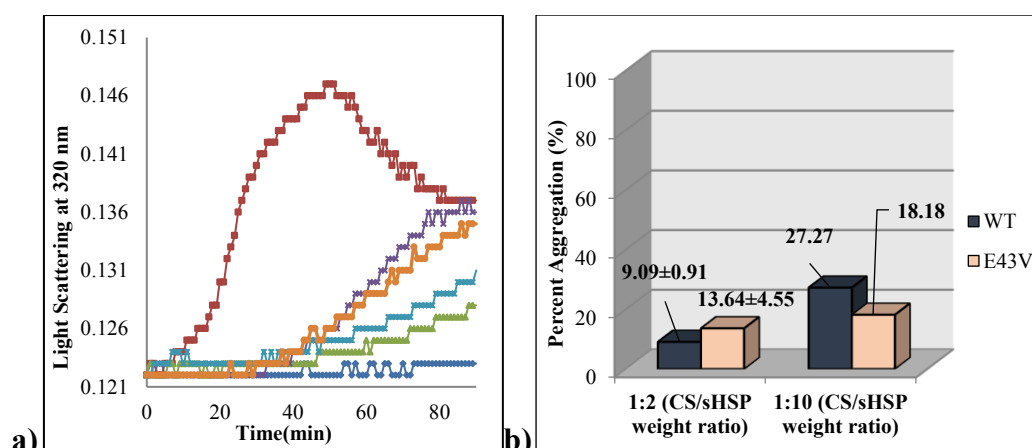
**Figure 3.51. Effect of the *Tpv* sHSP 14.3 Y34F mutant on the thermal aggregation of the CS.** a) CS aggregation was monitored by measuring light scattering at 320 nm at 45°C. (—◆—) Blank, (—■—) CS alone, (—▲—) CS + WT (1:2 w/w ratio), (—×—) CS + WT (1:10 w/w ratio), (—\*—) CS + Y34F (1:2 w/w ratio), (—○—) CS + Y34F (1:10 w/w ratio). b) Percent aggregation when the aggregation of CS alone (Control) was taken as 100 % (100±4.55). Each experiment was repeated at least three times.



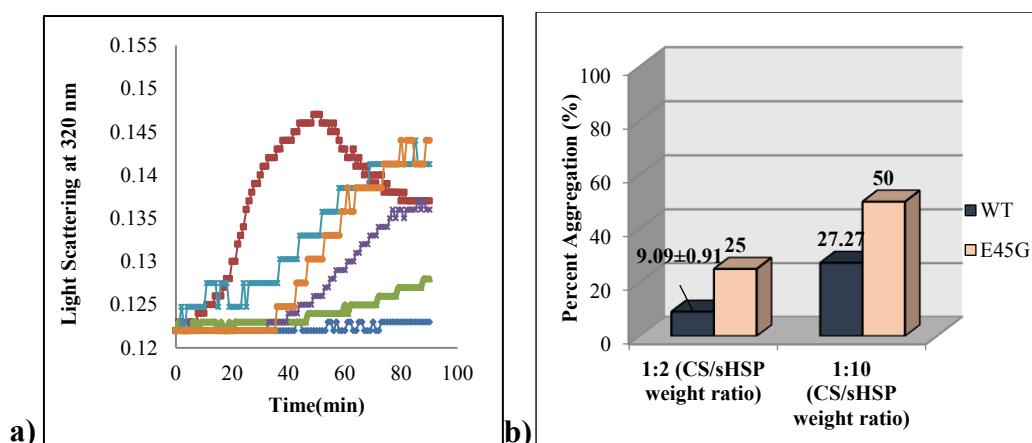
**Figure 3.52. Effect of the *Tpv* sHSP 14.3 G107A mutant on the thermal aggregation of the CS.** **a)** CS aggregation was monitored by measuring light scattering at 320 nm at 45°C. (—◆—) Blank, (—■—) CS alone, (—▲—) CS + WT (1:2 w/w ratio), (—×—) CS + WT (1:10 w/w ratio), (—\*—) CS + G107A (1:2 w/w ratio), (—○—) CS + G107A (1:10 w/w ratio). **b)** Percent aggregation when the aggregation of CS alone (Control) was taken as 100 % (100±4.55). Each experiment was repeated at least three times.

### 3.9.2 Effect of the *Tpv* sHSP14.3 Mutants that Lost Charge by Introduction of Hydrophobicity on CS Thermal Aggregation

When hydrophobic substitution on  $\beta 3$  strand by E43 to V mutation was introduced, at w/w ratio of CS/ sHSP of 1:2, chaperone activity of E43V mutant was 4.5 % less than the WT (Figure 3.53). The mutant sHSP reduced the light scattering to 13.6 %, while effect of the WT sHSP on decreasing light scattering was 9.1 %. At a 1:10 CS:sHSP ratio(w/w), repression of the CS by mutant sHSP was 9% higher than the wild type. E45 on  $\beta 3$  strand, which is highly conserved negatively charged residue in eukaryotic and prokaryotic sHSP, when replaced by hydrophobic G, chaperone activity of the *Tpv* sHSP14.3 was significantly reduced. Protection ability of this mutant was 18.18 % and 25% less than the wild type at 1:2(w/w) and 1:10 w/w ratio of CS/sHSP, respectively (Figure 3.54).



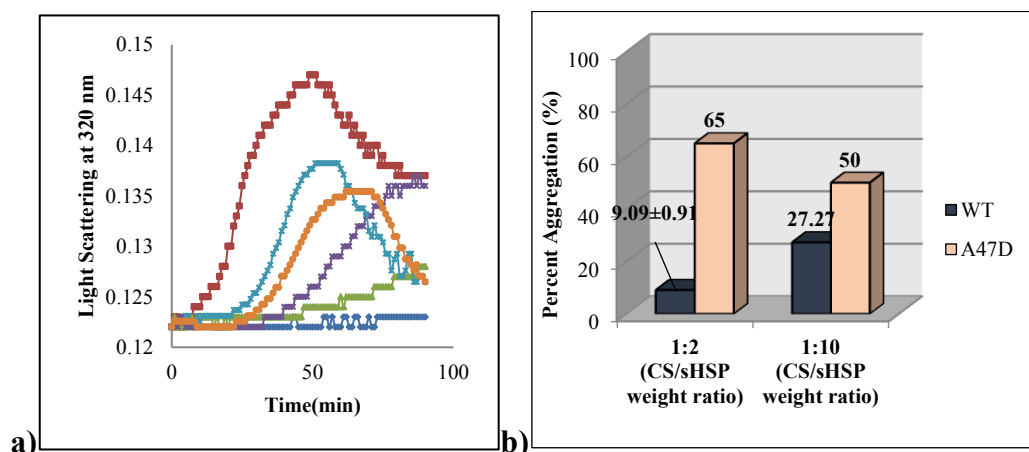
**Figure 3.53. Effect of the *Tpv* sHSP 14.3 E43V mutant on the thermal aggregation of the CS.** **a)** CS aggregation was monitored by measuring light scattering at 320 nm at 45°C. (—◆—) Blank, (—■—) CS alone, (—▲—) CS + WT (1:2 w/w ratio), (—×—) CS + WT (1:10 w/w ratio), (—\*—) CS + E43V (1:2 w/w ratio), (—○—) CS + E43V (1:10 w/w ratio). **b)** Percent aggregation when the aggregation of CS alone (Control) was taken as 100 % (100±4.55). Each experiment was repeated at least three times.



**Figure 3.54. Effect of the *Tpv* sHSP 14.3 E45G mutant on the thermal aggregation of the CS.** **a)** CS aggregation was monitored by measuring light scattering at 320 nm at 45°C. (—◆—) Blank, (—■—) CS alone, (—▲—) CS + WT (1:2 w/w ratio), (—×—) CS + WT (1:10 w/w ratio), (—\*—) CS + E45G (1:2 w/w ratio), (—○—) CS + E45G (1:10 w/w ratio). **b)** Percent aggregation when the aggregation of CS alone (Control) was taken as 100 % (100±4.55). Each experiment was repeated at least three times.

### 3.9.3 Effect of the *Tpv* sHSP14.3 Mutants with Gain of Charge (Loss of Hydrophobicity) on CS Thermal Aggregation

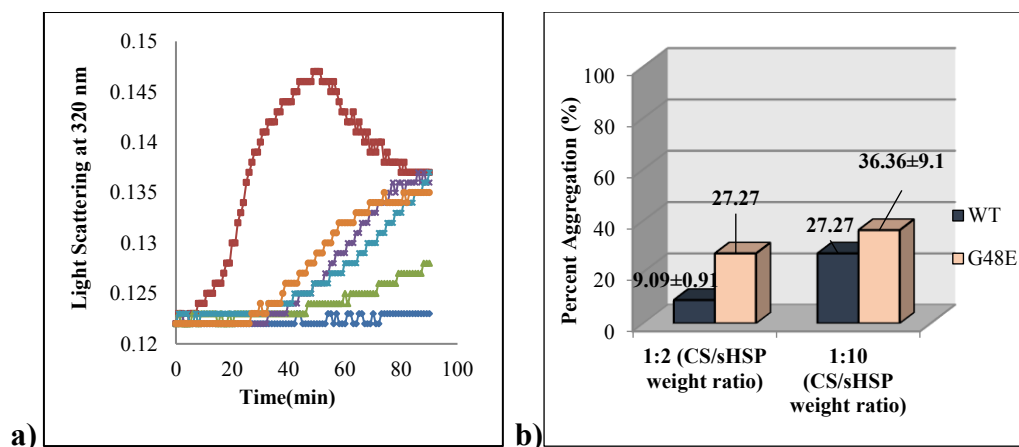
Three more *Tpv* sHSP 14.3 single mutants were constructed by replacing the hydrophobic amino acid within the ACD with the charged ones: A47D, G48E, and I108K. Among these mutants, the A47D mutant, which involves replacement of a hydrophobic residue by an acidic residue in the loop connecting two antiparallel  $\beta$ 3- $\beta$ 4 strands, did not as effectively protected the CS as the WT against heat induced aggregation at CS/sHSP ratio of 1:2 (35% protection) and 1:10 (50% protection) w/w ratios. The mutant sHSP reduced the light scattering to 65 % and 50 %, which was 7- fold and 1.8- fold less than the that of WT sHSP at 1:2 and 1:10 CS/sHSP w/w ratio, respectively (Figure 3.55).



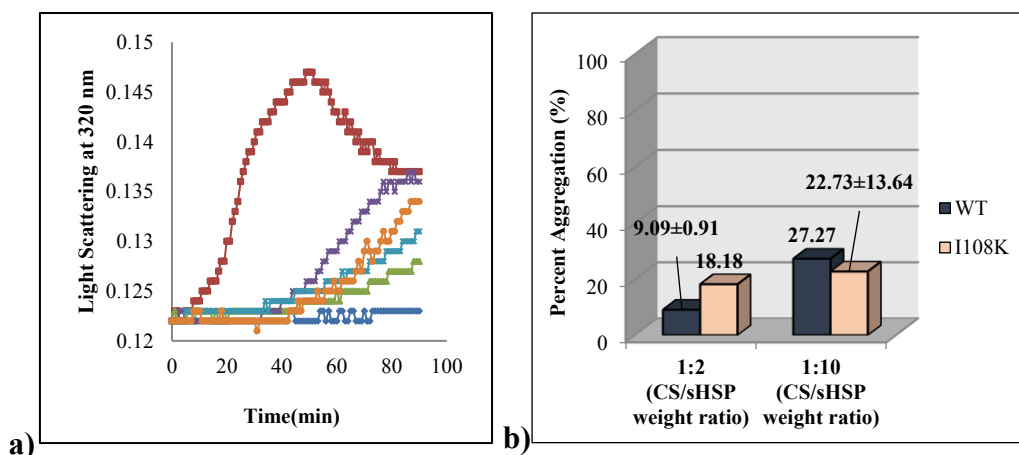
**Figure 3.55. Effect of the *Tpv* sHSP 14.3 A47D mutant on the thermal aggregation of the CS.** a) CS aggregation was monitored by measuring light scattering at 320 nm at 45°C. (—◆—) Blank, (—■—) CS alone, (—▲—) CS + WT (1:2 w/w ratio), (—×—) CS + WT (1:10 w/w ratio), (—\*—) CS + A47D (1:2 w/w ratio), (—○—) CS + A47D (1:10 w/w ratio). b) Percent aggregation when the aggregation of CS alone (Control) was taken as 100 % (100±4.55). Each experiment was repeated at least three times.

The ability of G48E and I108K single mutants to keep the CS soluble against heat induced aggregation was compared with wild type. CS aggregation assay showed that G48E decreased the CS aggregation about 73% and 64% at w/w ratio of 1/2 and 1/10 (CS/sHSP), respectively (Figure 3.56). Thermal aggregation measurements of CS at 1:2 w/w CS/sHSP ratio showed that light scattering in the presence of I108K was decreased to 18 % which was two-fold more than that of WT sHSP at that ratio.

At the 1:10 w/w ratio, CS aggregation was prevented by 77% on the addition of I108K (Figure 3.57). Degree of protection at this ratio was found to be slightly higher (4.6%) than WT.



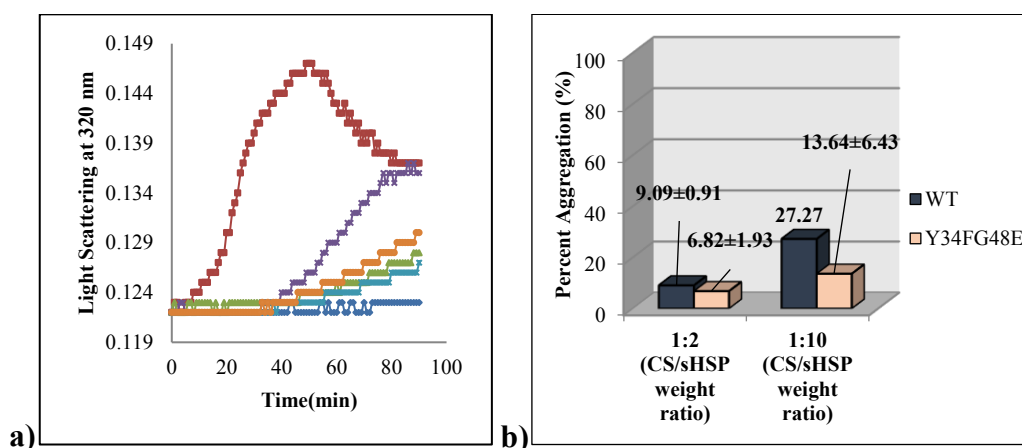
**Figure 3.56. Effect of the *Tpv* sHSP 14.3 G48E mutant on the thermal aggregation of the CS.** a) CS aggregation was monitored by measuring light scattering at 320 nm at 45°C. (◆) Blank, (■) CS alone, (▲) CS + WT (1:2 w/w ratio), (✕) CS + WT (1:10 w/w ratio), (✱) CS + G48E (1:2 w/w ratio), (●) CS + G48E (1:10 w/w ratio). b) Percent aggregation when the aggregation of CS alone (Control) was taken as 100 % (100±4.55). Each experiment was repeated at least three times.



**Figure 3.57. Effect of the *Tpv* sHSP 14.3 I108K mutant on the thermal aggregation of the CS** a) CS aggregation was monitored by measuring light scattering at 320 nm at 45°C. (◆) Blank, (■) CS alone, (▲) CS + WT (1:2 w/w ratio), (✕) CS + WT (1:10 w/w ratio), (✱) CS + I108K (1:2 w/w ratio), (●) CS + I108K (1:10 w/w ratio). b) Percent aggregation when the aggregation of CS alone (Control) was taken as 100 % (100±4.55). Each experiment was repeated at least three times.

### 3.9.4 Effect of the *Tpv* sHSP14.3 Double mutants on CS Thermal Aggregation

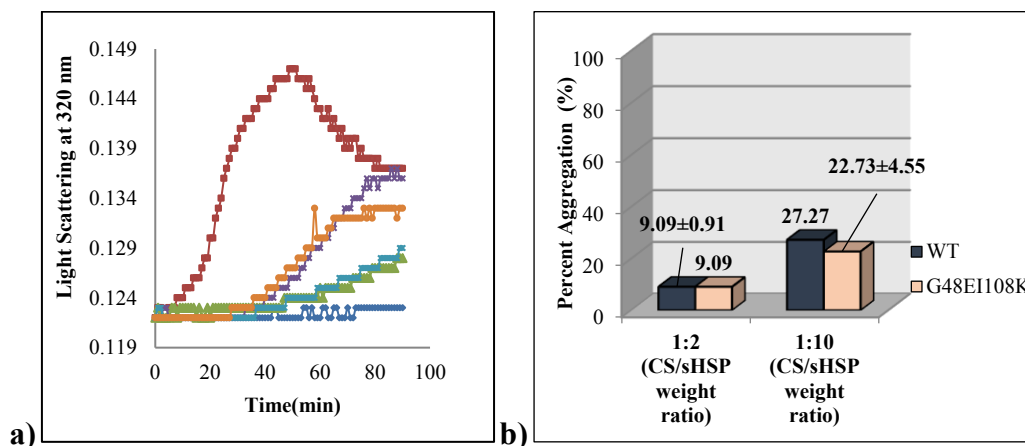
The ability of Y34FG48E double mutant protein to prevent thermal aggregation of CS was better than that of WT protein and single mutants at both 1:2 CS/sHSP and 1:10 of CS/sHSP w/w ratios. Significant suppression of aggregation to ~ 7% and 13.7 % were observed in the presence of the mutant protein at 1:2 and 1:10 of CS/sHSP w/w ratios, respectively (Figure 3.58). At both ratios of CS/sHSP, efficacy of the double mutant on preventing thermal aggregation of CS was better than related single mutants.



**Figure 3.58. Effect of the *Tpv* sHSP 14.3 Y34FG48E double mutant on the thermal aggregation of the CS.** a) CS aggregation was monitored by measuring light scattering at 320 nm at 45°C. (—◆—) Blank, (—■—) CS alone, (—▲—) CS + WT (1:2 w/w ratio), (—×—) CS + WT (1:10 w/w ratio), (—\*—) CS + Y34FG48E (1:2 w/w ratio), (—○—) CS + Y34FG48E (1:10 w/w ratio). b) Percent aggregation when the aggregation of CS alone (Control) was taken as 100 % (100±4.55). Each experiment was repeated at least three times.

On the other hand, in other double mutant, charged residues were introduced in the place of hydrophobic ones into the positions which are far apart from each other (G48E I108K). Effect of the mutation on the functional properties of *Tpv* sHSP14.3 was found to be different than single mutants, but more similar to WT. At the CS/sHSP w/w ratio of 1:2, aggregation of CS was reduced to 9 % like WT sHSP. Increasing the concentration of G48E I108K by 5-fold did not increase its protection efficiency in a concentration dependent manner, but still this mutant showed 4.5% higher chaperone activity than WT (Figure 3.59). It is noticeable that I108 to K

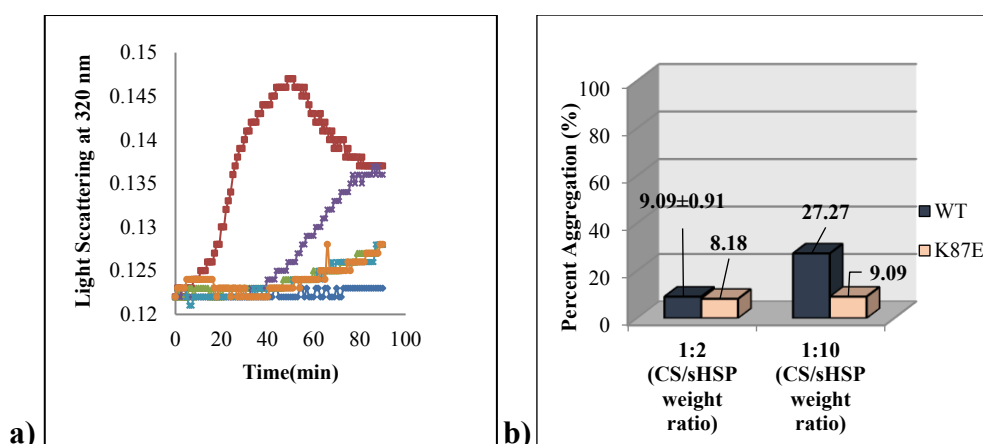
mutation suppresses the negative effect of the G48 to E mutation so that the chaperone activity of the double mutant was quite similar to the single variant I108K and WT.



**Figure 3.59. Effect of the *Tpv* sHSP 14.3 G48EI108K double mutant on the thermal aggregation of the CS. a)** CS aggregation was monitored by measuring light scattering at 320 nm at 45°C. (—◆—) Blank, (—■—) CS alone, (—▲—) CS + WT (1:2 w/w ratio), (—×—) CS + WT (1:10 w/w ratio), (—\*—) CS + G48EI108K (1:2 w/w ratio), (—○—) CS + G48EI108K (1:10 w/w ratio). **b)** Percent aggregation when the aggregation of CS alone (Control) was taken as 100 % (100±4.55). Each experiment was repeated at least three times.

### 3.9.5 Effect of the *Tpv* sHSP14.3 Mutant with Reversing of the Charge on CS Thermal Aggregation

Among the *Tpv* sHSP variants, K87E mutant has the highest chaperone activity for protection of the citrate synthase from heat aggregation. It effectively prevented the irreversible aggregation of CS at 45°C up to 8%. At both 1:2 CS/sHSP and 1:10 CS/sHSP w/w ratios, CS aggregation was prevented > 90%, which was more than WT sHSP and other mutant variants (Figure 3.60).



**Figure 3.60. Effect of the *Tpv* sHSP 14.3 K87E mutant on the thermal aggregation of the CS. a)** CS aggregation was monitored by measuring light scattering at 320 nm at 45°C. (—◆—) Blank, (—■—) CS alone, (—▲—) CS + WT (1:2 w/w ratio), (—×—) CS + WT (1:10 w/w ratio), (—\*—) CS + K87E (1:2 w/w ratio), (—○—) CS + K87E (1:10 w/w ratio). **b)** Percent aggregation when the aggregation of CS alone (Control) was taken as 100 % ( $100 \pm 4.55$ ). Each experiment was repeated at least three times.

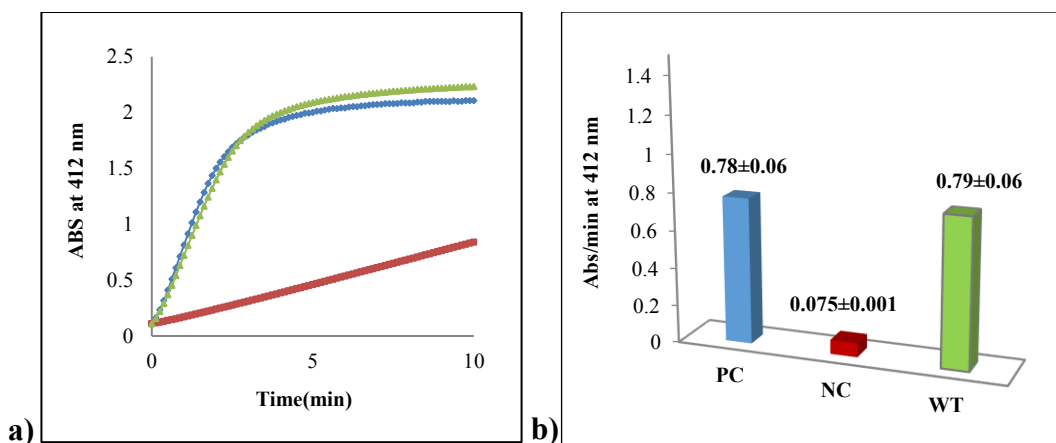
### 3.10 CS Activity Assay

Among the ACD mutants those display  $\leq 50$  % aggregation (except G107A) (*i.e.*, L33S, Y34F, E43V, G48E, K87E, G107A, I108K, Y34FG48E, G48EI108K) were tested for their ability to protect the activity of the substrate protein, CS from thermal inactivation. The activity assay was performed at 35°C, in the absence and presence of the selected sHSP variants at three different concentrations 220  $\mu\text{g}/\text{mL}$ , 110  $\mu\text{g}/\text{mL}$  and 64  $\mu\text{g}/\text{mL}$  (w/w ratios; 1:500, 1:250 and 1:147, respectively) while CS concentration was 0.44  $\mu\text{g}/\text{mL}$  as described in Materials and Methods part.

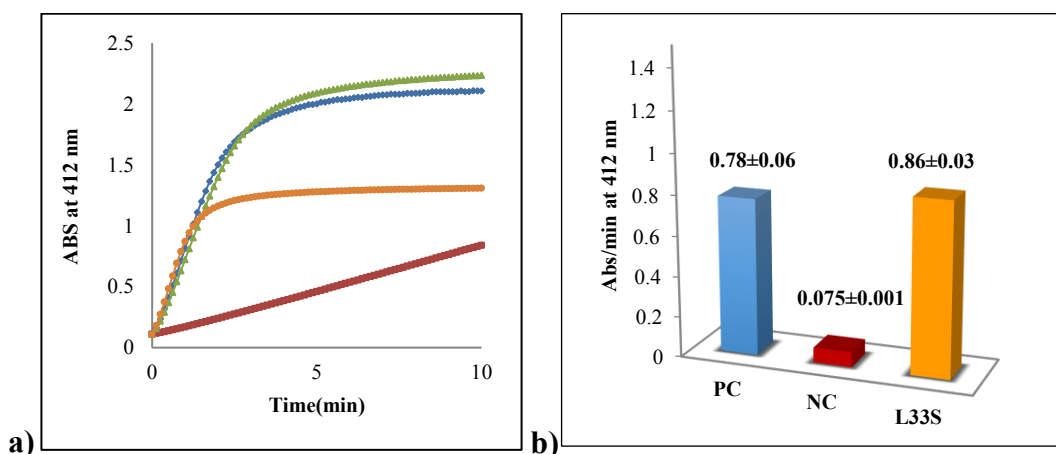
Our results showed that CS activity was dramatically reduced (11-fold) after heating at 47°C in the absence of chaperone. When excess amount of the chaperone (220  $\mu\text{g}/\text{mL}$ , 1:500 CS/sHSP w/w ratio) was used, WT sHSP was about 2% more than the chaperone activity of the positive control (Figure 3.61). Even higher increase (7 to 9.5 %) in the chaperone activity was observed with L33S, Y34F and G48E mutants with respect to WT sHSP (Figure 3.62-3.64). Chaperone activity also increased (3 %) in case of G48EI108K double mutation (Figure 3.65). However, heat protection efficiency was reduced by K87E and I108K mutations 27 % (Figure 3.66-3.67), by



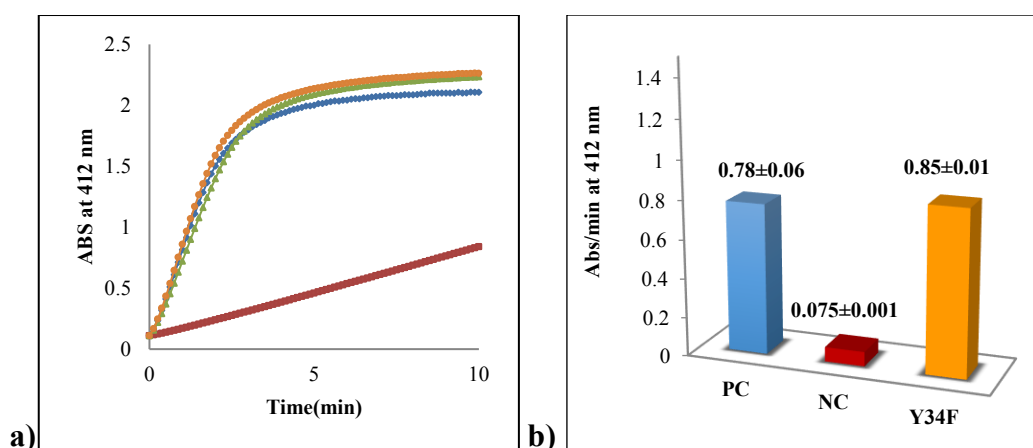
G107A mutation 22 % (Figure 3.68), by E43V mutation 5 % (Figure 3.69) and Y34FG48E double mutation about 17 % at 1:500 (w/w) substrate/chaperone ratio (Figure 3.70).



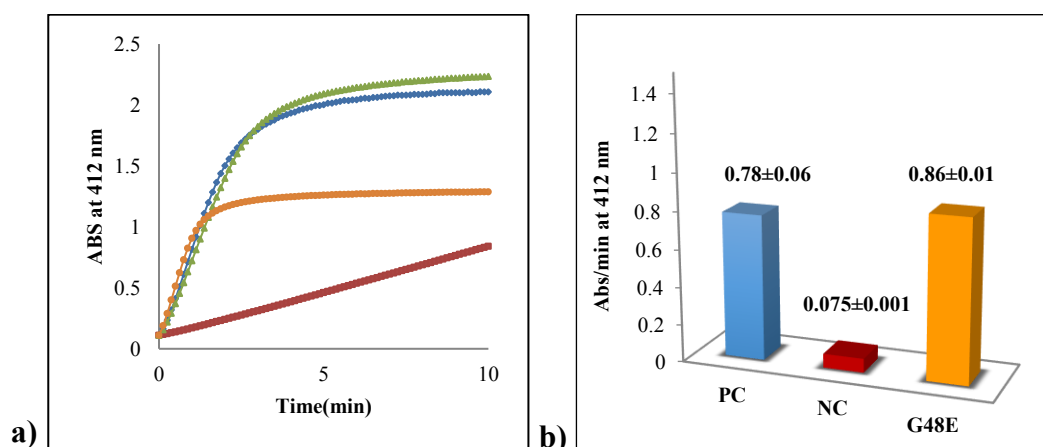
**Figure 3.61. Effect of *Tpv* sHSP 14.3 wild type on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat-treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. (—●—) PC, (—■—) NC, (—▲—) WT. Data shown are mean values, with STD values indicating the standard deviation of at least three independent experiments.



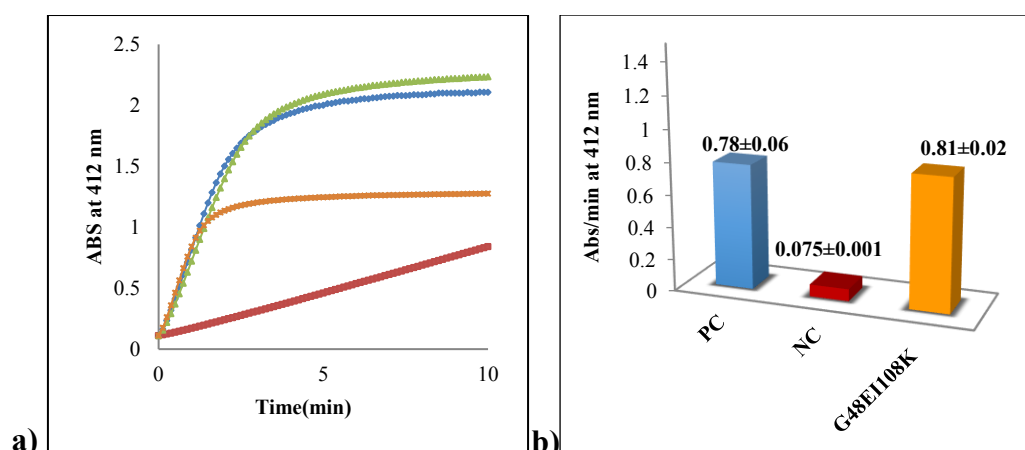
**Figure 3.62. Effect of *Tpv* sHSP 14.3 L33S mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat-treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone (—●—) PC, (—■—) NC, (—▲—) WT, (—○—) L33S. Data shown are mean values, with STD values indicating the standard deviation of at least three independent experiments.



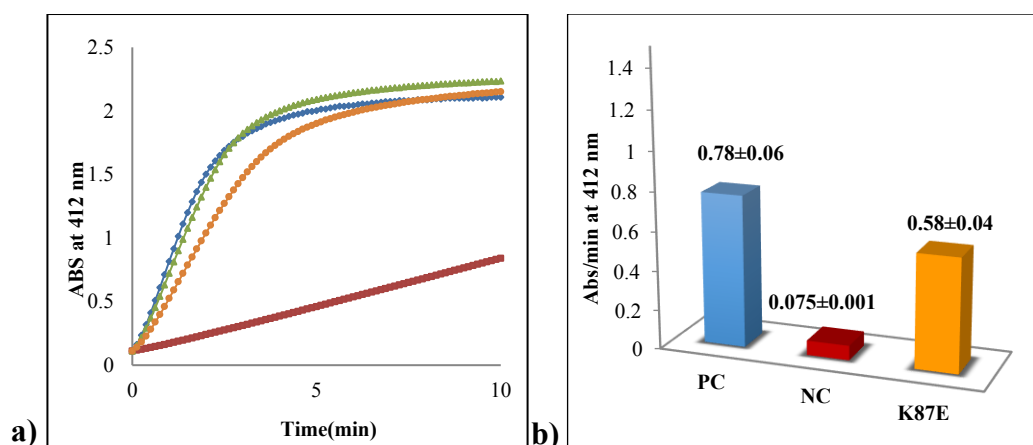
**Figure 3.63. Effect of *Tpv* sHSP 14.3 Y34F mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone. (—◆—) PC, (—■—) NC, (—▲—) WT, (—○—) Y34F. Data shown are mean values with STD values indicating the standard deviation of at least three independent experiments.



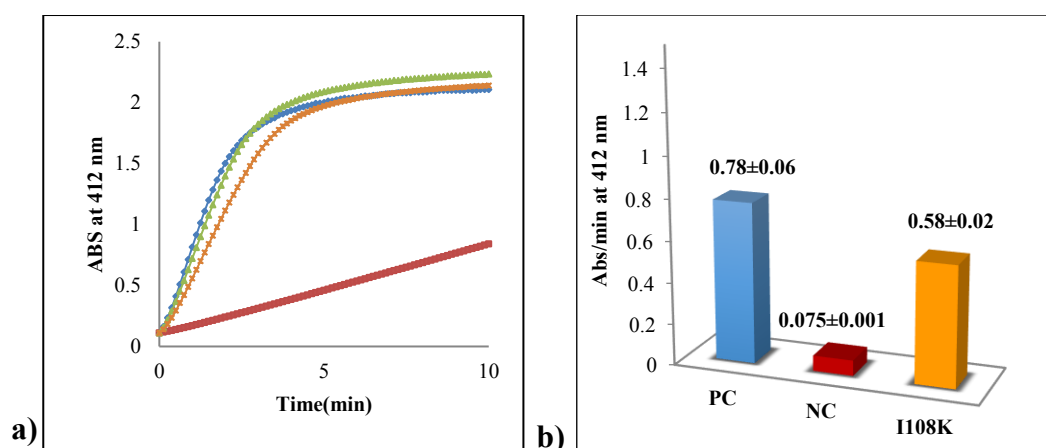
**Figure 3.64. Effect of *Tpv* sHSP 14.3 G48E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone (—◆—) PC, (—■—) NC, (—▲—) WT (—○—) G48E. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



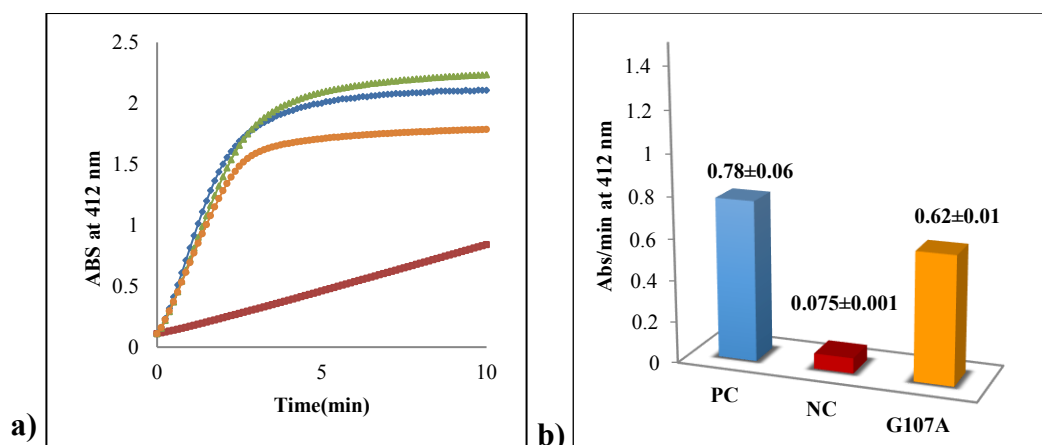
**Figure 3.65. Effect of *Tpv* sHSP 14.3 G48E1108K mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone. (—◆—) PC, (—■—) NC, (—▲—) WT (—○—) G48E1108K. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



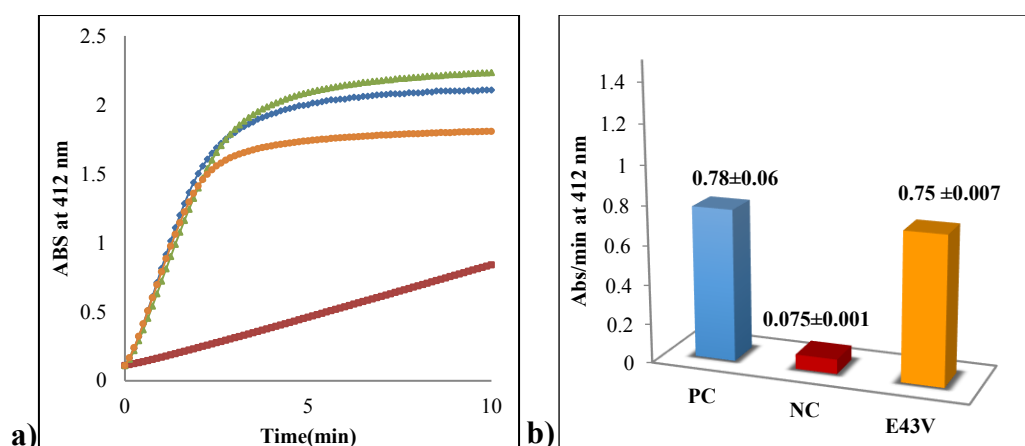
**Figure 3.66. Effect of *Tpv* sHSP 14.3 K87E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone. (—◆—) PC, (—■—) NC, (—▲—) WT, (—○—) K87E. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



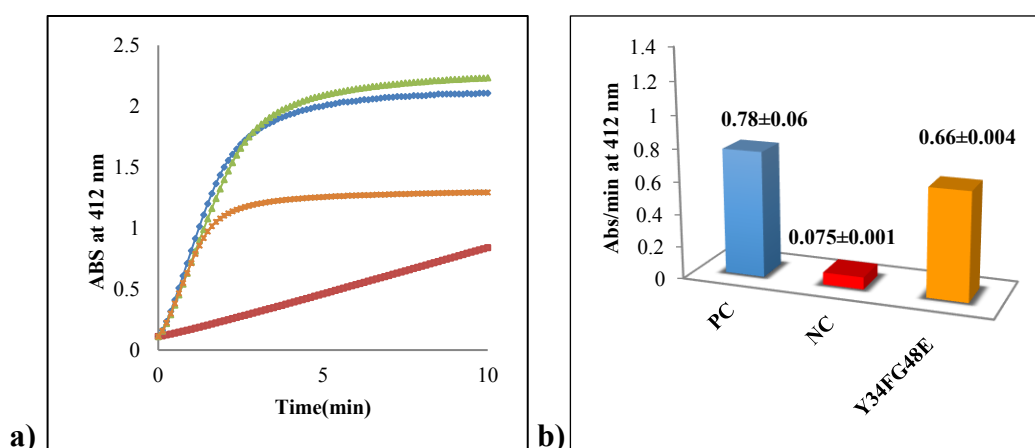
**Figure 3.67. Effect of *Tpv* sHSP 14.3 I108K mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a) The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone. (—●—) PC, (—■—) NC, (—▲—) WT, (—◆—) I108K. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



**Figure 3.68. Effect of *Tpv* sHSP 14.3 G107A mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a) The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone. (—●—) PC, (—■—) NC, (—▲—) WT, (—◆—) G107A. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



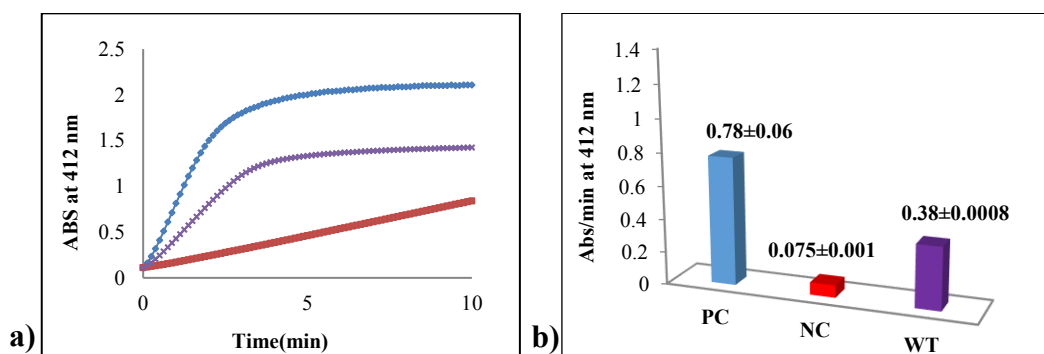
**Figure 3.69.** Effect of *Tpv* sHSP 14.3 E43V mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C. The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat- treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. (—◆—) PC, (—■—) NC, (—▲—) WT, (—●—) E43V. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



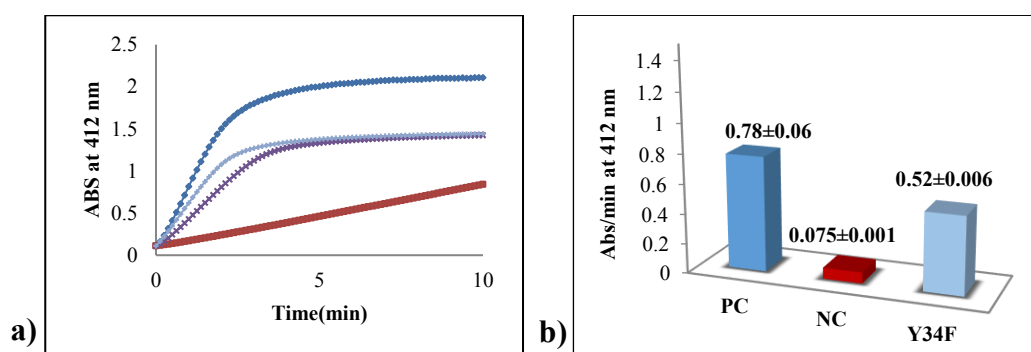
**Figure 3.70.** Effect of *Tpv* sHSP 14.3 Y34FG48E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:500 (w/w) at 47°C. The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat- treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. (—◆—) PC, (—■—) NC, (—▲—) WT, (—●—) Y34FG48E. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

When WT *Tpv* sHSP 14.3 was used (1:250 CS/sHSP w/w ratio), enzyme activity of CS reduced almost 2-folds (Figure 3.71). At this molar ratio, chaperone activity of Y34F and G48E single mutants was 1.4-fold and 1.3-fold higher than WT,

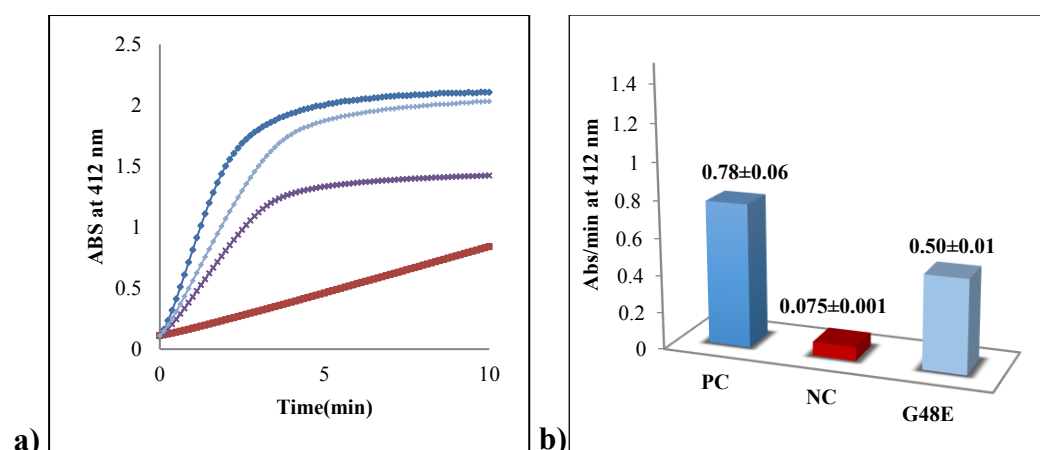
respectively (Figure 3.72-3.73). On the other hand, L33S, E43V, and G107A mutations almost did not change the protection capacity of the *Tpv* sHSP 14.3 (Figure 3.74-3.76). However, heat induced inactivation of CS was about 10% less in case of K87E and I108K mutations, than the WT sHSP (Figure 3.77-3.78). The decrease in chaperone activity was 14% and 18 % less than the WT, with Y34FG48E and G48EI108K double mutations, respectively (Figure 3.79-3.80).



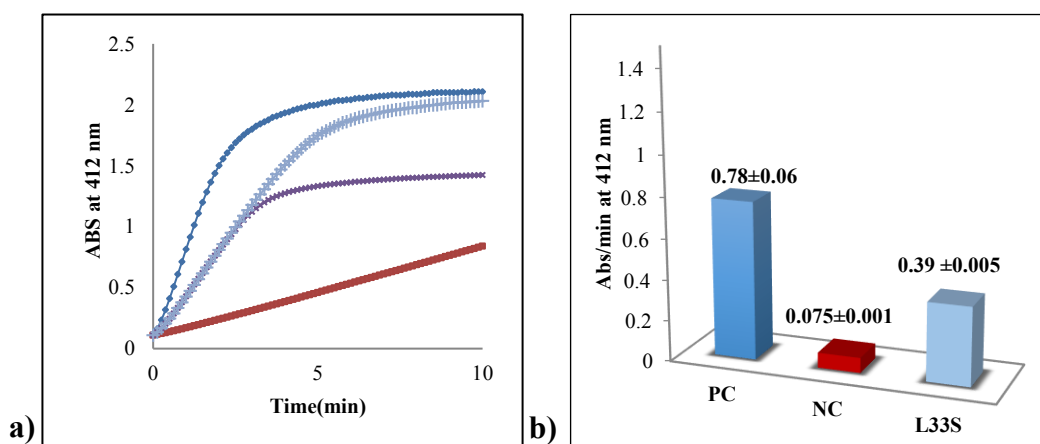
**Figure 3.71. Effect of *Tpv* sHSP 14.3 WT on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat-treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. (—♦—) PC, (—■—) NC, (—×—) WT 1:250 w/w ratio. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



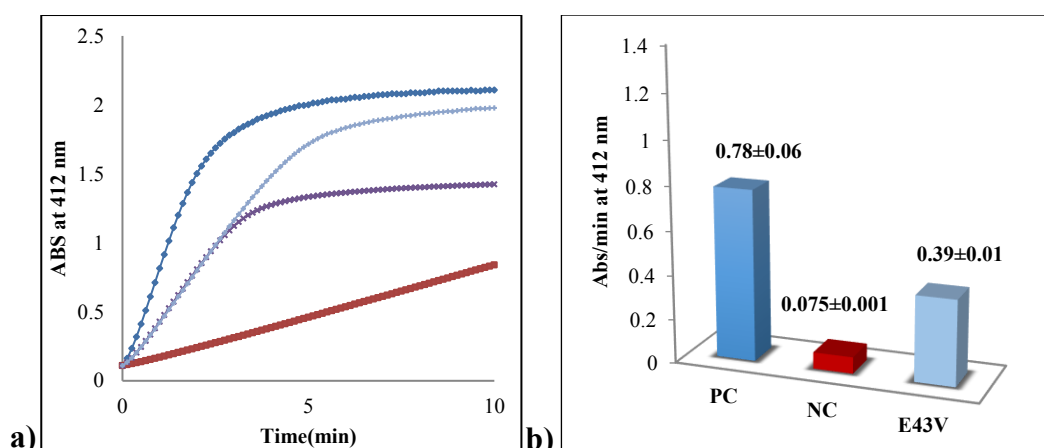
**Figure 3.72. Effect of *Tpv* sHSP 14.3 Y34F mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat-treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. (—♦—) PC, (—■—) NC, (—×—) WT, (—+—) Y34F. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



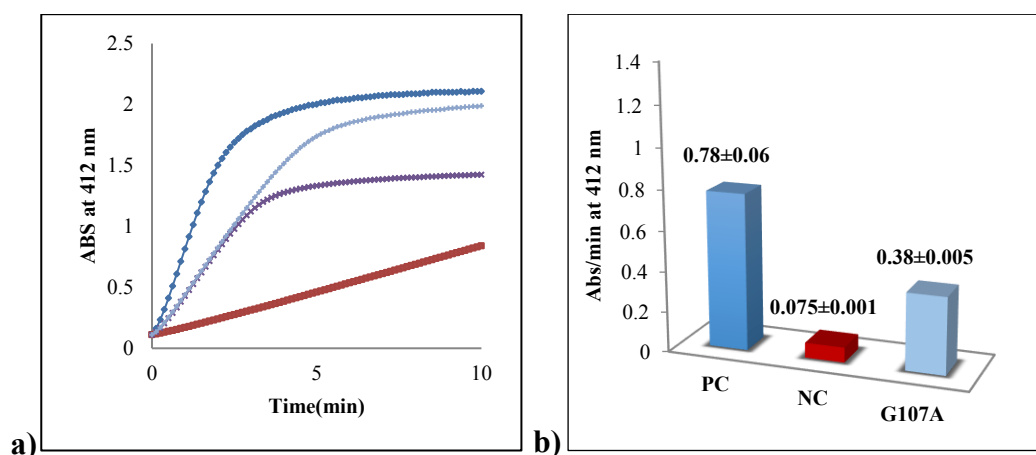
**Figure 3.73. Effect of *Tpv* sHSP 14.3 G48E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1/250 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone. (—◆—) PC, (—■—) NC, (—×—) WT, (—+—) G48E. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



**Figure 3.74. Effect of *Tpv* sHSP 14.3 L33S mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone. (—◆—) PC, (—■—) NC, (—×—) WT, (—+—) L33S. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

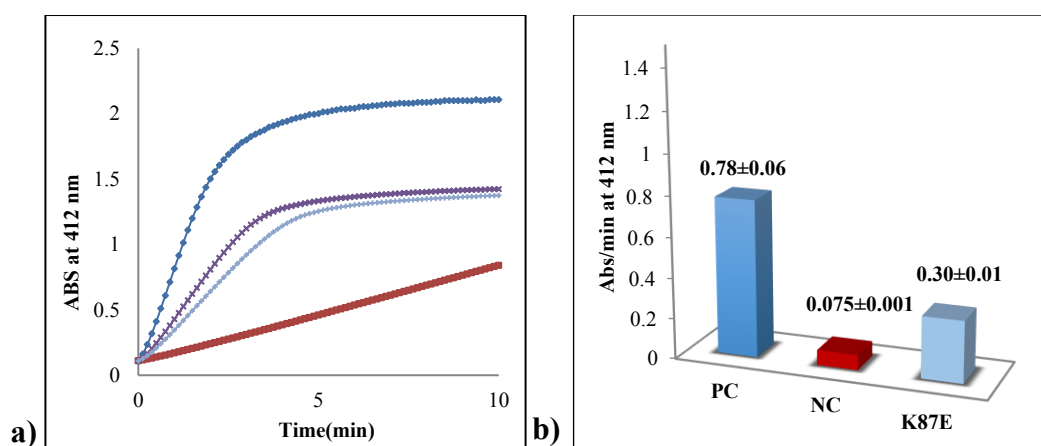


**Figure 3.75. Effect of *Tpv* sHSP 14.3 E43V mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat- treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. (—♦—) PC, (—■—) NC, (—×—) WT, (—+—) E43V. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

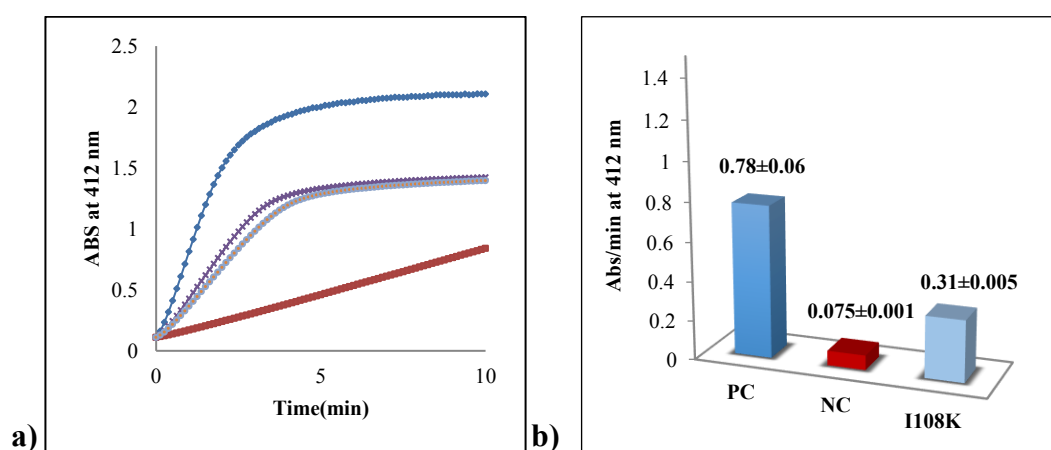


**Figure 3.76. Effect of *Tpv* sHSP 14.3 G107A mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat- treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. (—♦—) PC, (—■—) NC, (—×—) WT, (—+—) G107A. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

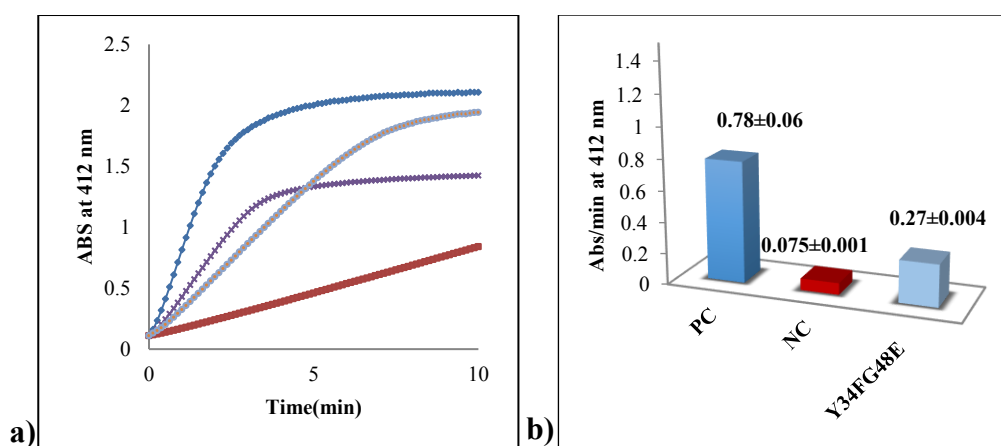




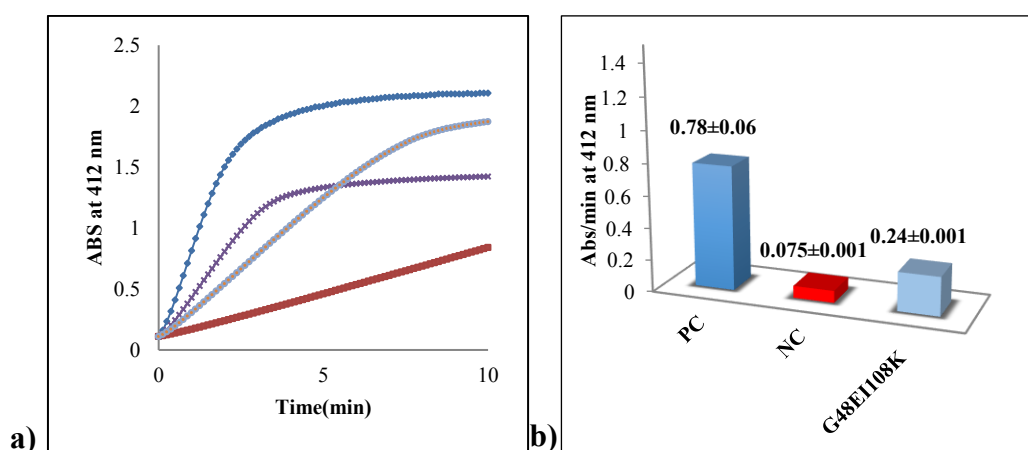
**Figure 3.77. Effect of *Tpv* sHSP 14.3 K87E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat- treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. (—◆—) PC, (—■—) NC, (—×—) WT, (—+—) K87E. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



**Figure 3.78. Effect of *Tpv* sHSP 14.3 I108K mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat- treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. (—◆—) PC, (—■—) NC, (—×—) WT, (—+—) I108K. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



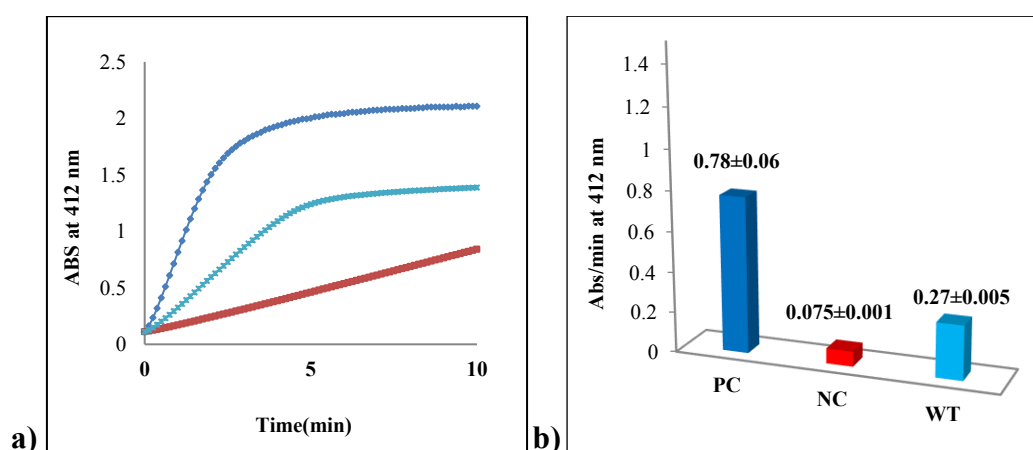
**Figure 3.79. Effect of *Tpv* sHSP 14.3 Y34FG48E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone. (—♦—) PC, (—■—) NC, (—×—) WT, (—+—) Y34FG48E. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



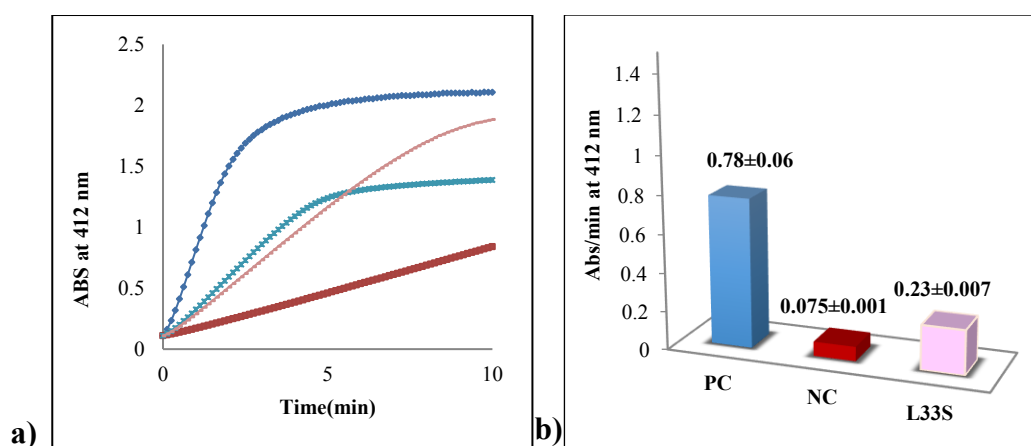
**Figure 3.80. Effect of *Tpv* sHSP 14.3 G48E1108K mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:250 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone. (—♦—) PC, (—■—) NC, (—×—) WT, (—+—) G48E1108K. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

The mutants with same or higher chaperone activity as compared to WT (L33S, Y34F, E43V, G48E and G107A) were included in an activity assay where CS: Chaperone the w/w ratio is reduced 1:147. The WT sHSP CS activity protection

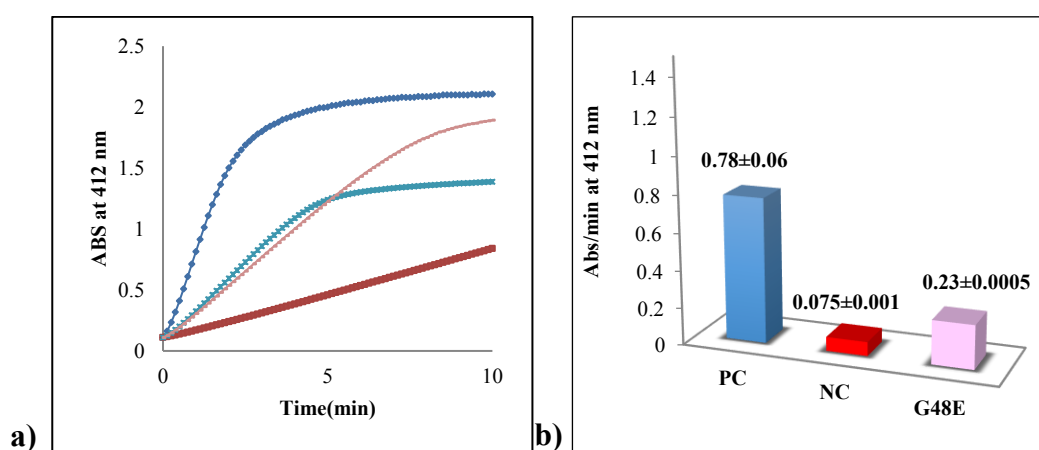
efficiency at this ratio was 15% less than 1: 250 CS/ sHSP w/w ratio (Figure 3.81). Protection of the CS activity by L33S, G48E mutants was 5% less than the WT, while E43V and G107A mutations reduced the chaperone activity about 10% as compared to WT (Figure 3.82-3.85). Y34F mutation resulted in only 1.6% decrease in the CS activity (Figure 3.86). These results indicated that excess amount of WT (220  $\mu\text{g}/\text{mL}$ ) and most of the mutant proteins are required to restore CS activity comparable to PC. At 1:250 and 1:147 CS/sHSP (w/w) ratio, Y34F and G48E showed highest chaperone activity among seven *Tpv* sHSP 14.3 ACD mutants. The L33S, E43V and G107A mutants displayed almost same efficiency as WT sHSP for protection of the CS activity at high temperature.



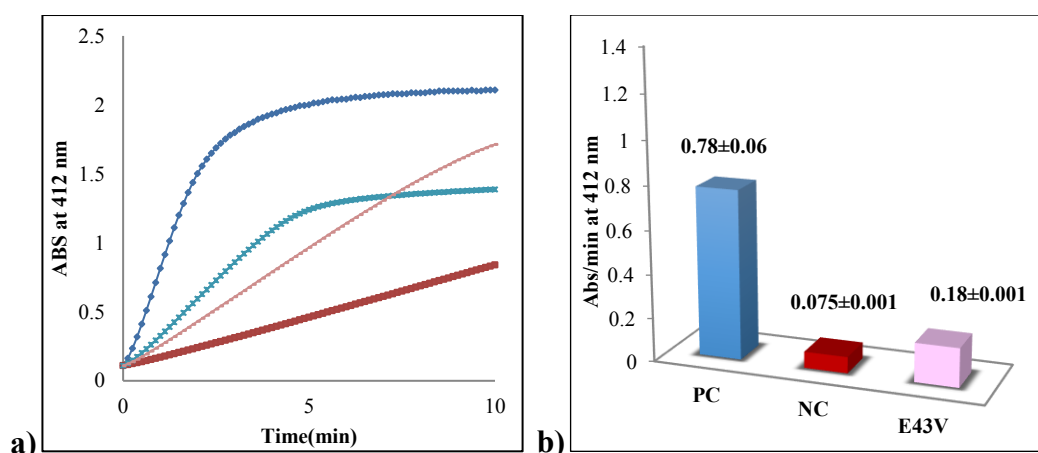
**Figure 3.81. Effect of *Tpv* sHSP 14.3 WT on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:147 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat-treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. (—●—) PC, (—■—) NC, (—■—) WT. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



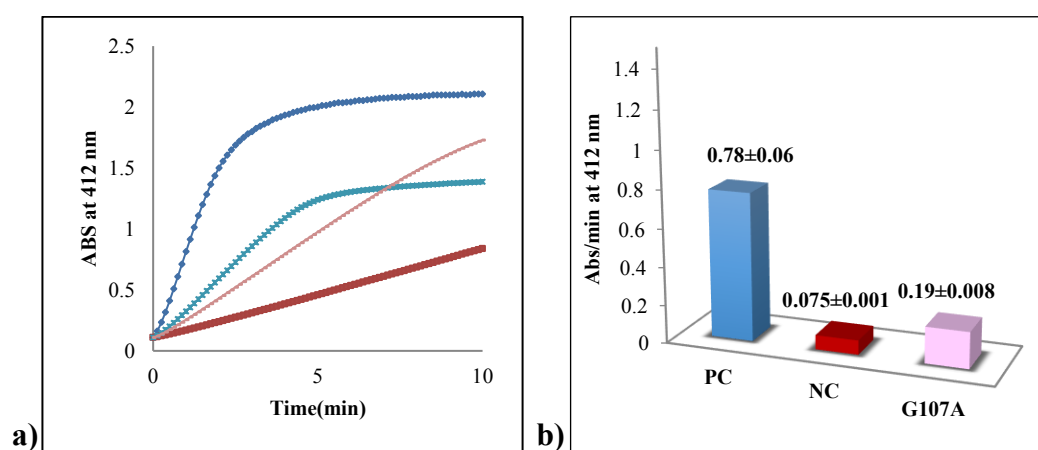
**Figure 3.82. Effect of *Tpv* sHSP 14.3 L33S mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:147 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone. (—◆—) PC, (—■—) NC, (—■—) WT, (—■—) L33S. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



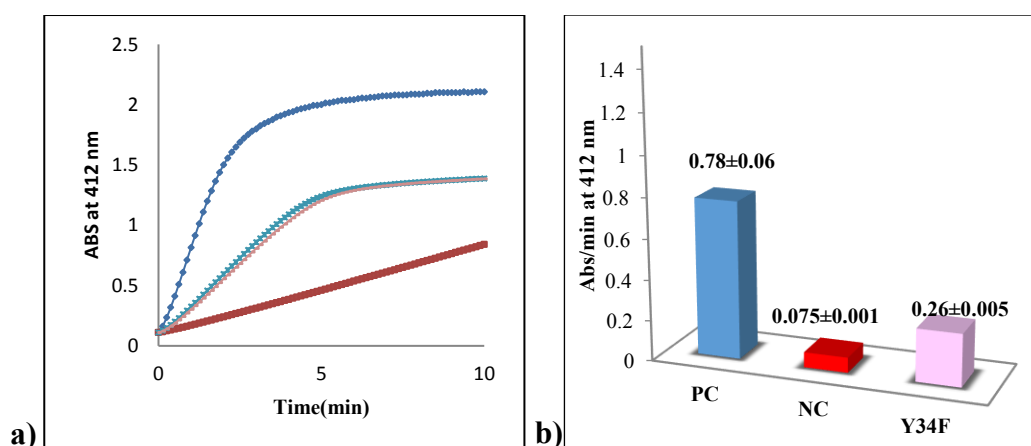
**Figure 3.83. Effect of *Tpv* sHSP 14.3 G48E mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:147 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone. (—◆—) PC, (—■—) NC, (—■—) WT, (—■—) G48E. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



**Figure 3.84. Effect of *Tpv* sHSP 14.3 E43V mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:147 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat- treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. (—●—) PC, (—■—) NC, (—■—) WT, (—●—) E43V. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



**Figure 3.85. Effect of *Tpv* sHSP 14.3 G107A mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:147 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). PC: Positive Control, activity measured before heat- treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. (—●—) PC, (—■—) NC, (—■—) WT, (—●—) G107A. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



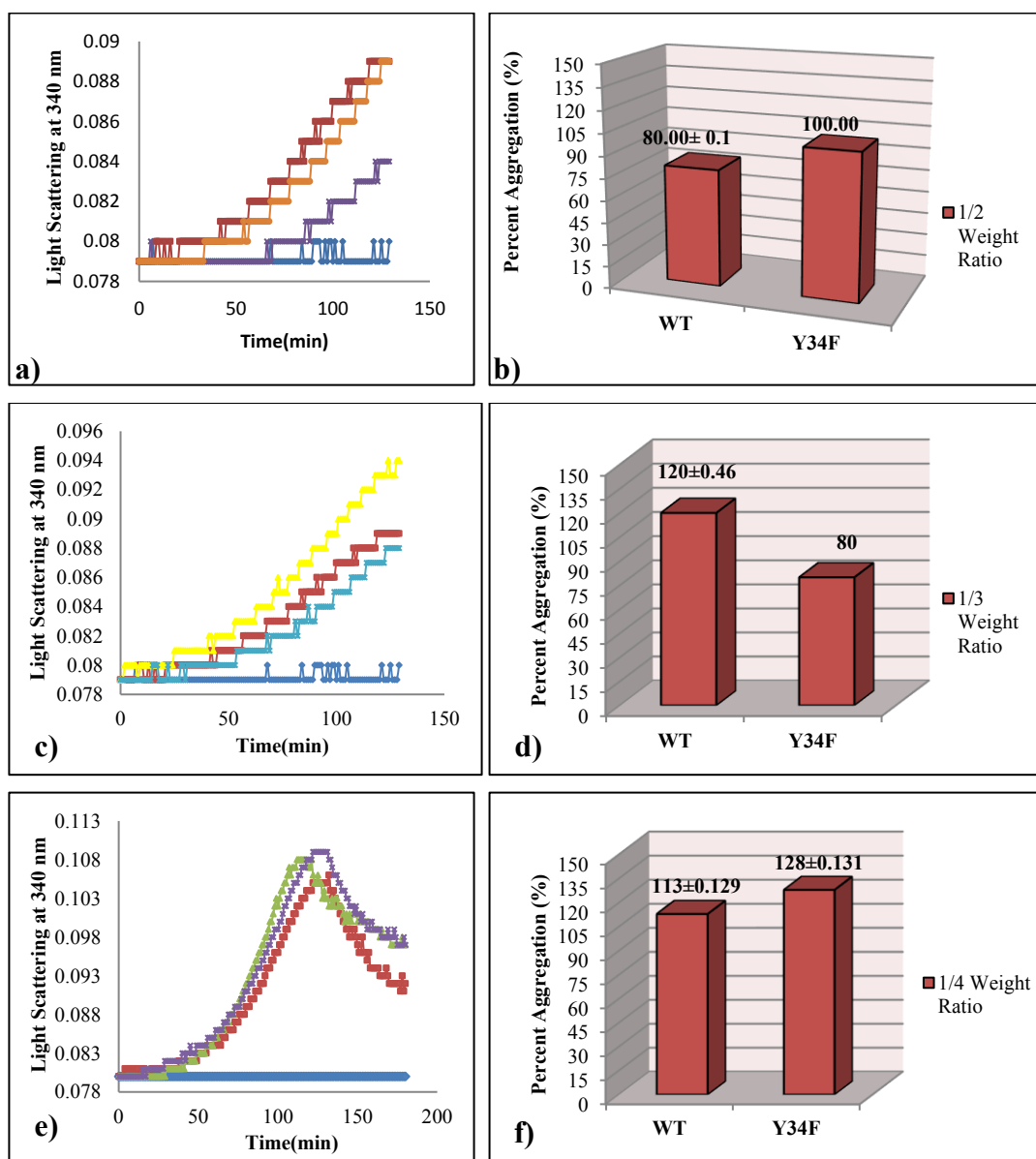
**Figure 3.86. Effect of *Tpv* sHSP 14.3 Y34F mutant on the prevention of CS from thermal inactivation at a CS to sHSP ratio of 1:147 (w/w) at 47°C.** The remaining activity was monitored by continuous measuring of the absorbance at 412 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (b). **PC:** Positive Control, activity measured before heat- treatment. **NC:** Negative Control, remaining activity after heat treatment in the absence of chaperone. (—◆—) PC, (—■—) NC, (—□—) WT, (—□—) Y34F. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

### 3.11 Yeast Alcohol Dehydrogenase Aggregation Assay

The effect of targeted mutations on the chaperone action of the *Tpv* sHSP 14.3 were also studied by anti-aggregation assay using the second model substrate protein, yeast alcohol dehydrogenase (yADH). The selected mutant sHSPs (*i.e.*, L33S, Y34F, E43V, G48E, K87E, I108K, Y34FG48E and G48EI108K) were included into the aggregation assay experiments based on their good heat stability, purification quality, and protection ability of the mutants against thermal aggregation of the CS. These selected mutants had appreciable chaperone effect on the prevention of CS aggregation, except G48E mutant sHSP. Thermal aggregation assay was performed by monitoring light scattering at 340 nm, as described in the Materials and Methods. The experiments were performed at three different substrate / chaperone w/w ratio: 1:2, 1:3 and 1:4.

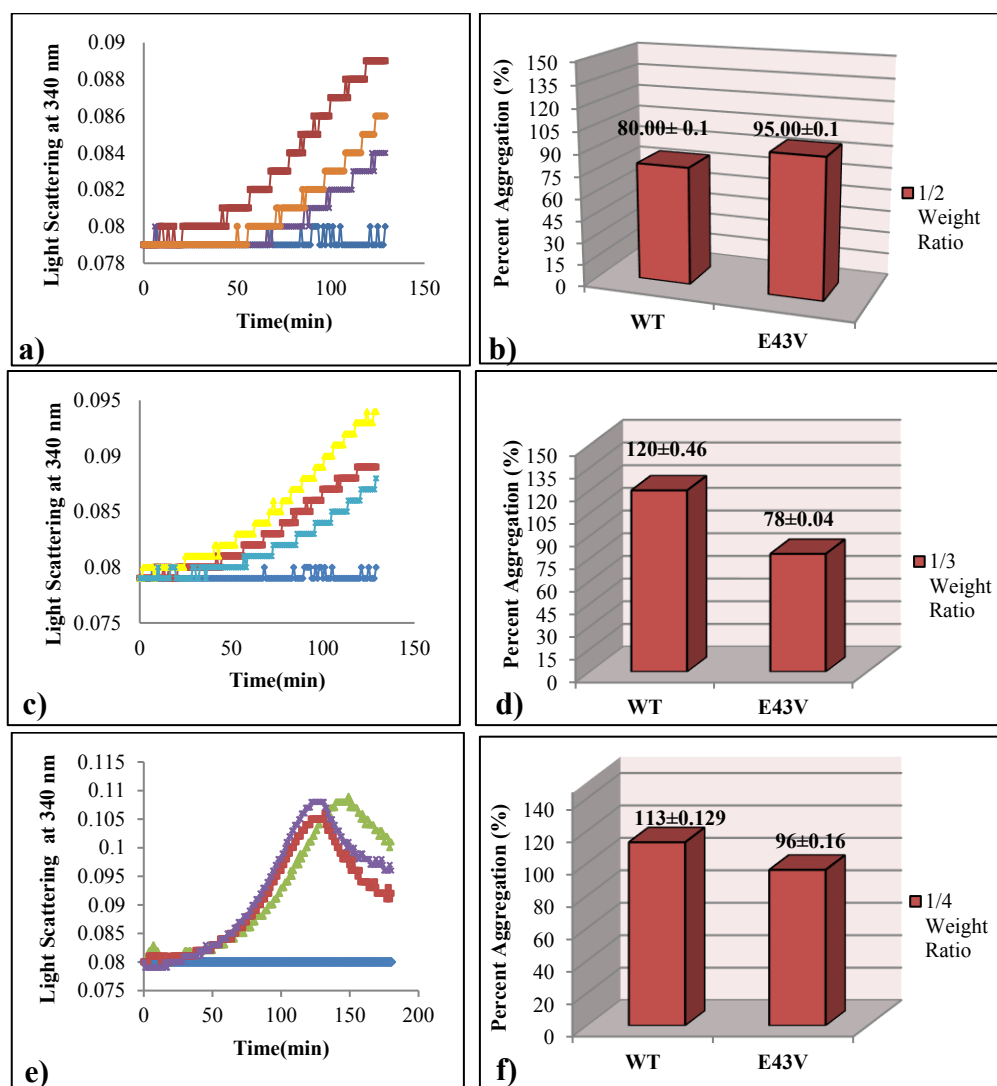
Thermal aggregation of the ADH was enhanced by Y34F mutation as compared to control and WT, at excess concentration (1:4 ADH/sHSP w/w ratio) while the same mutations decreased aggregation 20% at 1:3 ADH/sHSP w/w ratio. At 1:2 w/w

ADH:sHSP ratio, possibly due to insufficient amount of the Y34F mutant variant ADH aggregation was higher than WT sHSP (Figure 3.87).



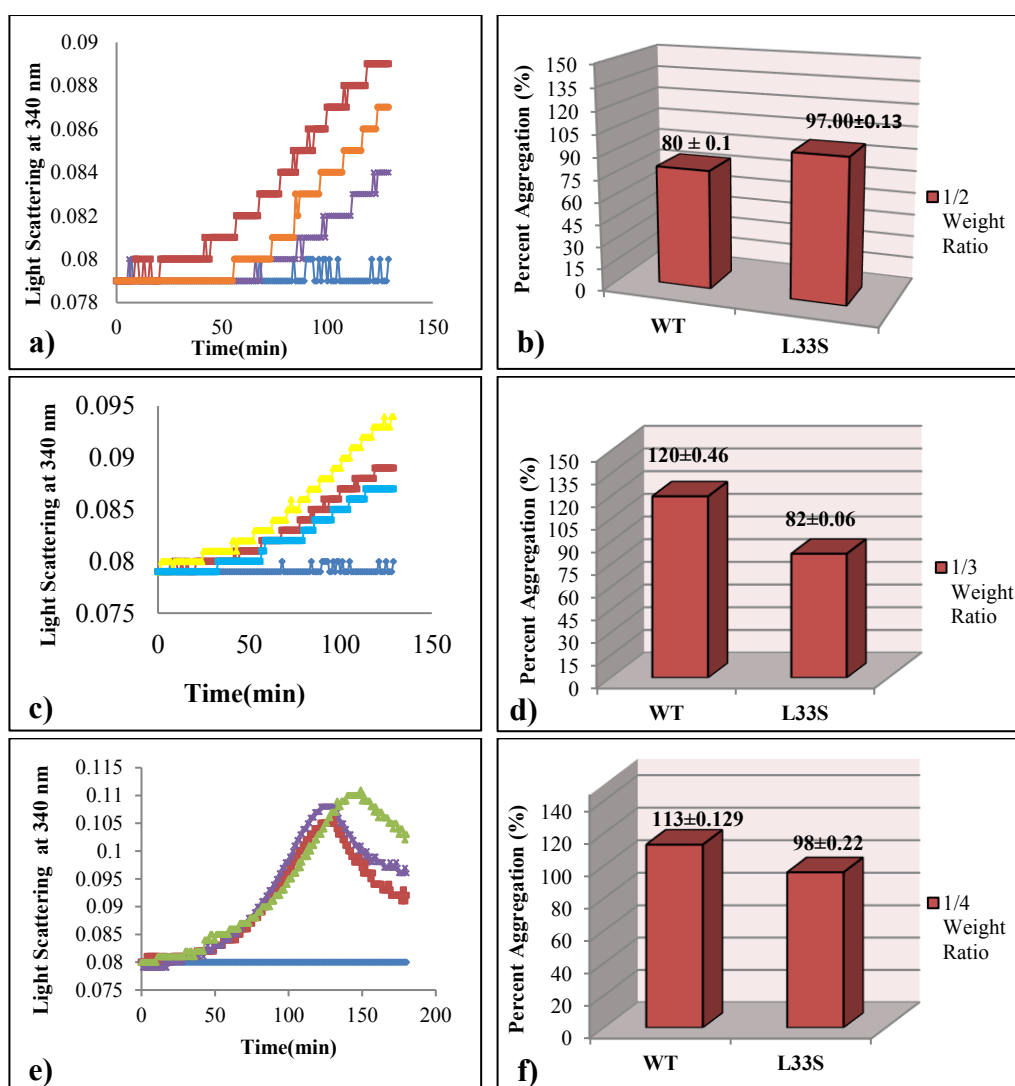
**Figure 3.87. ADH aggregation assay with Y34F mutant sHSP.** Heat induced aggregation of ADH in the absence or presence of the WT *Tpv* sHSP 14.3 and Y34F proteins were monitored by measuring light scattering at 340 nm. The tested w/w ADH/sHSP ratio were 1/2(a), 1/3(c) and 1/4(e). (—♦—) Blank, (—■—) ADH alone, (—×—) ADH + WT 1:2 w/w ratio, (—○—) ADH + Y34F 1:2 w/w ratio, (—★—) ADH + WT 1:3 w/w ratio, (—□—) ADH + Y34F 1:3 w/w ratio, (—▲—) ADH + Y34F 1:4 w/w ratio, (—×—) ADH + WT 1:4 w/w ratio. Interpretation of the results were done based on percent aggregation where the aggregation of ADH alone was taken as 100 % ( $100 \pm 0.102$ ) (b-d-f). Each experiment was repeated at least three times.

The single mutant E43V, which lost charge by introduction of hydrophobicity, performed better protection with respect to the control and WT at high concentrations (1:3 and 1:4 ADH/sHSP w/w ratio), while it was not as efficient as the WT *Tpv* sHSP14.3 at low concentration (1:2 ADH/sHSP w/w ratio) (Figure 3.88). Similar to the E43V mutant, L33S also showed better protection at 1:3 ADH/sHSP w/w ratio (Figure 3.89)



**Figure 3.88. ADH aggregation assay with E43V mutant sHSP.** Heat induced aggregation of ADH in the absence or presence of the WT *Tpv* sHSP 14.3 and E43V proteins were monitored by measuring light scattering at 340 nm. The tested w/w ADH/sHSP ratio were 1/2(a), 1/3(c) and 1/4(e). (—●—) Blank, (—■—) ADH alone, (—\*—) ADH + WT 1:2 w/w ratio, (—○—) ADH + E43V 1:2 w/w ratio, (—▲—) ADH + WT 1:3 w/w ratio, (—□—) ADH + E43V 1:3 w/w ratio, (—△—) ADH + E43V 1:4 w/w ratio, (—◇—) ADH + WT 1:4 w/w ratio. Interpretation of the results were done based on percent aggregation where the aggregation of ADH alone was taken as 100 % ( $100 \pm 0.102$ ) (b-d-f). Each experiment was repeated at least three times.

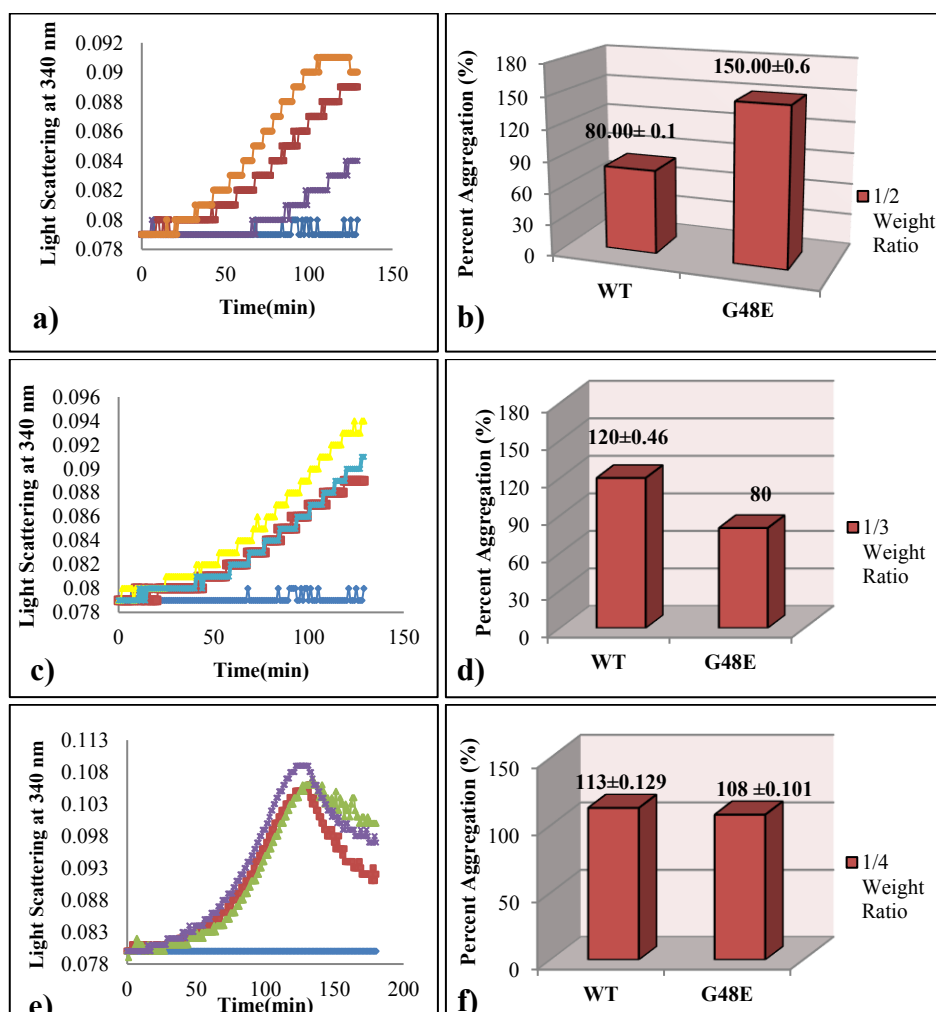




**Figure 3.89. ADH aggregation assay with L33S mutant sHSP.** Heat induced aggregation of ADH in the absence or presence of the WT *Tpv* sHSP 14.3 and L33S proteins were monitored by measuring light scattering at 340 nm. The tested w/w ADH/sHSP ratio were 1/2(a), 1/3(c) and 1/4(e). (—●—) Blank, (—■—) ADH alone, (—×—) ADH + WT 1:2 w/w ratio, (—○—) ADH + L33S 1:2 w/w ratio, (—▲—) ADH + WT 1:3 w/w ratio, (—□—) ADH + L33S 1:3 w/w ratio, (—△—) ADH + L33S 1:4 w/w ratio, (—×—) ADH + WT 1:4 w/w ratio. Interpretation of the results were done based on percent aggregation where the aggregation of ADH alone was taken as 100 % (100±0.102) (b-d-f). Each experiment was repeated at least three times.

Thermal aggregation measurements of G48E showed that this mutant did not provide sufficient protection of the ADH against heat induced aggregation at low chaperone concentration (*i.e.*, at 1:2 ADH/sHSP w/w ratio). On the other hand, at 1:3 ADH/sHSP w/w ratio, G48E mutant sHSP reduced the substrate protein aggregation

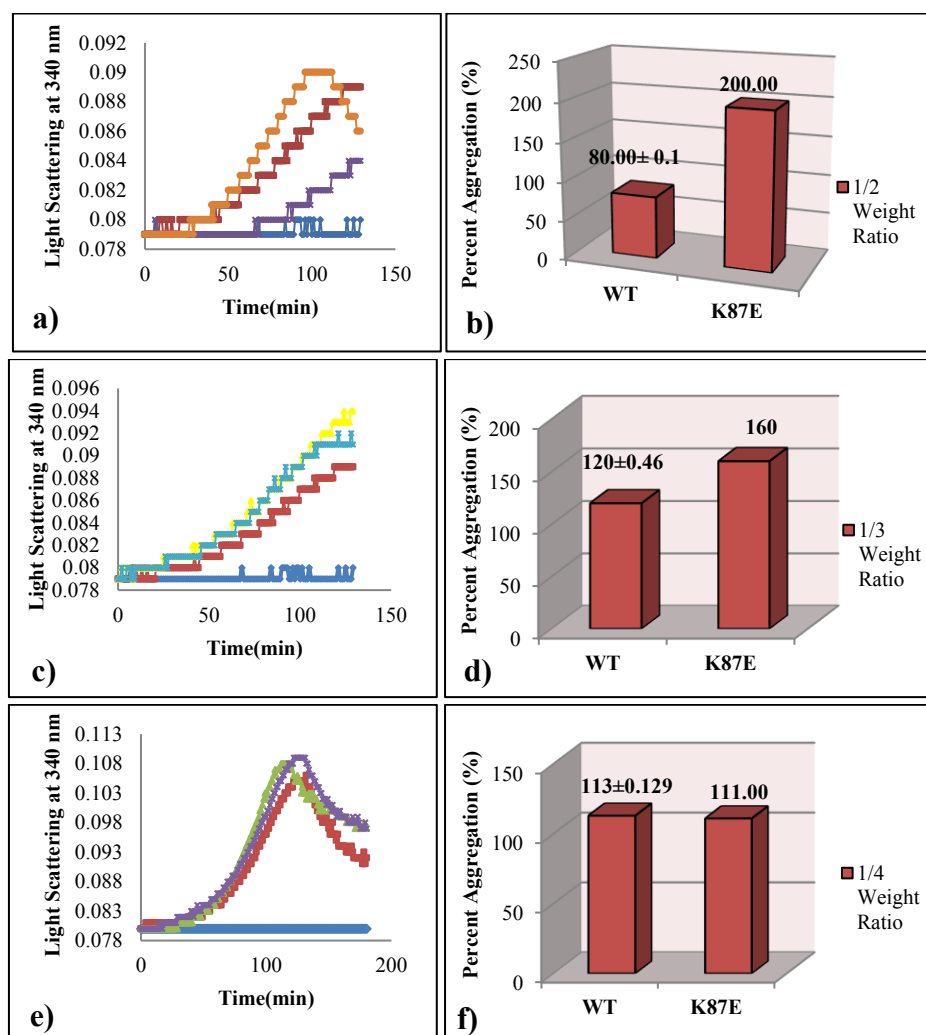
20% as compared to control. However, ADH aggregation was almost same as the control at 1:4 ADH/sHSP w/w ratio (Figure 3.90).



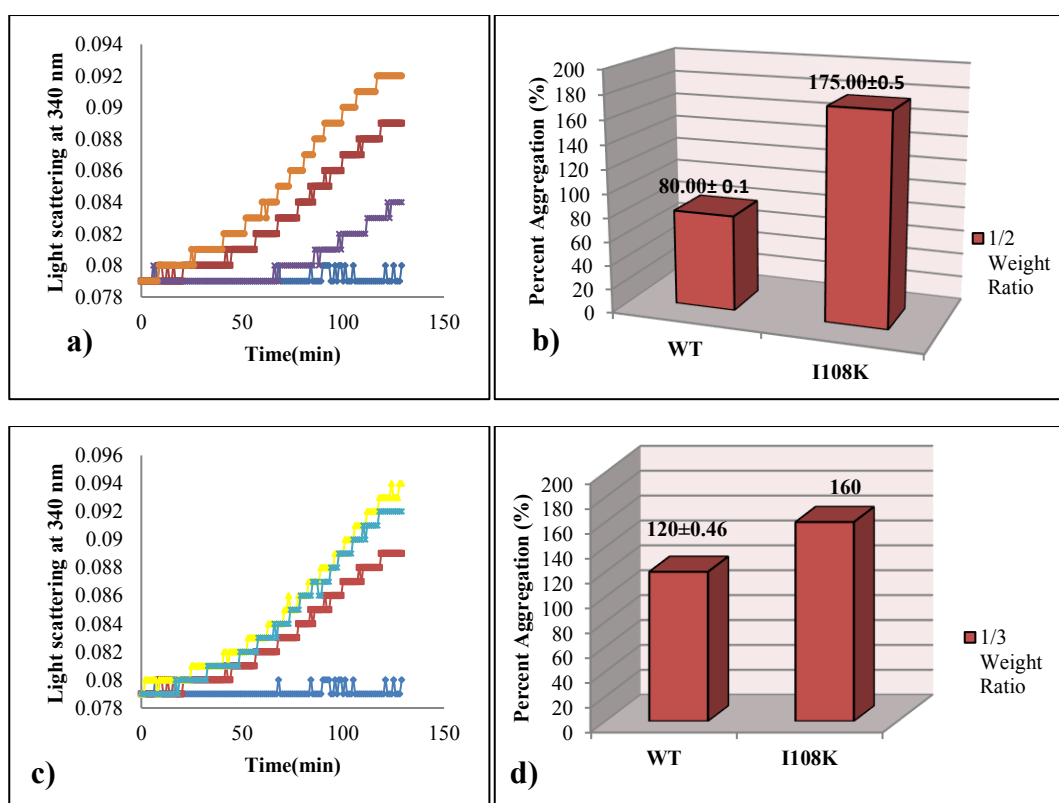
**Figure 3.90. ADH aggregation assay with G48E mutant sHSP.** Heat induced aggregation of ADH in the absence or presence of the WT *Tpv* sHSP 14.3 and G48E proteins were monitored by measuring light scattering at 340 nm. The tested w/w ADH/sHSP ratio were 1/2(a), 1/3(c) and 1/4(e). (—●—) Blank, (—■—) ADH alone, (—\*—) ADH + WT 1:2 w/w ratio, (—○—) ADH + G48E 1:2 w/w ratio, (—▲—) ADH + WT 1:3 w/w ratio, (—□—) ADH + G48E 1:3 w/w ratio, (—△—) ADH + G48E 1:4 w/w ratio, (—×—) ADH + WT 1:4 w/w ratio. Interpretation of the results were done based on percent aggregation where the aggregation of ADH alone was taken as 100 % (100±0.102) (b-d-f). Each experiment was repeated at least three times.

On the other hand, K87E and I108K mutations increased heat-induced aggregation of the ADH as compared to the control and WT at 1:2 and 1:3 ADH: sHSP w/w ratios (Figure 3.91-3.92). In case of the K87E mutation, ADH aggregation was not different than WT at 1:4 ADH: sHSP w/w ratio. Since the mutant sHSP already

increased the light scattering above 150 % at 1:2 and 1:3 ADH/sHSP w/w ratios, 1:4 ADH: sHSP w/w ratio was not studied for this mutant (Figure 3.92).



**Figure 3.91. ADH aggregation assay with K87E mutant sHSP.** Heat induced aggregation of ADH in the absence or presence of the WT *Tpv* sHSP 14.3 and K87E proteins were monitored by measuring light scattering at 340 nm. The tested w/w ADH/sHSP ratio were 1/2(a), 1/3(c) and 1/4(e). (—●—) Blank, (—■—) ADH alone, (—×—) ADH + WT 1:2 w/w ratio, (—○—) ADH + K87E 1:2 w/w ratio, (—▲—) ADH + WT 1:3 w/w ratio, (—□—) ADH + K87E 1:3 w/w ratio, (—△—) ADH + K87E 1:4 w/w ratio, (—×—) ADH + WT 1:4 w/w ratio. Interpretation of the results were done based on percent aggregation where the aggregation of ADH alone was taken as 100 % (100 ± 0.102) (b-d-f). Each experiment was repeated at least three times.

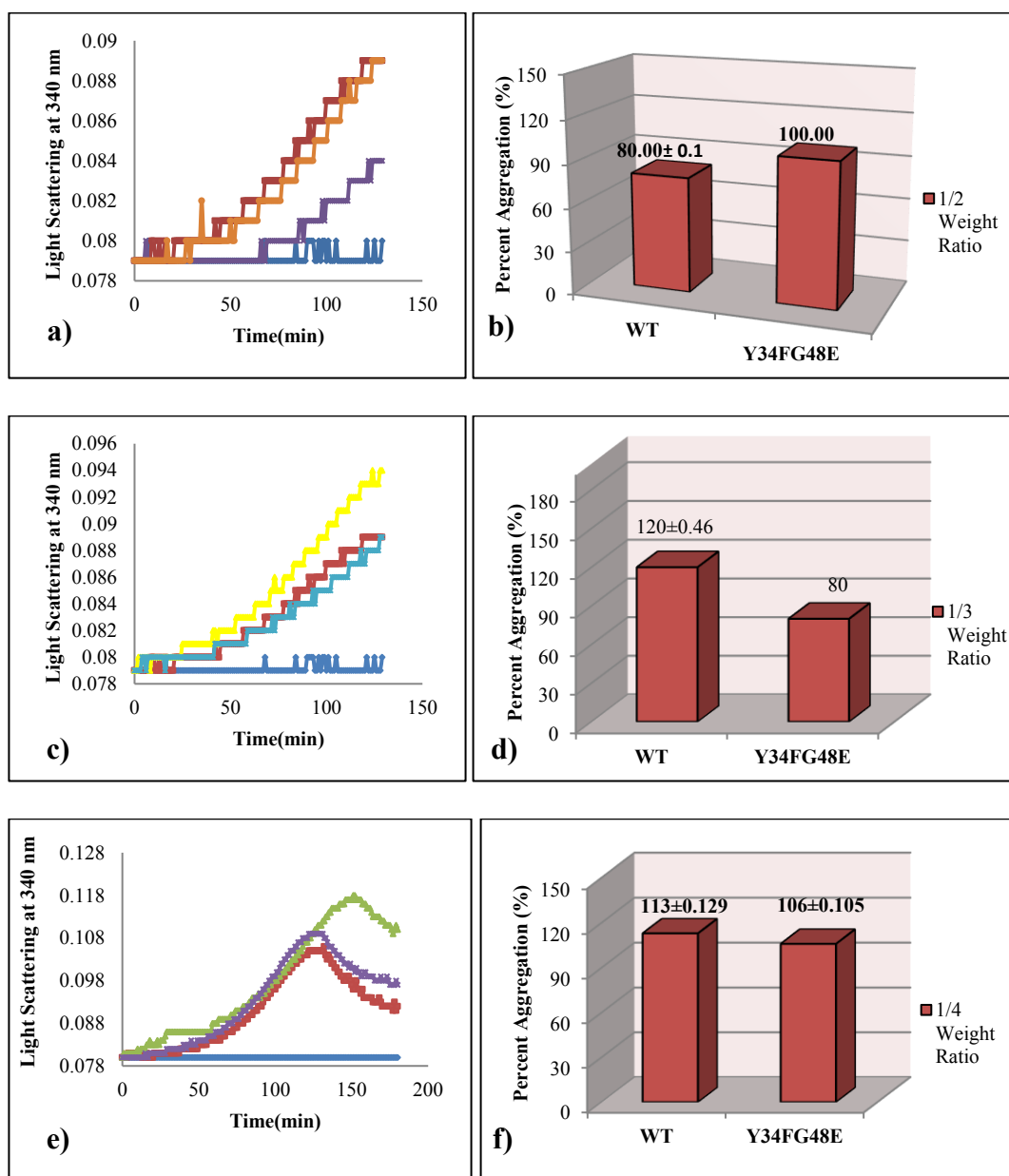


**Figure 3.92. ADH aggregation assay with I108K mutant sHSP.** Heat induced aggregation of ADH in the absence or presence of the WT *Tpv* sHSP 14.3 and I108K proteins were monitored by measuring light scattering at 340 nm. The tested w/w ADH/sHSP ratio were 1/2(a), 1/3(c). (—●—) Blank, (—■—) ADH alone, (—\*—) ADH + WT 1:2 w/w ratio, (—○—) ADH + I108K 1:2 w/w ratio, (—▲—) ADH + WT 1:3 w/w ratio, (—□—) ADH + I108K 1:3 w/w ratio. Interpretation of the results were done based on percent aggregation where the aggregation of ADH alone was taken as 100 % ( $100 \pm 0.102$ ) (b-d). Each experiment was repeated at least three times.

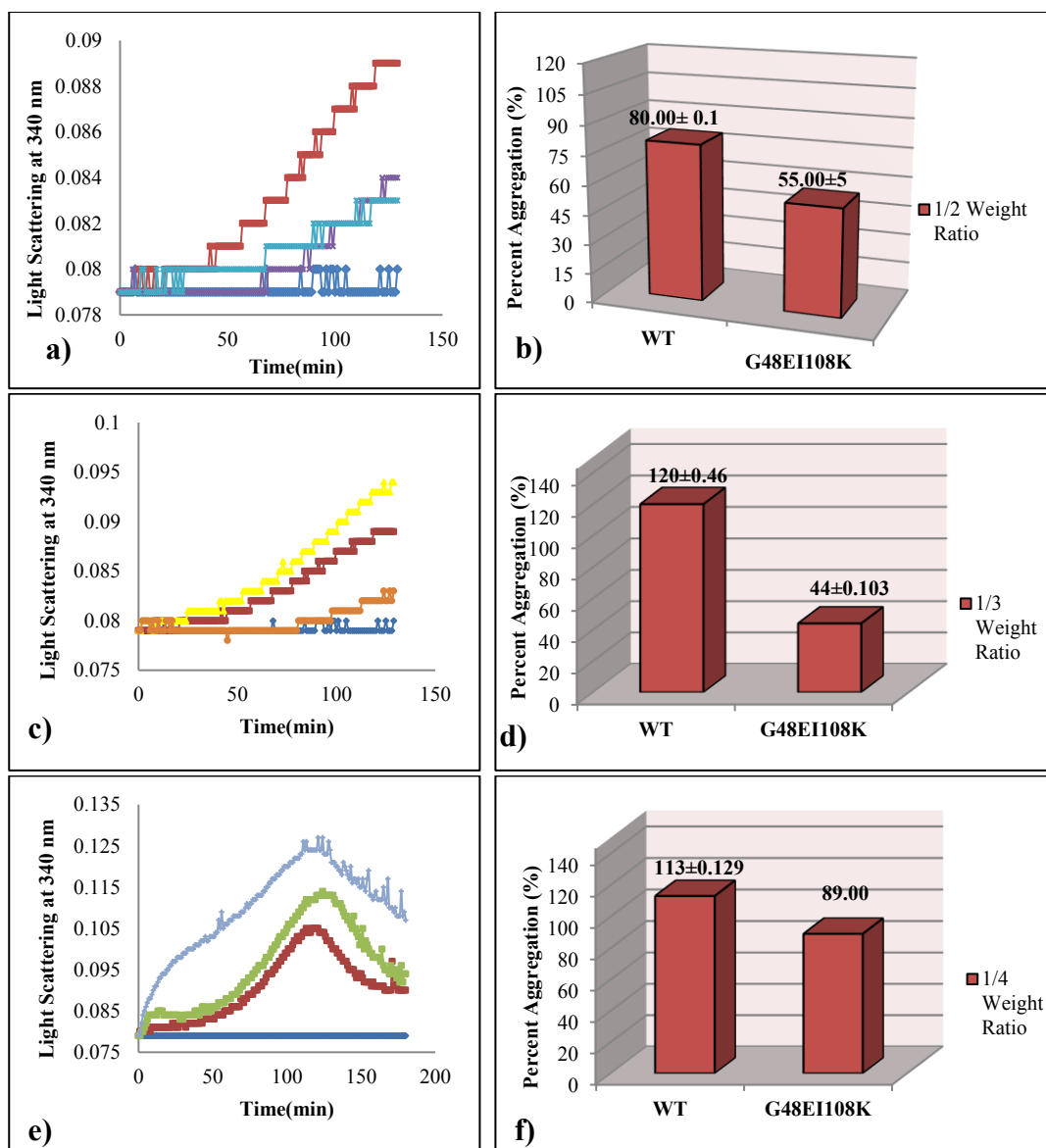
In addition, the efficiencies of the Y34FG48E double mutant sHSP in protection of the ADH aggregation was almost same as control and slightly higher than WT (7%) at 1:4 ADH: sHSP w/w ratio. Similar to the its respective single variants (*i.e.*, Y34F, G48E), the double mutant sHSP also reduced aggregation 20% with respect to control at 1:3 ADH/sHSP w/w ratio. On the other hand, at low concentration (1:2 ADH/sHSP w/w ratio), the double mutant sHSP had insignificant effect on prevention of ADH aggregation, which was not found to be different than control (Figure 3.93).

Among seven mutant variants, at all three-molar ratio, G48EI108K double mutant provided the highest heat-protection effect against aggregation of ADH, as compared to the WT *Tpv* sHSP14.3 and control (without sHSP14.3). ADH

aggregation was reduced to 56% and 45% when 3-fold and 2-fold excess mutant sHSP was added to the reaction mixture, respectively (Figure 3.94).



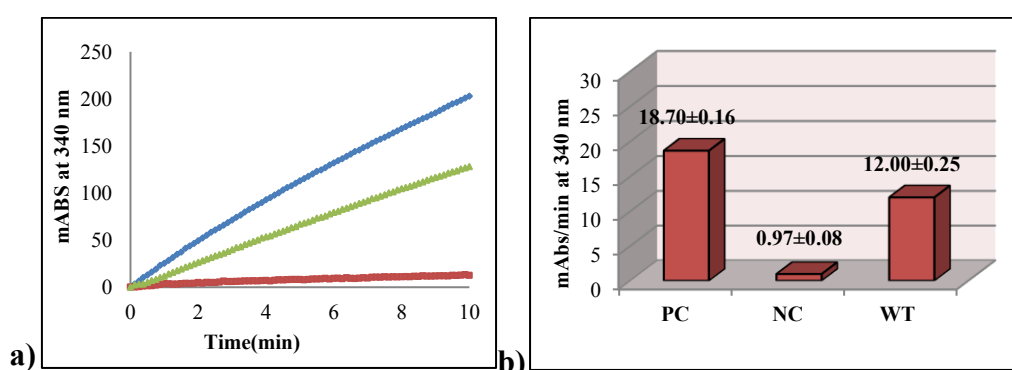
**Figure 3.93. ADH aggregation assay with Y34FG48E mutant sHSP.** Heat induced aggregation of ADH in the absence or presence of the WT *Tpv* sHSP 14.3 and Y34FG48E proteins were monitored by measuring light scattering at 340 nm. The tested w/w ADH/sHSP ratio were 1/2(a), 1/3(c) and 1/4(e). (—●—) Blank, (—■—) ADH alone, (—\*—) ADH + WT 1:2 w/w ratio, (—○—) ADH + Y34FG48E 1:2 w/w ratio, (—▲—) ADH + WT 1:3 w/w ratio, (—□—) ADH + Y34FG48E 1:3 w/w ratio, (—△—) ADH + Y34FG48E 1:4 w/w ratio, (—×—) ADH + WT 1:4 w/w ratio. Interpretation of the results were done based on percent aggregation where the aggregation of ADH alone was taken as 100 % (100±0.102) (b-d-f). Each experiment was repeated at least three times.



**Figure 3.94. ADH aggregation assay with G48E1108K mutant sHSP.** Heat induced aggregation of ADH in the absence or presence of the WT *Tpv* sHSP 14.3 and G48E1108K proteins were monitored by measuring light scattering at 340 nm. The tested w/w ADH/sHSP ratio were 1/2(a), 1/3(c) and 1/4(e). (—●—) Blank, (—■—) ADH alone, (—\*—) ADH + WT 1:2 w/w ratio, (—◆—) ADH + G48E1108K 1:2 w/w ratio, (—▲—) ADH + WT 1:3 w/w ratio, (—○—) ADH + G48E1108K 1:3 w/w ratio, (—+—) ADH + G48E1108K 1:4 w/w ratio, (—■—) ADH + WT 1:4 w/w ratio. Interpretation of the results were done based on percent aggregation where the aggregation of ADH alone was taken as 100 % (100±0.102) (b-d-f). Each experiment was repeated at least three times.

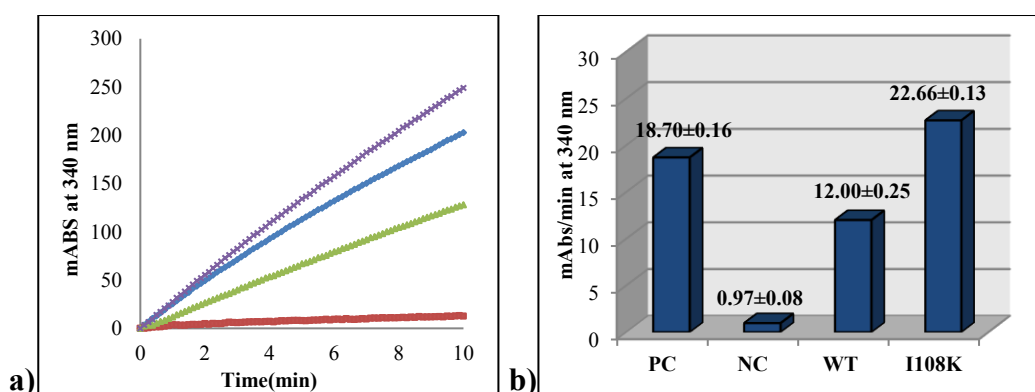
### 3.12 Chaperone Activity Assay with Yeast Alcohol Dehydrogenase

The selected *Tpv* ACD mutants (*i.e.*, L33S, Y34F, E43V, G48E, K87E, I108K, Y34FG48E and G48EI108K) were also investigated for their chaperone activities to protect the ADH activity from heat inactivation. Our results showed that the activity of ADH decreased sharply (19.2-fold) after incubation at 47°C for 20 min in the absence of chaperone (negative control, NC). On the other hand, the ADH activity increased about 12- fold in the presence of WT *Tpv* sHSP 14.3 with respect to the negative control (Figure 3.95).



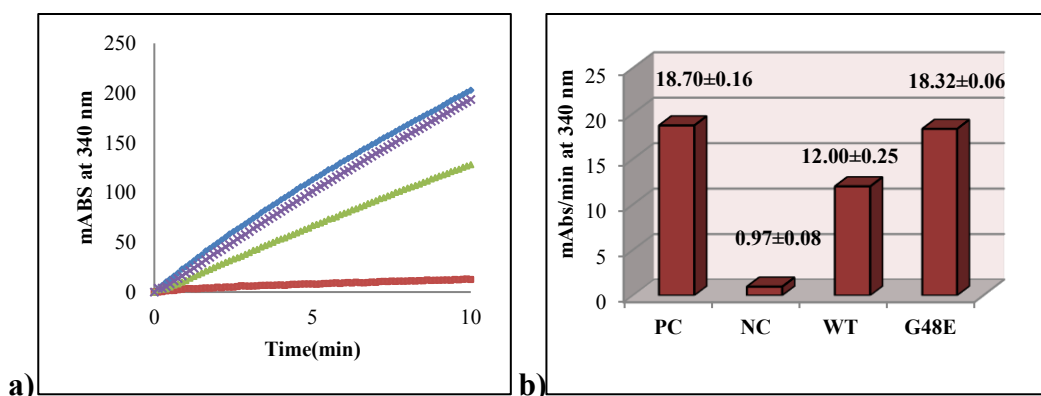
**Figure 3.95. ADH activity assay in the absence and presence of WT *Tpv* sHSP 14.3.** The ADH was heated at 47°C for 20 min in the presence or absence of WT. The rate of reduction of NAD<sup>+</sup> measured at 340 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (—●—) PC, (—■—) NC, (—▲—) WT (b). PC: Positive Control, activity measured before heat-treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

Among eight *Tpv* sHSP 14.3 ACD mutants, I108K mutant showed the highest chaperone activity. The thermal inactivation of the ADH at 47°C was strongly repressed by the I108K mutant sHSP to an extent of 1.2-fold more than that of positive control. Its heat protection effect was also considerably higher than the WT sHSP (Figure 3.96).



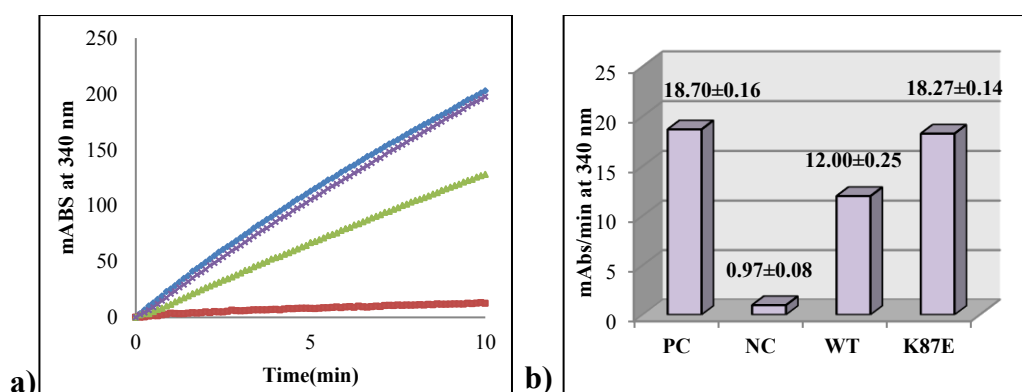
**Figure 3.96. ADH activity assay in the absence and presence of I108K *Tpv* sHSP 14.3.** The ADH was heated at 47°C for 20 min in the presence or absence of I108K. The rate of reduction of NAD<sup>+</sup> measured at 340 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (—●—) PC, (—■—) NC, (—▲—) WT (—×—) I108K (b). PC: Positive Control, activity measured before heat-treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

The G48E and K87E single mutants displayed about 98% protection against the heat induced inactivation of ADH as compared to PC. The protection effect of these mutants was 1.53-fold higher than WT (Figure 3.97-3.98).



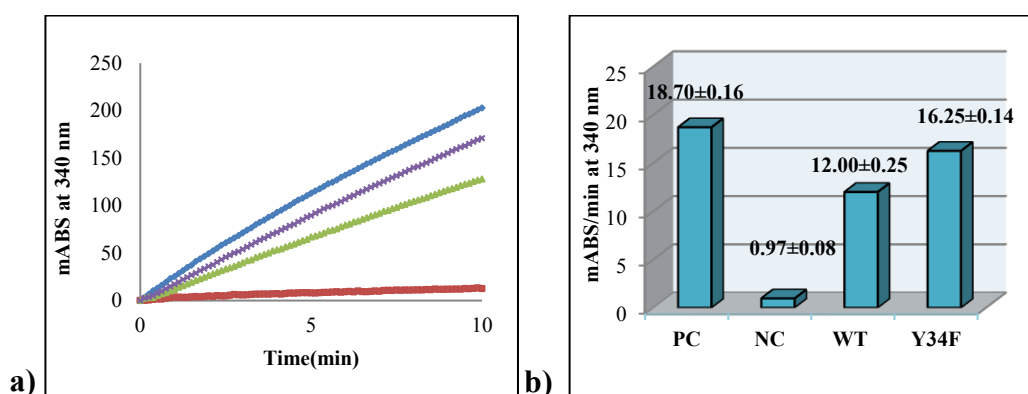
**Figure 3.97. ADH activity assay in the absence and presence of G48E *Tpv* sHSP 14.3.** The ADH was heated at 47°C for 20 min in the presence or absence of G48E. The rate of reduction of NAD<sup>+</sup> measured at 340 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (—●—) PC, (—■—) NC, (—▲—) WT (—×—) G48E. (b) PC: Positive Control, activity measured before heat-treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



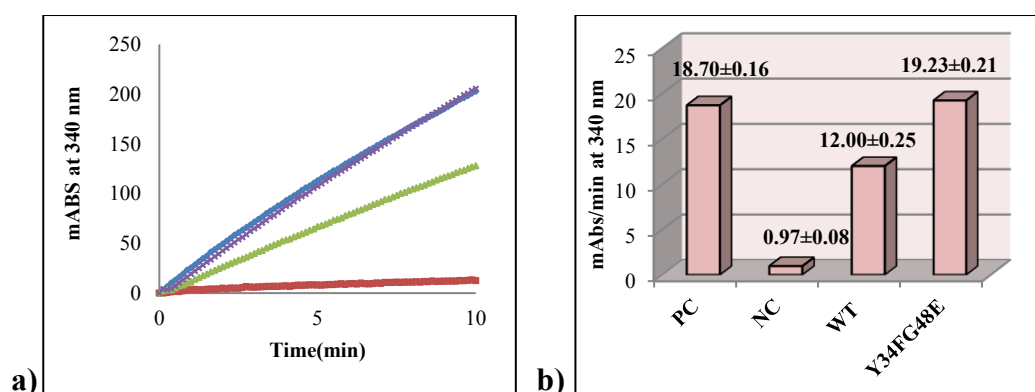


**Figure 3.98. ADH activity assay in the absence and presence of K87E *Tpv* sHSP 14.3.** The ADH was heated at 47°C for 20 min in the presence or absence of K87E. The rate of reduction of NAD<sup>+</sup> measured at 340 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (—♦—) PC, (—■—) NC, (—▲—) WT (—×—) K87E. (b) PC: Positive Control, activity measured before heat-treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

Besides, introduction of the highly hydrophobic amino acid to generate single mutant Y34F resulted in enhanced chaperone activity which is 1.4 – fold higher than WT sHSP (Figure 3.99). On the other hand, protection effect of the Y34FG48E double mutant was found to be better than that of respective single variants. This double mutant sHSP strongly repressed thermal inactivation of the ADH (*i.e.*, its activity was 1.6-fold higher than WT) and its chaperone activity was almost equal to PC (Figure 3.100).

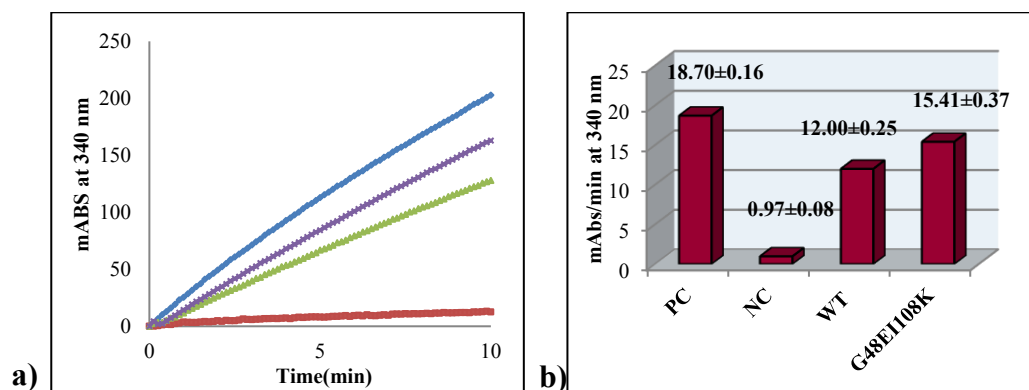


**Figure 3.99. ADH activity assay in the absence and presence of Y34F *Tpv* sHSP 14.3.** The ADH was heated at 47°C for 20 min in the presence or absence of Y34F. The rate of reduction of NAD<sup>+</sup> measured at 340 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (—♦—) PC, (—■—) NC, (—▲—) WT (—×—) Y34F. (b) PC: Positive Control, activity measured before heat-treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

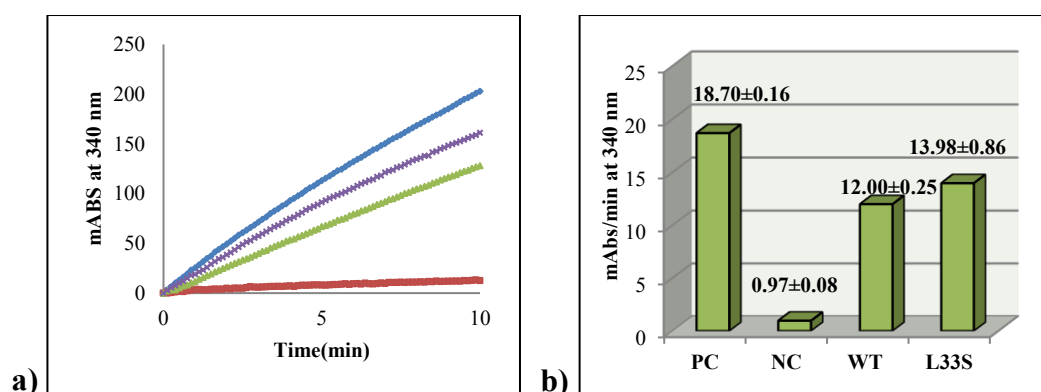


**Figure 3.100. ADH activity assay in the absence and presence of Y34FG48E *Tpv* sHSP 14.3.** The ADH was heated at 47°C for 20 min in the presence or absence of Y34FG48E. The rate of reduction of NAD<sup>+</sup> measured at 340 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (—●—) PC, (—■—) NC, (—▲—) WT (—×—) Y34FG48E (b) PC: Positive Control, activity measured before heat- treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

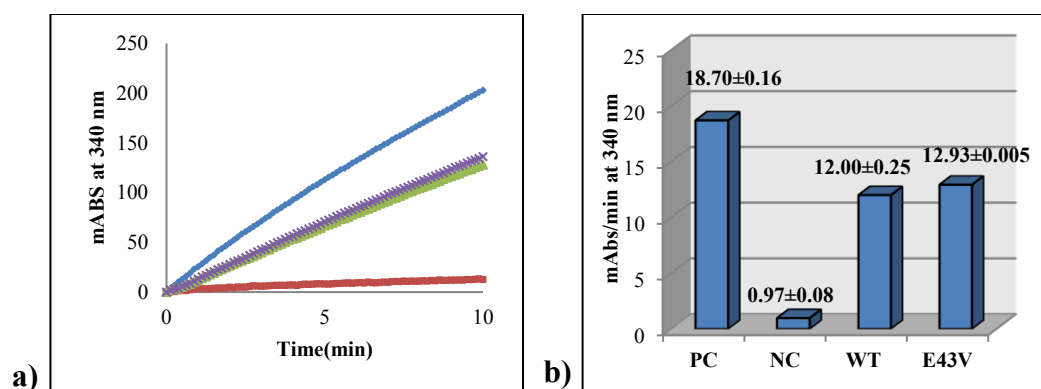
ADH activity assay also revealed that protection effect of the G48E1108K double mutant was 20 % higher than that of WT in preventing the ADH from inactivation at high temperature. However, its chaperone activity was less than the respective single variants (Figure 3.101). On the other hand, L33S and E43V mutations did not alter the heat protection efficiency of *Tpv* sHSP 14.3 sHSP which was almost same as the wild type (Figure 3.102-3.103).



**Figure 3.101. ADH activity assay in the absence and presence of G48E1108K *Tpv* sHSP 14.3.** The ADH was heated at 47°C for 20 min in the presence or absence of G48E1108K. The rate of reduction of NAD<sup>+</sup> measured at 340 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (—●—) PC, (—■—) NC, (—▲—) WT (—×—) G48E1108K. (b) PC: Positive Control, activity measured before heat- treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.



**Figure 3.102. ADH activity assay in the absence and presence of L33S *Tpv* sHSP 14.3.** The ADH was heated at 47°C for 20 min in the presence or absence of L33S. The rate of reduction of NAD<sup>+</sup> measured at 340 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (—♦—) PC, (—■—) NC, (—▲—) WT (—×—) L33S. (b) PC: Positive Control, activity measured before heat-treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

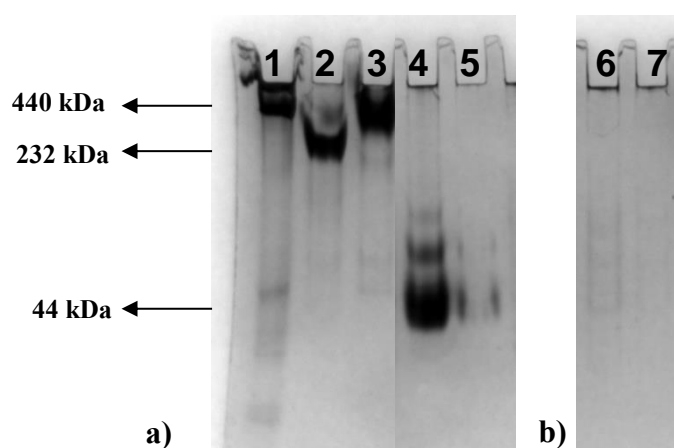


**Figure 3.103. ADH activity assay in the absence and presence of E43V *Tpv* sHSP 14.3.** The ADH was heated at 47°C for 20 min in the presence or absence of E43V. The rate of reduction of NAD<sup>+</sup> measured at 340 nm (a). The slope of the initial linear increase in absorption was used to calculate the rate of reaction (—♦—) PC, (—■—) NC, (—▲—) WT (—×—) E43V. (b) PC: Positive Control, activity measured before heat-treatment. NC: Negative Control, remaining activity after heat treatment in the absence of chaperone. Data shown are mean values with STD values indicating the standard deviation of the at least three independent experiments.

### 3.13 Native Polyacrylamide Gel Electrophoresis

In order to assess the oligomeric states of the wild type and mutant *Tpv* sHSP 14.3 proteins, Polyacrylamide Gel Electrophoresis was performed under native conditions. In this system, proteins are separated based on their charge to mass ratios. First of all, wild type and G48E mutant proteins were analyzed by Native PAGE.

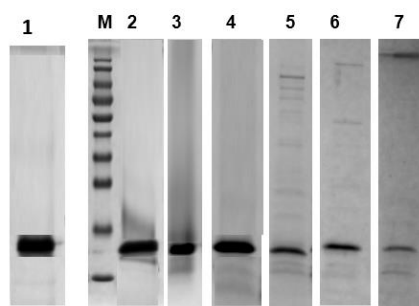
Five protein standards (ranging in molecular weight from ~13.7 to 440 kDa) were used to estimate the molecular weight of the WT and mutant variant forms. However, *Tpv* sHSP 14.3 WT and G48E mutant protein remained at the top of the gel and did not migrate through the stacking gel properly (Figure 3.104). Therefore, stained proteins were not detected on the gel clearly. Therefore, the sample resolution and separation were not good. This could be related to the net charge of the protein such that wild type protein (pI=8.5) has zero charge in native gel buffer at pH 8.49 and G48E (pI=7.70) has slightly negative charge (*i.e.*, -0.981) at this pH. Besides, other *Tpv* sHSP 14.3 ACD mutants, Y34F(pI=8.523), I108K(pI=8.794), and G107A (pI=8.504), have neutral charge at pH 8.49. Also, E43V mutant protein(pI=8.816) has basic character with a net positive charge (+1.02) at this pH. On the the hand, K87E mutant protein(pI=6.346) has negative charge (-1.93). Considering the charge properites of the mutant proteins, this technique was not found suitable to be used for the separation of the oligomeric forms. In order to solve this problem, Blue Native PAGE was performed using Coomassie Blue G-250, which denotes negative charges to all proteins regardless of their charge and size in native states and shifts the pI values of the proteins to more negative values.



**Figure 3.104. Oligomeric state analysis by native gel.** a) Molecular weight standards: 1: Ferritin :440 kDa, 2: Catalase: 232 kDa, 3: Aldolase:158 kDa, 4:Ovalbumin:44 kDa, 5: Ribonuclease A: 13.7 kDa. b) 6-7: Oligomeric state analysis of *Tpv* sHSP14.3 WT and G48E mutant protein, respectively by 4%-20 % native gel.

### 3.14 Blue Native PAGE Analysis

Among the *Tpv* ACD mutants generated, those which showed high stability with good protein concentration at higher temperatures between 60-70°C (*i.e.*, Y34F, G48E, K87E, G107A, I108K and G48E/I108K mutants) were selected for BN-PAGE analysis. The protein samples (in Lysis Buffer) were exchanged by low salt containing buffer (50 mM NaCl+ 50 mM imidazole, pH 7.06) by using ultrafiltration (filter units 5 kDa MWCO, Centriscart I, Sartorius, 10 kDa MWCO, Amicon, Millipore, and 30kDa MWCO, Pierce Protein Concentrator, ThermoScientific). The recovery of protein samples after ultrafiltration is shown in the Figure 3.105.



**Figure 3.105. SDS PAGE analysis of WT *Tpv* sHSP14.3 and G48E, I108K, K87E, E43V, Y34F and G48E/I108K mutant sHSPs after buffer exchange.** M: PageRuler™ Prestained Protein Ladder (10-170 kDa), **1 and 2:** WT sHSP(14.3 kDa) and G48E mutant protein retentate(14.405 kDa) obtained by 30kDa MWCO filter unit, respectively. **3:** I108K mutant protein retentate (14.348 kDa) obtained by 10 kDa MWCO filter unit.**4:** K87E mutant protein retentate (14.334 kDa) obtained by 5 kDa MWCO filter unit. **5:** G107A mutant protein retentate (14.339 kDa) obtained by 5 kDa MWCO filter unit. **6:** Y34F mutant protein retentate (14.317 kDa) obtained by 5 kDa MWCO filter unit.**7:** G48E/I108K mutant protein retentate (14.42 kDa) obtained by 5 kDa MWCO filter unit.

Oligomeric states of *Tpv* sHSP14.3 WT and its variants were evaluated by blue native polyacrylamide gel electrophoresis (BN-PAGE). In this method, addition of an anionic dye, Coomassie blue G-250, changes the isoelectric point of the proteins to more negative values so that all proteins including basic ones gain ability to migrate to the anode regardless of their original pI values (Schägger, 2004). The oligomeric forms of the sHSP samples without heat treatment and by pre-heating at 70°C for 10 min were analyzed by BN-PAGE. Ferritin (880 and 440 kDa), Catalase (232 kDa), Aldolase (158 kDa) and Ribonuclease A (13.7 kDa) were used as standards (See Appendix F). The results of *Tpv* sHSP 14.3 WT sHSP oligomers were

found in equilibrium between large oligomeric (*i.e.*, 60-mer, 54-mer, 50-mer and 24-mer) and small oligomeric species (*i.e.*, <12-mer, tetramer and dimer) (Figure 3.106-a). The apparent molecular weight of higher-order oligomers of the WT sHSP were between 326 and 864 kDa (24-60 mer). The most notable small oligomeric species were 12-mer oligomers and dimers. There were also some intermediate forms between them but less apparent in the heat-treated WT sHSP sample.

As revealed by BN-PAGE result dimeric state was dominating to other oligomeric species of the G48E mutant variant (Figure 3.106-b). Clear bands for 24-mer and 12-mer oligomeric forms were also observed when the mutant protein sample is not heated or pre-heated before loading. Besides, there was a distinct but weak band corresponding to 8 mer. However, the bands corresponding the large oligomers were hardly visible. Presence of smaller oligomeric species (from dimer to 24-mer oligomer), but disappearance of large oligomers (*e.g.*, 60-mer oligomers) may indicate the perturbation of oligomer assembly as a result of G48E point mutation.

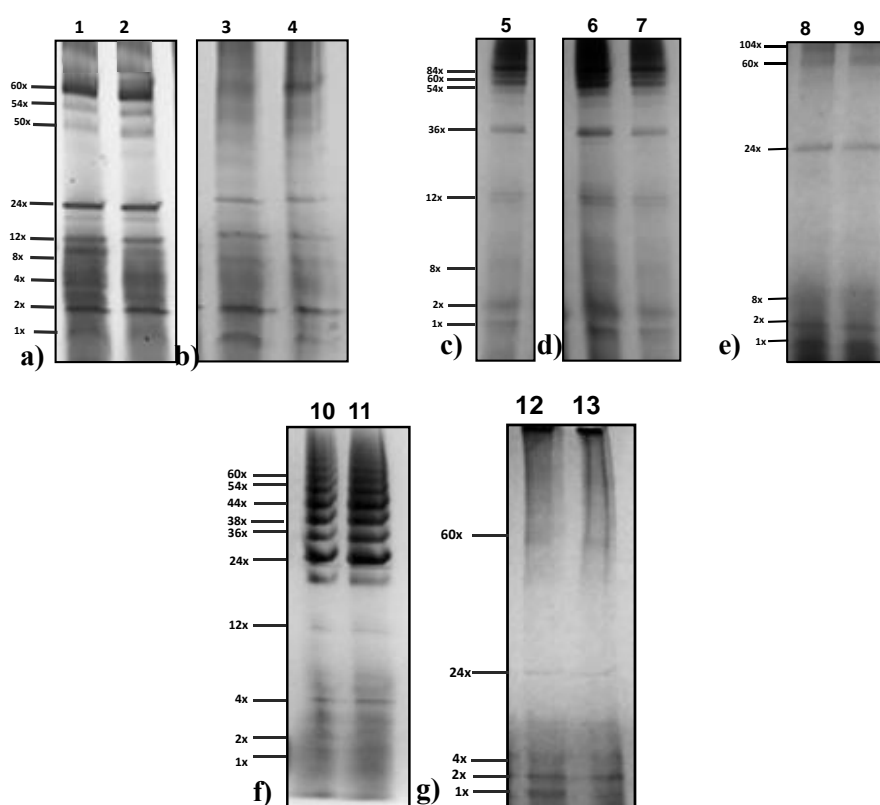
Oligomeric forms of the Y34F mutant analyzed by BN-PAGE is shown in the Figure 3.106-c. This result indicated that Y34F mutation shifted the equilibrium of oligomeric species from smaller oligomers (<36 mer) to larger oligomers. Different from WT, a series of bands corresponding to assemblies of molecular weights more than 880 kDa (>60 mer) were observed. Although there were bands corresponding to 12-mer oligomers, they were not visible enough (Figure 3.106-c).

The BN-PAGE analysis of I108K mutant showed abundance of larger oligomeric forms (36-84 mer), besides dimeric species. Different from the WT, larger size oligomers, those having subunits more than 60 mer was detected in an oligomer of I108K. In addition to these, I108K mutant sHSP had intermediate size oligomers of 12 and 8 subunits (Figure 3.106-d).

Oligomer size distribution analysis of the double mutant G48EI108K by BN-PAGE showed that the bands of monomeric, dimeric and larger oligomeric forms (*i.e.*, 60-mer) were more pronounced (Figure 3.106-e). The relative amount of intermediate forms between 60-mer and dimer, including the archaeal sHSP specific 24-subunit oligomers was significantly less than that of the WT sHSP.

Oligomer distribution of K87E mutant *Tpv* sHSP14.3 as revealed by BN-PAGE analysis showed that K87E mutation favored the formation of higher order oligomeric ensembles (>24-mer, or >880 kDa molecular weight species). The band corresponding to characteristic 24-mer was more pronounced among the others (Figure 3.106-f).

Oligomer size distribution analysis of G107A mutant by BN-PAGE revealed that this mutant sHSP behaved similar to WT by forming 60-mer, 24 mer and smaller species (4-mer, 2 mer and monomer). Expansion into the higher oligomeric ensembles was not observed for G107A mutant. This result suggests that smaller species are more favorable for G107A mutant (Figure 3.106-g).



**Figure 3.106. Blue Native Gel Analysis for WT and selected *Tpv* sHSP14.3 mutants.** The oligomeric forms of the sHSP samples without heat treatment and pre-heating at 70°C for 10 min were analyzed by 4%-13% BN-PAGE. **a-b)** BN-PAGE analysis of the WT *Tpv* sHSP14.3 and G48E mutant sHSP without heat treatment (**1,3**) and with heat treatment (**2,4**) respectively. **c) 5:** BN-PAGE analysis of the Y34F mutant sHSP without heat treatment. **d-e)** BN-PAGE analysis of the I108K mutant sHSP and G48E/I108K mutant sHSP without heat treatment (**6,8**) and with heat treatment (**7,9**) respectively. **f-g)** BN-PAGE analysis of the K87E mutant sHSP and G107A mutant sHSP without heat treatment (**10,12**) and with heat treatment (**11,13**) respectively.

### 3.15 Primary and Secondary Structure Analysis of *Tpv* sHSP14.3 Protein

The alpha crystallin domain of *Tpv* sHSP consists of 82 amino acids in its primary sequence (Figure 3.107). It has high proportion of hydrophobic amino acids (37.8 %), and neutral amino acids (29.27 %). In its sequence, 17.07 % and 15.85 % basic and acidic amino acids, respectively. Thirteen of the negatively charged residues (D, E) are found in the ACD of this protein. Out of 14 basic amino acids, 10 of them are lysine residues. The amino acid composition is closely related to the stability of the proteins. The thermophiles have high content of acidic and basic amino acids which provide an advantage to them in adapting to high growth optimum temperature (Luo & Robb, 2011).

Jpred created in Jalview (ver. 2.11.1.3) was used for secondary structure prediction of *Tpv* sHSP 14.3 (Figure 3.107). Highly conserved residues were found in the core alpha crystallin domain spanning residues L33-K114 in *Tpv* sHSP 14.3 protein. This domain is composed of eight beta strands (termed  $\beta$ 2- $\beta$ 9) and the characteristic motif "G-x-L" is seen towards the end of the ACD. N- terminal region of the protein is poorly conserved and contains small quantities of  $\alpha$ -helices. On the other hand, highly variable C- terminal region is unstructured.

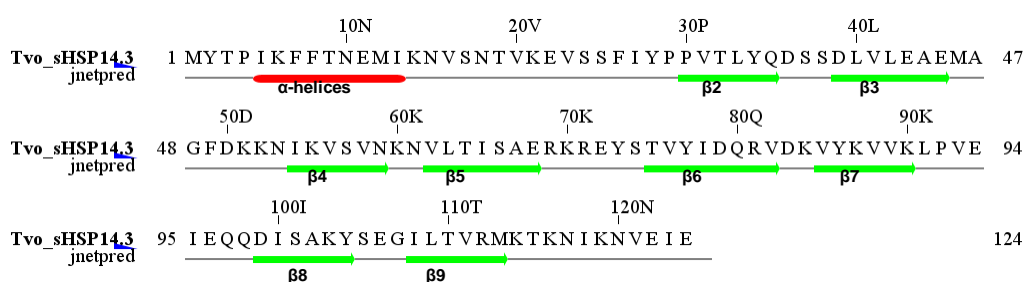


Figure 3.107. Secondary structure prediction of the *Tpv* sHSP14.3 protein.

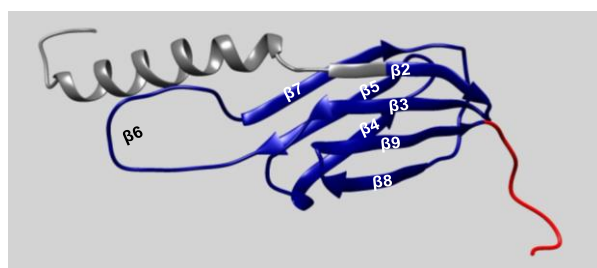
### 3.16 Computational Modelling of 3-D Structure of *Tpv* sHSP14.3

Three dimensional structure of *Thermoplasma volcainum* sHSP14.3 was generated by homology modelling using crystal structures of the *S.tokodaii* (PDB entry 3AAC



and 3VQM), *X. axonopodis* (PDB entry 3GLA) and *D. Radiodurans* (PDB entry 4FEI) as templates. In the monomeric structure, short helices in the beginning of the N-terminal domain between the Phe8 and Ile27 are present. The residues Met1-Thr32 were positioned in the NTD of the protein. Two residues, Val41 and Thr32, are located on the  $\beta$ 2 strand near the beginning of the alpha crystallin domain (Leu33-Lys114). In the ACD, monomeric structure has seven beta strands and it is like an immunoglobulin fold consisting of two anti parallel layer such that  $\beta$ 2 (residues spanning Val31-Asp36),  $\beta$ 3 (Leu40-Glu45),  $\beta$ 8 (Ser101-Glu109),  $\beta$ 9 (Ile108-Met113) form one layer and  $\beta$ 4 (Lys51-Asn59),  $\beta$ 5 (Val62-Glu68) and  $\beta$ 7 (Lys84-Lys90) together with distinct  $\beta$ 6 (Arg69-Asp83) for the other sheet (Figure 3.108). Secondary structure analysis of the ACD of *Tpv* sHSP14.3 integrated with multiple sequence alignment of the sequences from archaeal species showed similar results with that of our generated 3-D model. Also, beginning of the  $\beta$ 4 strand, small helical structure was found both in the 3-D model and in the primary sequence analysis (Figure 3.109).

The dimer is created with the same monomer of *Tpv* sHSP14.3 based on the crystal structure of *S.tokodaii* (PDB entry 3AAC) using the Chimera Tool. In the dimer, a large  $\beta$ 6 loop protrudes from a  $\beta$  sheet between  $\beta$ 5 and  $\beta$ 7 and positioned close to the  $\beta$ 2 sheet of the adjacent monomer making the dimer interface (Figure3.110). This is the potential dimerization points involved in monomer-monomer interactions as seen in bacterial type sHSPs (Haslbeck *et al.*, 2019). Also, residues in the core alpha crystallin domain of *Tpv* sHSP 14.3 targeted for mutagenesis (L33, Y34, E43, E45, A47, G48, G107 and I108) were located at the putative dimer interface as predicted by the National Center for Biotechnology Information's Conserved Domain Database (CCD) search. Besides, the C-terminal domain of the protein, CTD, is unstructured and it is positioned between the residues Thr115-Glu124.



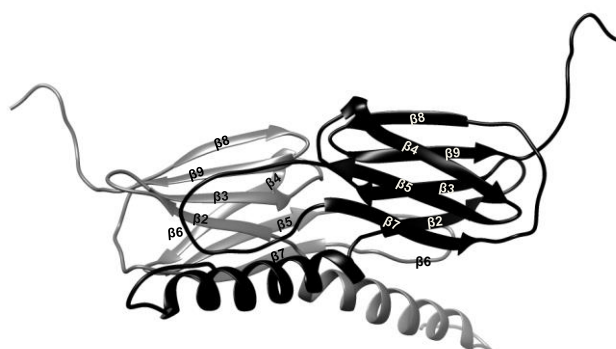
**Figure 3.108. 3D Structure of *Tpv* HSP 14.3 generated by homology modelling.** Ribbon model of the monomeric *Tpv* sHSP 14.3. Dark gray, blue, and red indicate the NTD, ACD and CTD of the *Tpv* sHSP 14.3 protein.

|            |    | $\beta 2$ | $\beta 3$ | helix   | $\beta 4$ | $\beta 5$ | $\beta 6$       |            |            |        |     |
|------------|----|-----------|-----------|---------|-----------|-----------|-----------------|------------|------------|--------|-----|
| <i>Tvo</i> | 33 | -LYQDSS   | DLVLEAE   | MAGFD   | KKNI      | IKVSVN    | -KNVLTISAERKREY | -----STVY  | -IDQRV     | 82     |     |
| <i>Tac</i> | 33 | -MYQDGT   | DLVLEAE   | MPGFD   | KKDI      | IKITVE    | -KNVLTIRAERKREY | -----RSVY  | -IDQRV     | 82     |     |
| <i>Fac</i> | 33 | -MYEDNG   | YIGIEAD   | LPGFS   | REDI      | IKVTLE    | -KNAIVIRAERE    | IKP        | -----EGTVF | -ENQRP | 83  |
| <i>Pto</i> | 33 | -MYQENGY  | IYIDL     | DMPGFK  | KDNI      | ISVTLE    | -KSYVVINASRE    | INK        | -----GGTVF | -ENQRP | 83  |
| <i>Sto</i> | 28 | DMYEEGG   | YLVVVAD   | LAGFN   | KEKI      | KARVSG    | QNELIIEAERE     | ITE        | -----PGVKY | -LTQRP | 80  |
| <i>Sso</i> | 29 | DMYEEGG   | ELVVVAD   | LAGFN   | KDKI      | ISVRLS    | AQNELIINAERE    | IQY        | -----IGTKY | -ATQRP | 81  |
| <i>Mja</i> | 47 | -IIEGDQ   | HIKVI     | AWLPGVN | KEDI      | ILNAV     | G-DTLEIRAKR     | SPLMITESER | IYSEI      | PEE    | 102 |

|            |     |    |        |   |   |   |   |   |   |   |   |    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |     |     |
|------------|-----|----|--------|---|---|---|---|---|---|---|---|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----|-----|-----|
| <i>Tvo</i> | 83  | DK | VYKVVK | L | P | V | E | I | E | Q | - | QD | I | S | A | K | Y | S | E | G | I | L | T | V | R | M | K | - | 114 |     |     |
| <i>Tac</i> | 83  | DK | VFKVVR | L | P | V | D | V | D | Q | - | AS | I | S | A | K | Y | Q | D | G | L | L | T | V | R | M | K | - | 114 |     |     |
| <i>Fac</i> | 84  | EK | VFKRMS | L | P | M | E | V | D | T | E | Q  | E | F | S | A | K | Y | N | D | G | V | L | S | L | K | I | P | -   | 116 |     |
| <i>Pto</i> | 84  | SK | VFKRI  | Q | L | P | G | E | P | D | K | N  | A | D | V | S | A | K | Y | E | D | G | V | L | H | L | S | I | P   | -   | 116 |
| <i>Sto</i> | 81  | KY | VRKVI  | R | L | P | Y | N | V | A | K | D  | A | E | I | S | G | K | Y | E | N | G | V | L | T | I | R | I | P   | -   | 113 |
| <i>Sso</i> | 82  | LK | IHKVIR | L | P | V | K | V | K | R | D | S  | Q | V | T | A | K | Y | E | N | G | V | L | T | I | R | I | P | -   | 114 |     |
| <i>Mja</i> | 103 | EE | IYRTIK | L | P | A | T | V | K | E | - | EN | A | S | A | K | F | E | N | G | V | L | S | V | I | L | P | K | -   | 135 |     |

**Figure 3.109. Secondary structure analysis of the ACD of *Tpv* sHSP14.3 together with multiple sequence alignment of sHSPs from archaeal species.** *Tvo*: *Thermoplasma volcanium* GSS1, *Tac*: *Thermoplasma acidophilum*, *Fac*: *Ferroplasma acidiphilum*, *Pto*: *Picrophilus torridus*, *Sto*: *Sulfolobus tokodaii*, *Sso*: *Saccharolobus solfataricus*, *Mja*: *Methanocaldococcus jannaschii*. MSA is performed using Clustal W programme and secondary structure prediction is performed by PRALINE online tool. The image is viewed by Jalview. Beta strands and helices are colored by blue and red, respectively.



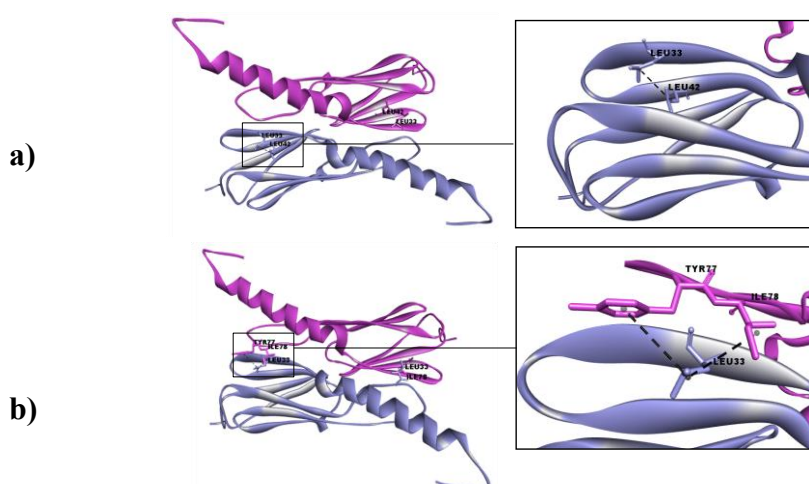
**Figure 3.110. Ribbon model of the dimeric *Tpv* sHSP 14.3.** Light gray indicating the one monomer and dark gray indicating the adjacent monomer.

### 3.17 Comparative Analysis of Intra- and Inter- Molecular Interactions of Wild Type *Tpv* sHSP 14.3 protein and its mutant variants

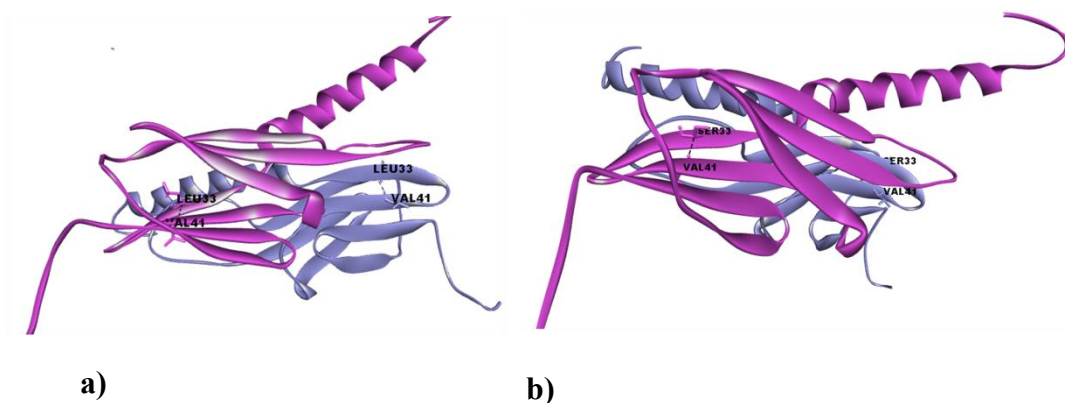
Detailed analysis of the intra-/inter- molecular bonds of the mutant sHSPs were performed in comparison with the WT sHSP by using DSV Programme. Following is a brief description of the changes in the bond interactions. Tables containing complete list of the molecular bonds in the mutant sHSP variants as compared to the WT *Tpv* sHSP14.3 are given in Appendix G.

#### 3.17.1 *Tpv* sHSP14.3 L33S Mutant Variant

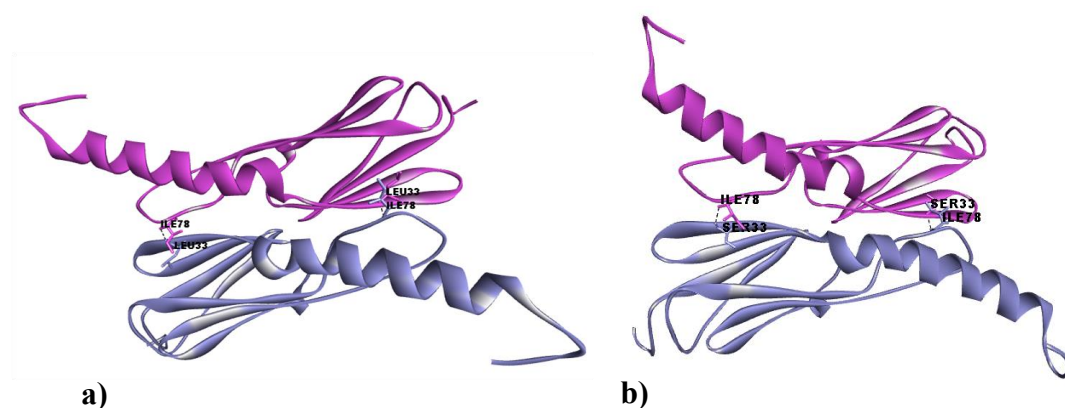
The substitution of the hydrophobic amino acid leucine with a polar amino acid serine at the position 33 (L33S) disrupted the intramolecular hydrophobic interactions between Leu33-Leu42 and as well as the intermolecular hydrophobic interaction between Leu33-Ile78 and Tyr77-Leu33 (Figure 3.111), but intramolecular hydrogen bonds remained same as compared to WT (Figure 3.112). Also, L33S mutant gained one more intermolecular hydrogen bond formed between nitrogen atom of Ile78 and oxygen atom of Ser33 (Figure 3.113). Besides, distance between the hydrogen bonds present in the L33S mutant decreased upon mutation (Appendix G, Table G.1).



**Figure 3.111. Intra (a) and intermolecular (b) hydrophobic bonds of *Tpv* sHSP14.3 WT at residue Leu33. Pink indicating the one monomer and blue indicating the partner monomer.**



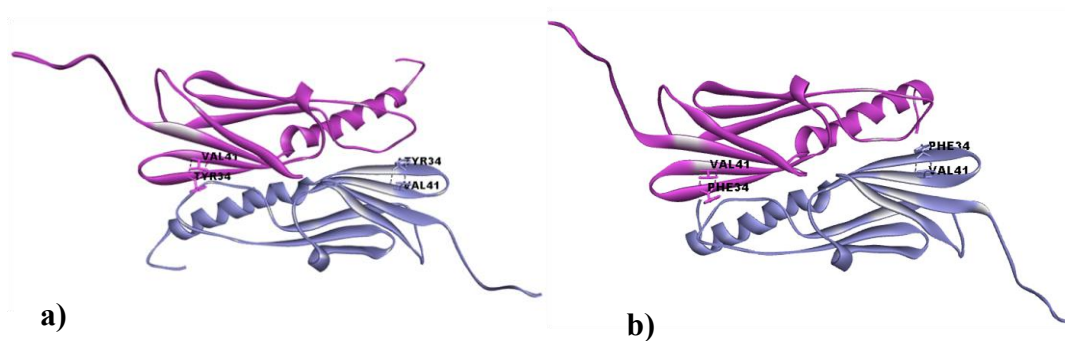
**Figure 3.112. Intramolecular hydrogen bonds of *Tpv* sHSP14.3 WT (a) and L33S mutant (b).** Pink indicating the one monomer and blue indicating the partner monomer.



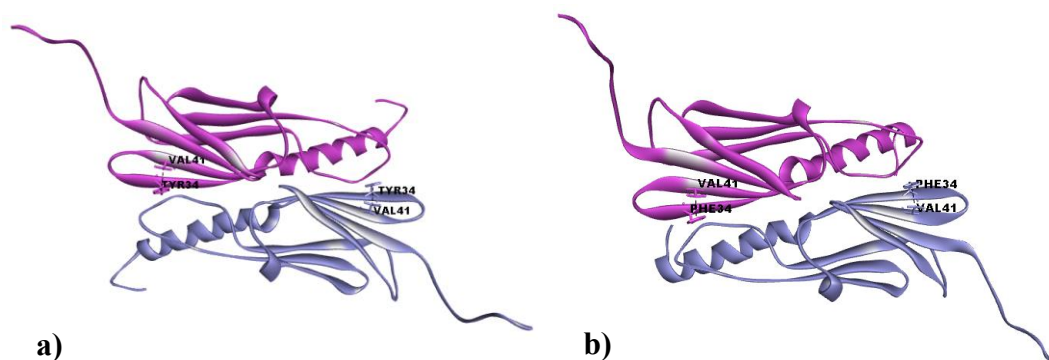
**Figure 3.113. Intermolecular hydrogen bonds of *Tpv* sHSP14.3 WT (a) and L33S mutant (b).** Pink indicating the one monomer and blue indicating the partner monomer.

### 3.17.2 *Tpv* sHSP14.3 Y34F Mutant Variant

Introduction of the highly hydrophobic phenylalanine in the place of tyrosine at residue 34 (Y34F mutation) did not change the putative intramolecular hydrogen bond and the intramolecular hydrophobic interactions but the distance of the bonds changed (Appendix G, Table G.2). Residue F34 or Y34 on  $\beta$ 2 strand and Val 41 on  $\beta$ 3 strand were involved in both intramolecular hydrogen and hydrophobic interactions (Figure 3.114-3.115).



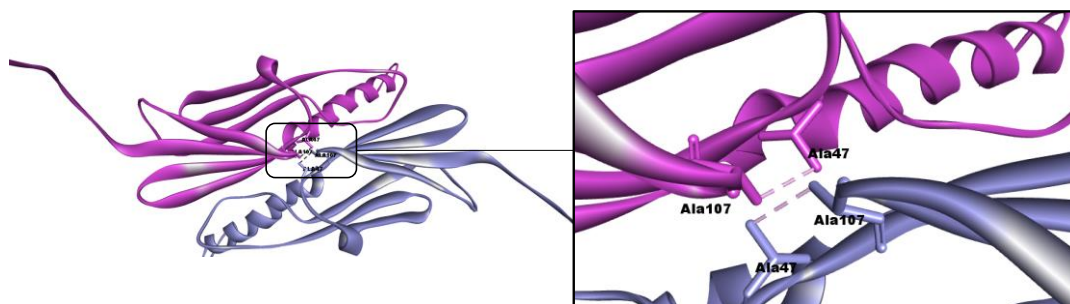
**Figure 3.114. Intramolecular hydrogen bonds of *Tpv* sHSP14.3 WT (a) and Y34F mutant (b).** Pink indicating the one monomer and blue indicating the partner monomer.



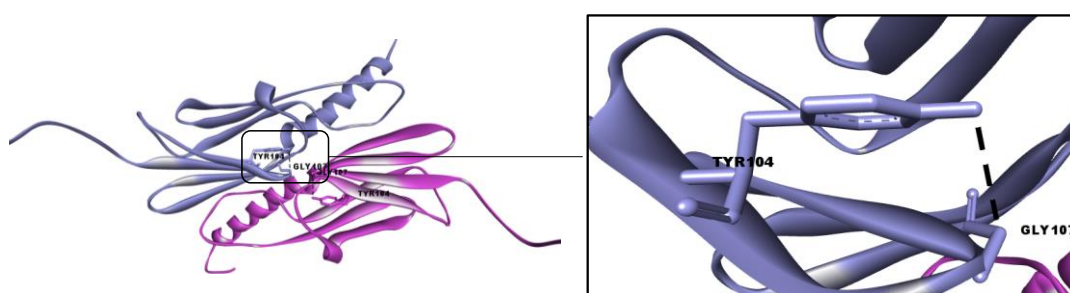
**Figure 3.115. Intramolecular hydrophobic bonds of *Tpv* sHSP14.3 WT (a) and Y34F mutant (b).** Pink indicating the one monomer and blue indicating the partner monomer.

### 3.17.3 *Tpv* sHSP14.3 G107A Mutant Variant

Substitution of glycine by hydrophobic amino acid alanine (G107A mutation) resulted in the formation of an intramolecular hydrophobic interactions between the Ala47 and Ala107, possibly as a result of increased hydrophathy at that site (Figure 3.116). But the intramolecular hydrogen bonds between the Gly107 and Tyr104, which was present in the wild type sHSP were disappeared by this mutation (Figure 3.117, Appendix G, Table G.3).



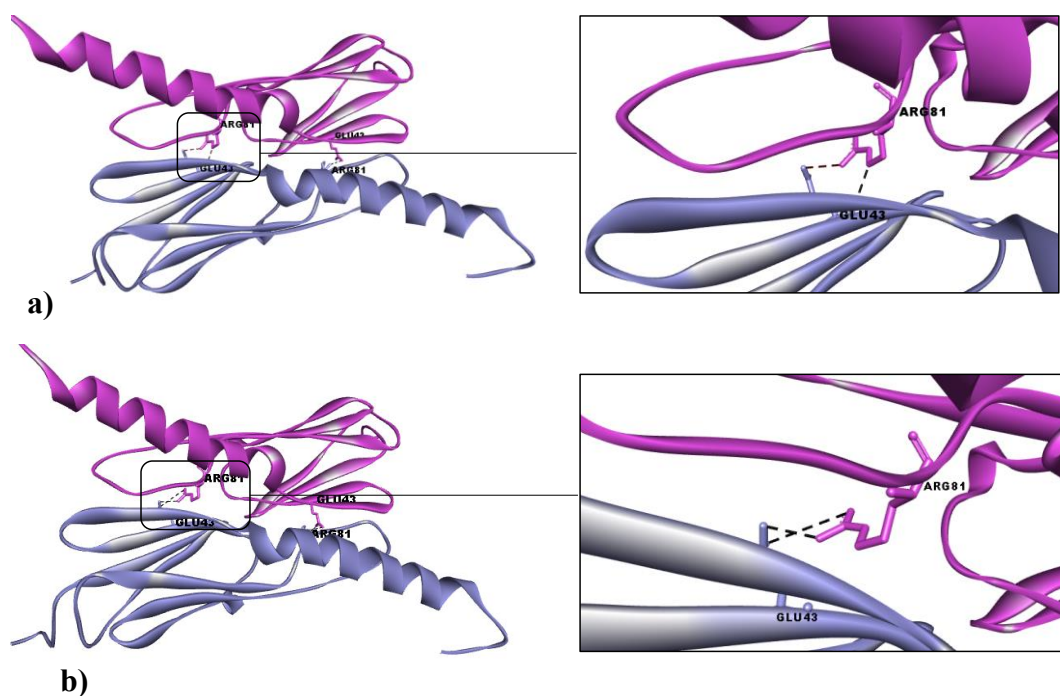
**Figure 3.116. Intramolecular hydrophobic interactions in G107A mutant.** Pink indicating the one monomer and blue indicating the partner monomer.



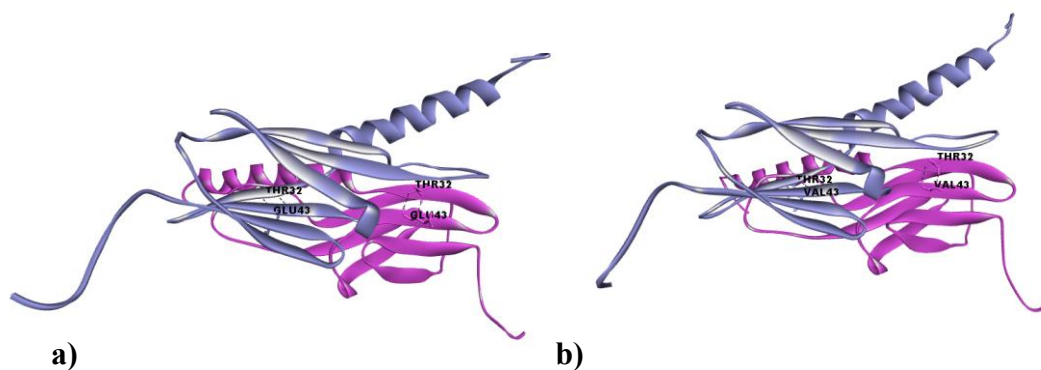
**Figure 3.117. Intramolecular hydrogen bond in the WT at position Gly107.** Pink indicating the one monomer and blue indicating the partner monomer.

### 3.17.4 *Tpv* sHSP14.3 E43V Mutant Variant

Loss negative charge at position 43 by E43V mutation resulted in the elimination of the intermolecular hydrogen bond and intermolecular electrostatic interactions with Arg81, which are present in the WT sHSP (Figure 3.118). These interactions could have play role for stabilizing the dimerization loop  $\beta_6$ , since Arg81 on  $\beta_6$  strand interacted with the Glu43 on the  $\beta_3$  strand, and might have role in dimerization. The residues Glu43 of WT and Val43 of mutant which are on  $\beta_3$  strand of the ACD make same intramolecular hydrogen bond with the nitrogen atom of Thr32 which is on the  $\beta_2$  strand (Figure 3.119). This bond between two adjacent  $\beta$ -strands can be important for maintaining the monomer stability



**Figure 3.118. Intermolecular hydrogen (a) and electrostatic (b) bonds of *Tpv* sHSP14.3 WT at residue Glu43.** Pink indicating the one monomer and blue indicating the partner monomer.

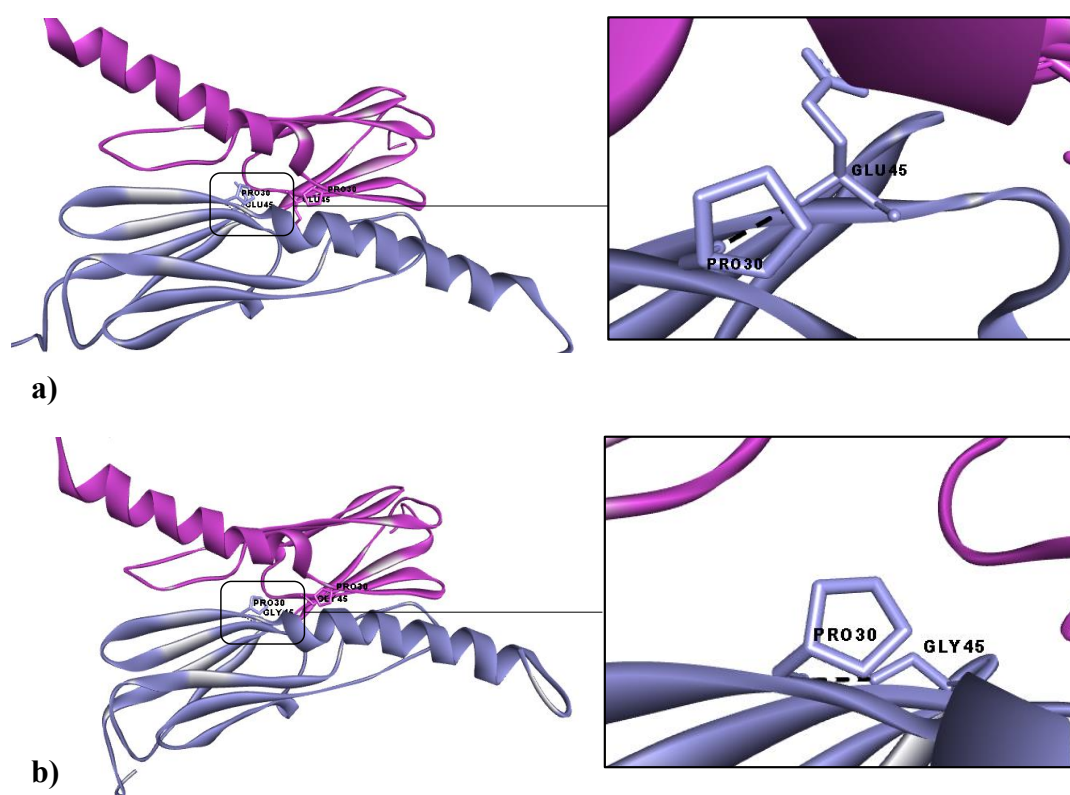


**Figure 3.119. Intramolecular hydrogen bonds of *Tpv* sHSP14.3 WT (a) and E43V mutant (b).** Pink indicating the one monomer and blue indicating the partner monomer.

### 3.17.5 *Tpv* sHSP14.3 E45G Mutant Variant

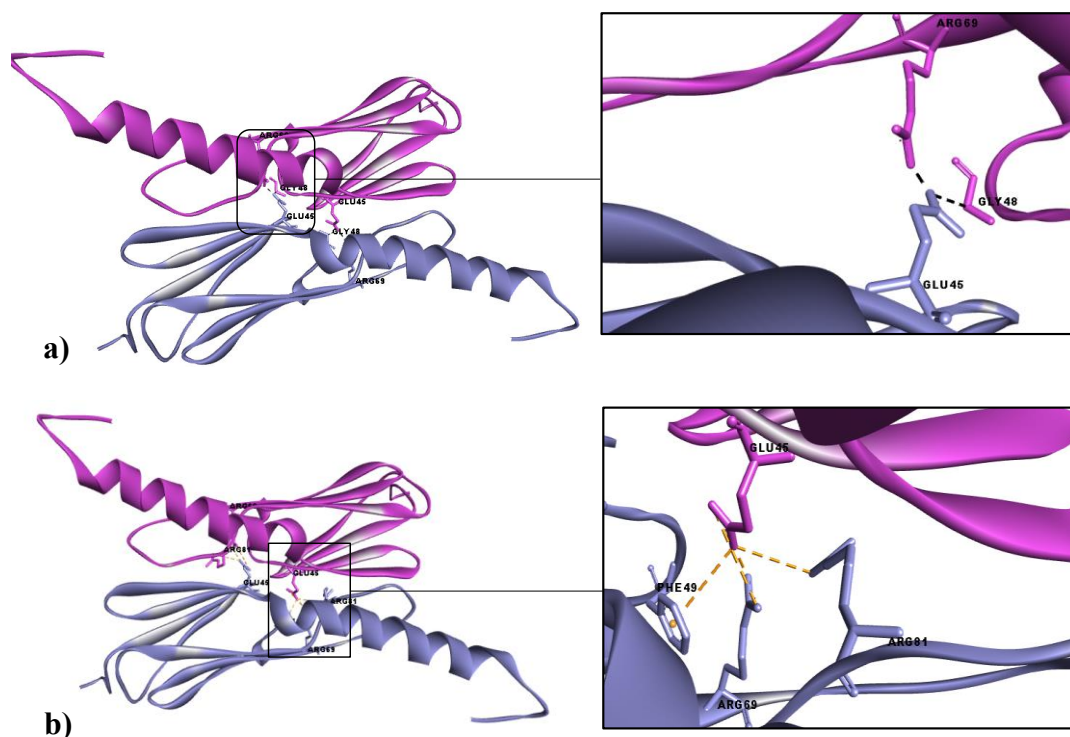
Loss of the negatively charge at position 45 (on  $\beta$ 3 strand) by E45G mutation did not change the intramolecular hydrogen bonding capacity with Pro30 on the  $\beta$ 2 strand (Figure 3.120). However, this mutation resulted in disappearance of intermolecular

salt bridges (*i.e.*, E45-R69 in WT) and intermolecular hydrogen bonds (R69-E45 and G48-E45 in WT), and attractive charge interaction (E45-R81 in WT) (Figure 3.121, Appendix G, Table G.5). The absence of above-mentioned electrostatic interactions mainly salt bridges between  $\beta 3$  and  $\beta 6$  strands would have impact on dimerization and/or its stability in the mutant sHSP. Besides, inter molecular pi-anion bond formed by interaction of negatively charged Glu45 and aromatic Phe 49 also disappeared after E45G mutation (Figure 3.121).



**Figure 3.120.** Intramolecular hydrogen bonds of *Tpv* sHSP14.3 WT (a) and E45G mutant (b). Pink indicating the one monomer and blue indicating the partner monomer.

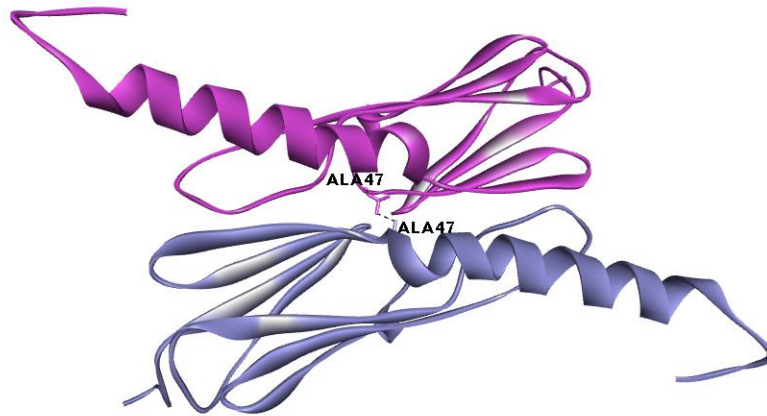




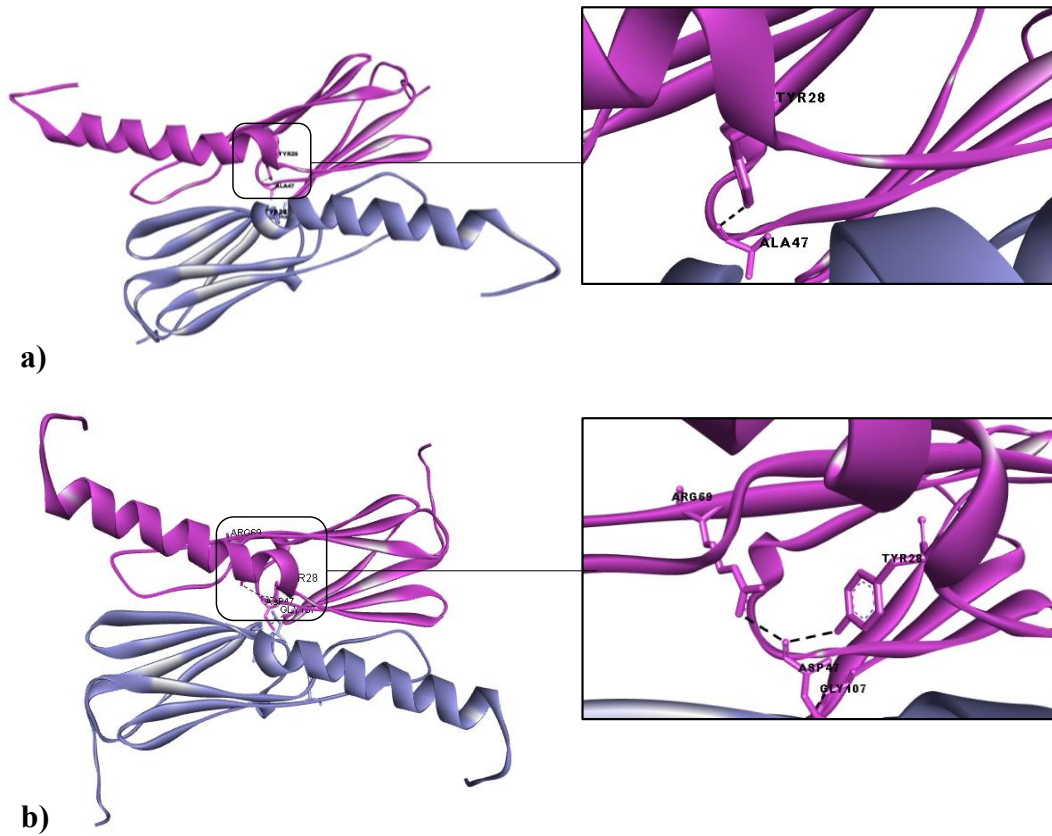
**Figure 3.121. Intermolecular hydrogen (a) and electrostatic (b) bonds of *Tpv* sHSP14.3 WT at residue Glu45.** Pink indicating the one monomer and blue indicating the partner monomer.

### 3.17.6 *Tpv* sHSP14.3 A47D Mutant Variant

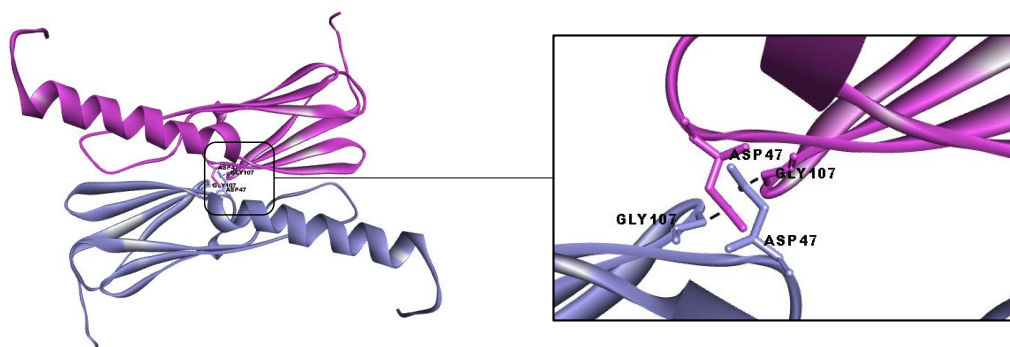
The removal of hydrophobic residue and instead introduction of negatively charged residue by A47D mutation resulted in loss of the intermolecular hydrophobic interactions between Ala47-Ala47 (Figure 3.122). The residue 47 was positioned at L34 loop between the  $\beta$ 3 and  $\beta$ 4 strands. This intermolecular hydrophobic bond present in the WT could contribute to the stability of the dimer. On the other hand, this mutation increased the number of intramolecular (newly formed Arg69-Asp47; Asp47-Gly107) hydrogen bonds from 1 to 3, besides introducing a new intermolecular H-bond (Gly107-Asp47) (Figure 3.123-3.124, Appendix G, Table G.6). Arg69 and Gly107 were found at the  $\beta$ 6 strand and  $\beta$ 8- $\beta$ 9 loop of the *Tpv* sHSP14.3 protein, respectively. Therefore, hydrogen bonds between the Asp47-Arg69 and Asp47-Gly107 can be important for stability of the dimeric structure.



**Figure 3.122.** Intermolecular hydrophobic bonds of *Tpv* sHSP14.3 WT at residue Ala47. Pink indicating the one monomer and blue indicating the partner monomer.



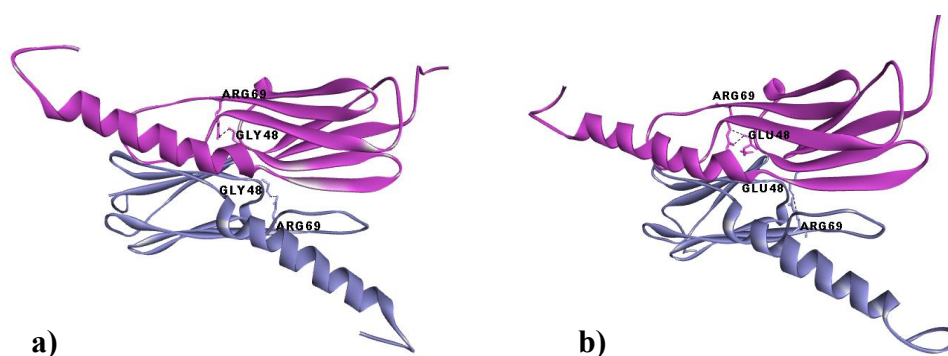
**Figure 3.123.** Intramolecular hydrogen bonds of *Tpv* sHSP14.3 WT (a) and A47D mutant (b). Pink indicating the one monomer and blue indicating the partner monomer.



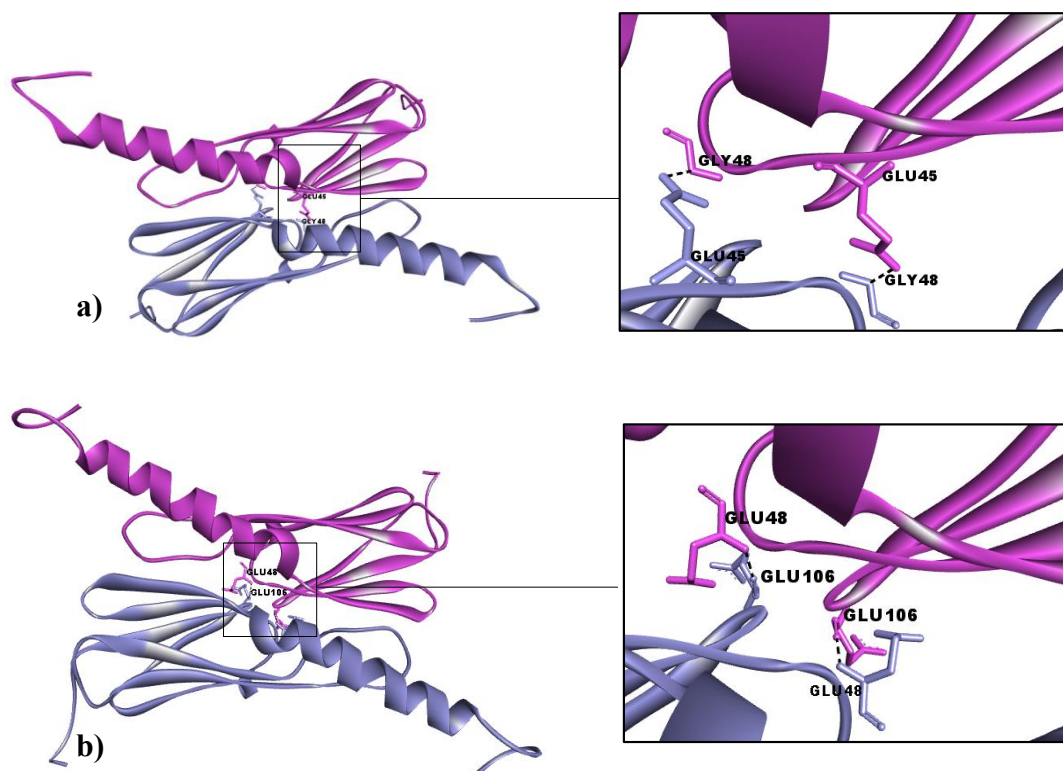
**Figure 3.124. Inter-molecular hydrogen bonds of *Tpv* sHSP14.3 A47D mutant at residue Asp47.** Pink indicating the one monomer and blue indicating the partner monomer.

### 3.17.7 *Tpv* sHSP14.3 G48E Mutant Variant

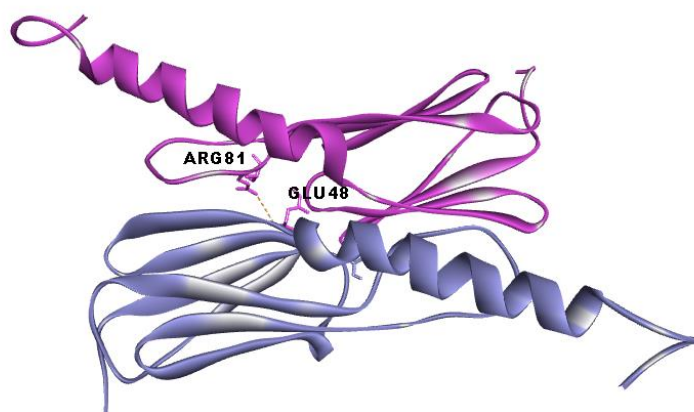
G48 residue is found in highly conserved P/A-G doublet of the L34 loop of the non-metazoan sHSPs. This residue in *Tpv* sHSP14.3 WT makes an intramolecular hydrogen bond with Arg69. Same interaction was also observed for G48E mutant protein, but with an additional carbon hydrogen bond between same residues (*i.e.*, R69-E48) that were not found in the WT (Figure 3.125). The distance between these bonds was increased as a result of the mutation (Appendix G, Table G.7). Also, intermolecular H-bonding between Gly48-Glu45 in the wild type is exchanged by the Glu48-Glu106 hydrogen bond in the mutant strain (Figure 3.126). Furthermore, by introduction of the charged residue at position 48 which was previously neutral resulted in the construction of an intramolecular electrostatic interaction between Arg81 and Glu48 in mutant strain (Figure 3.127).



**Figure 3.125. Intramolecular hydrogen bonds of *Tpv* sHSP14.3 WT (a) and G48E mutant (b).** Pink indicating the one monomer and blue indicating the partner monomer.



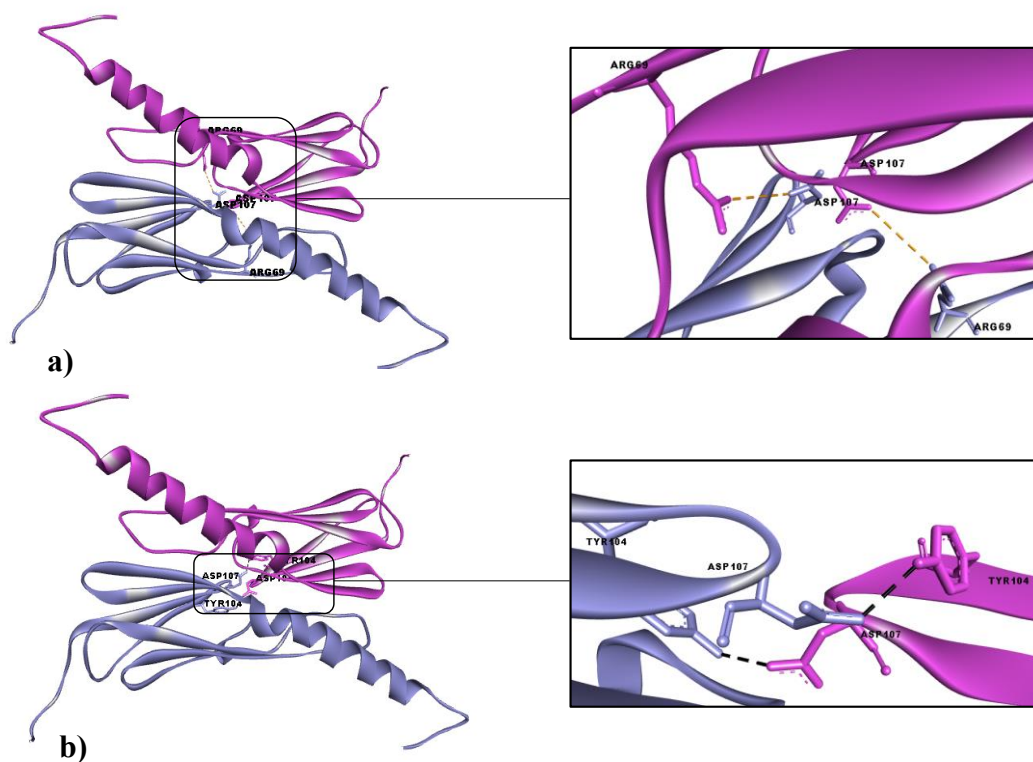
**Figure 3.126. Intermolecular hydrogen bonds of *Tpv* sHSP14.3 WT (a) and G48E mutant (b).** Pink indicating the one monomer and blue indicating the partner monomer.



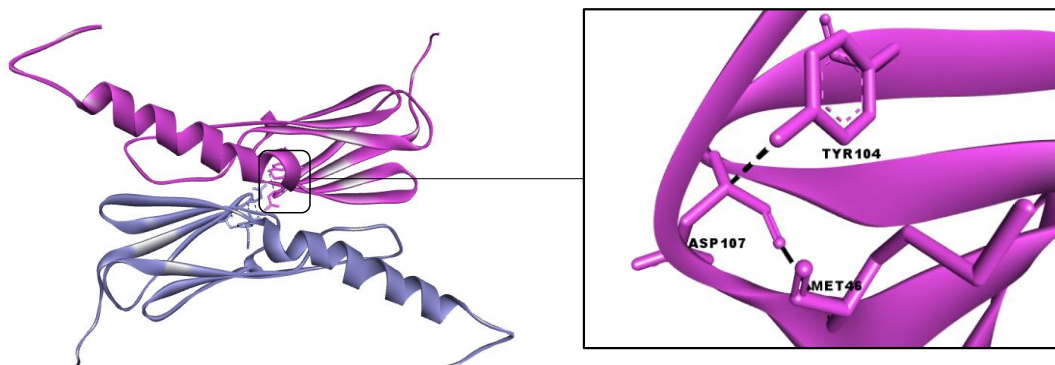
**Figure 3.127. Intramolecular electrostatic bonds of *Tpv* sHSP14.3 G48E mutant at residue Glu48.** Pink indicating the one monomer and blue indicating the partner monomer.

### 3.17.8 *Tpv* sHSP14.3 G107D Mutant Variant

Introduction of the negative charged residue at the position 107 (G107D) resulted in the formation of the intermolecular electrostatic interactions (Arg69-Asp107) and hydrogen bonds (Tyr104-Asp107), which are not available in the wild type (Figure 3.128). These inter-molecular interactions could contribute to the integrity of the dimer structure. Also, the number of the intramolecular hydrogen bonds increased by G107D mutation with formation of Met46-Asp107 bond (Figure 3.129, Appendix G, Table G.8). The interacting residues with the Asp107 were located at  $\beta$ 6 strand (Arg69), at  $\beta$ 3- $\beta$ 4 loop (Met46) and at  $\beta$ 8 strand (Tyr104) of the G107D mutant protein.



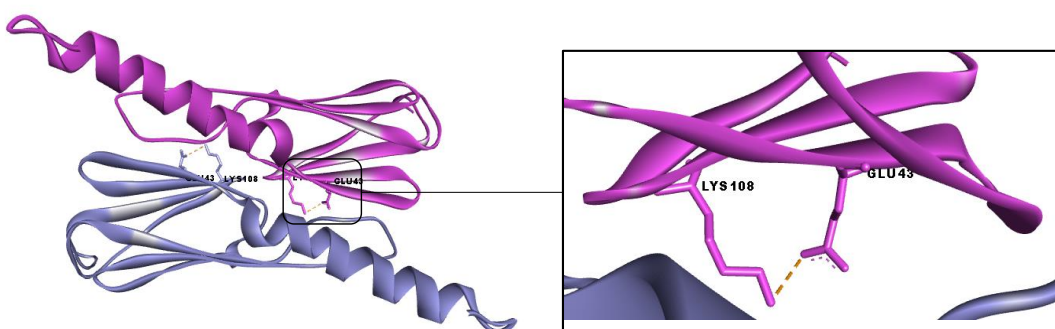
**Figure 3.128. Intermolecular electrostatic (a) and hydrogen (b) bonds of *Tpv* sHSP14.3 G107D mutant at residue Asp107.** Pink indicating the one monomer and blue indicating the partner monomer.



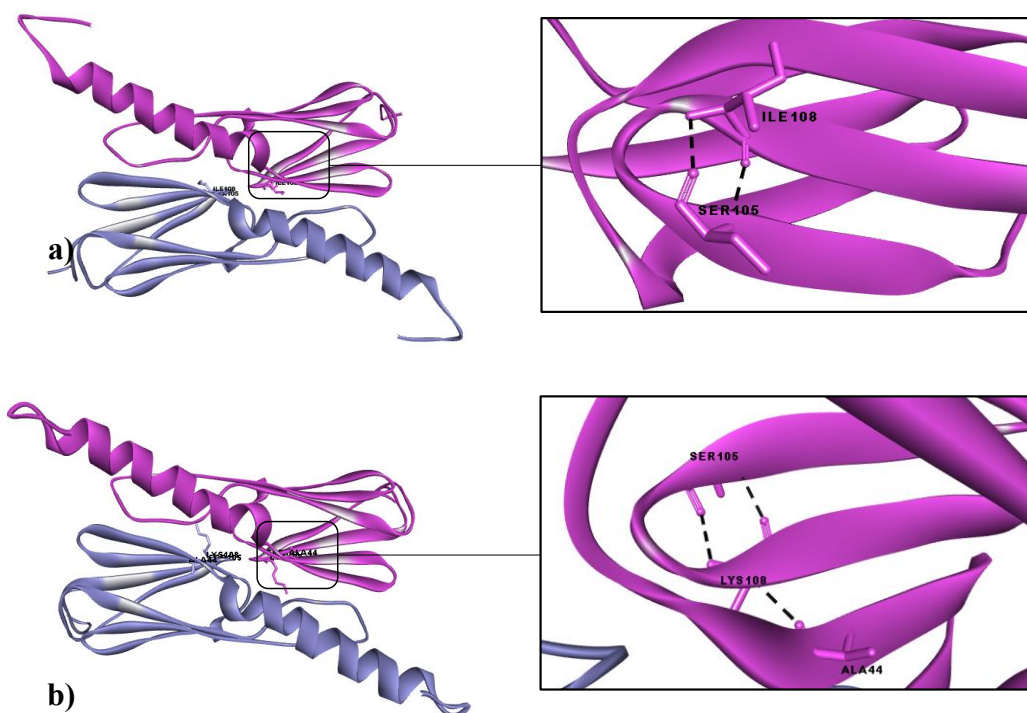
**Figure 3.129. Intramolecular hydrogen bonds in G107D mutant.** Pink indicating the one monomer and blue indicating the partner monomer.

### 3.17.9 *Tpv* sHSP14.3 I108K Mutant Variant

Introduction of the charged residue lysine in the place of hydrophobic amino acid isoleucine at position 108 (I108K mutation) resulted in the formation of the intramolecular electrostatic interactions between Lys108 and Glu43 (Figure 3.130). Also, this mutation led to the appearance of the new intramolecular hydrogen bonds (*i.e.*, Lys108-Ala44), besides already existing H-bonds of the residue 108 with Ser105 (Figure 3.131, Appendix G, Table G.9).



**Figure 3.130. Intramolecular electrostatic bonds of *Tpv* sHSP14.3 I108K at residue Lys108.** Pink indicating the one monomer and blue indicating the partner monomer.



**Figure 3.131. Intramolecular hydrogen bonds of *Tpv* sHSP14.3 WT (a) and I108K mutant (b).** Pink indicating the one monomer and blue indicating the partner monomer.

### 3.17.10 *Tpv* sHSP14.3 Y34FG48E Double Mutant Variant

Double mutation Y34FG48E like G48E single mutation but different from WT led to formation of a new intramolecular electrostatic interaction (between Arg69-Glu48) (Figure 3.132). Like Y34F single mutant, intramolecular H-bonds and hydrophobic interactions remained same in the double mutant and in the WT-sHSP. Also, intermolecular hydrogen bonds present in G48E single mutant was also found for the double mutant (Figure 3.133-3.135) (Appendix G, Table G.10).

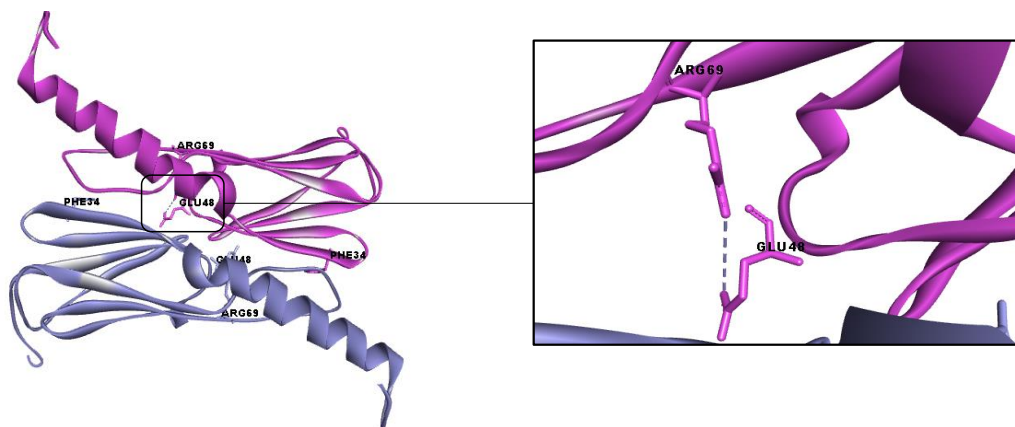


Figure 3.132. Intramolecular electrostatic interactions in the double mutant Y34FG48E. Pink indicating the one monomer and blue indicating the partner monomer.

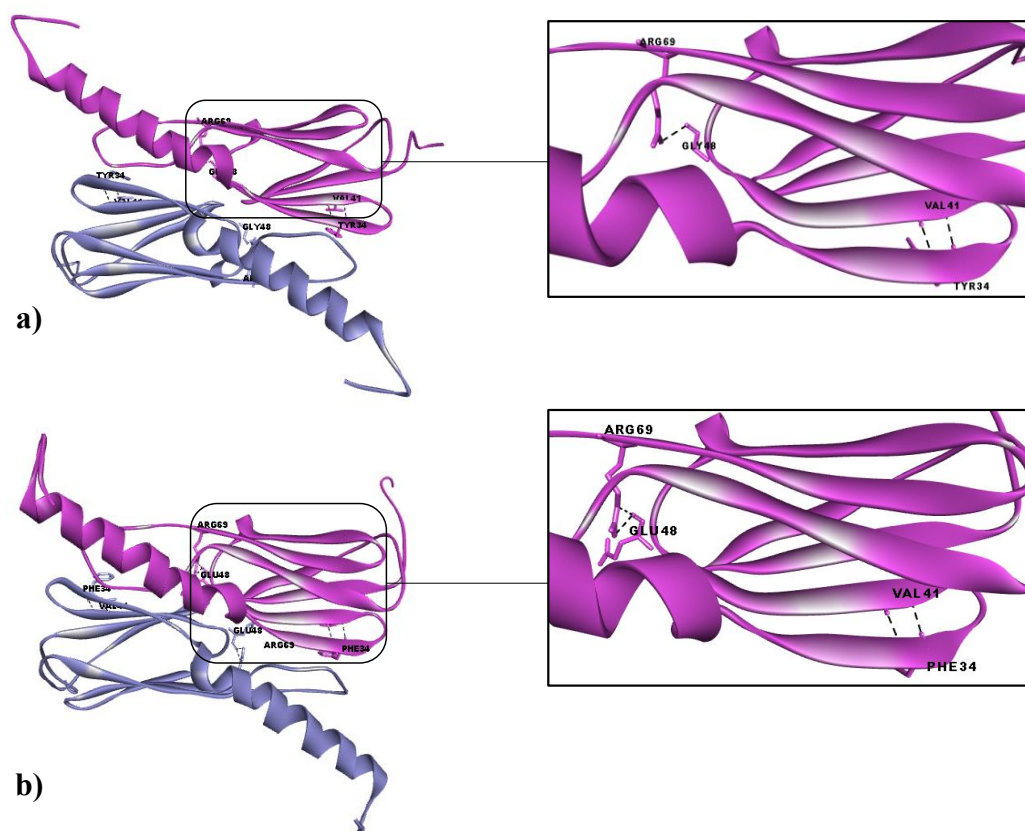
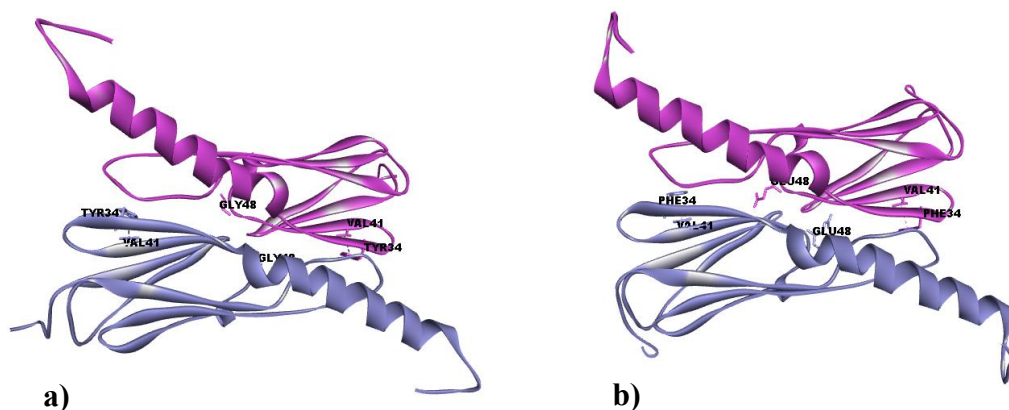
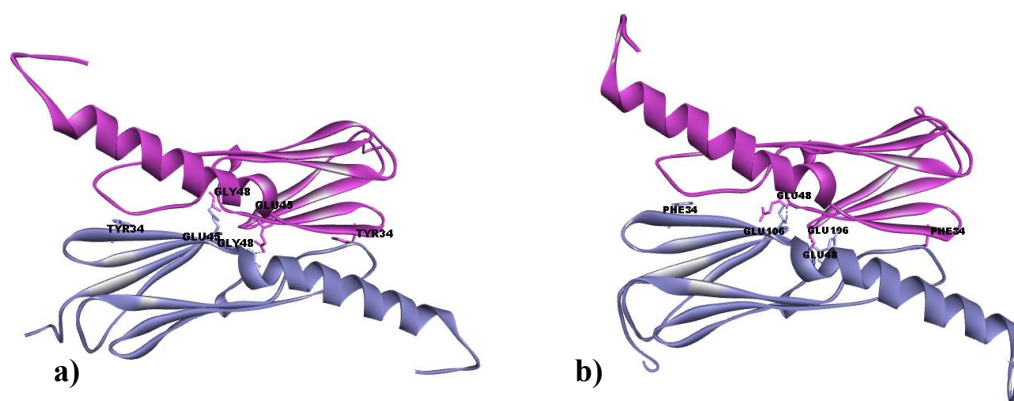


Figure 3.133. Intramolecular hydrogen bonds in the *Tpv* sHSP14.3 WT (a) and Y34FG48E double mutant (b). Pink indicating the one monomer and blue indicating the partner monomer.





**Figure 3.134.** Intramolecular hydrophobic bonds in the *Tpv* sHSP14.3 WT (a) and Y34FG48E double mutant (b). Pink indicating the one monomer and blue indicating the partner monomer.

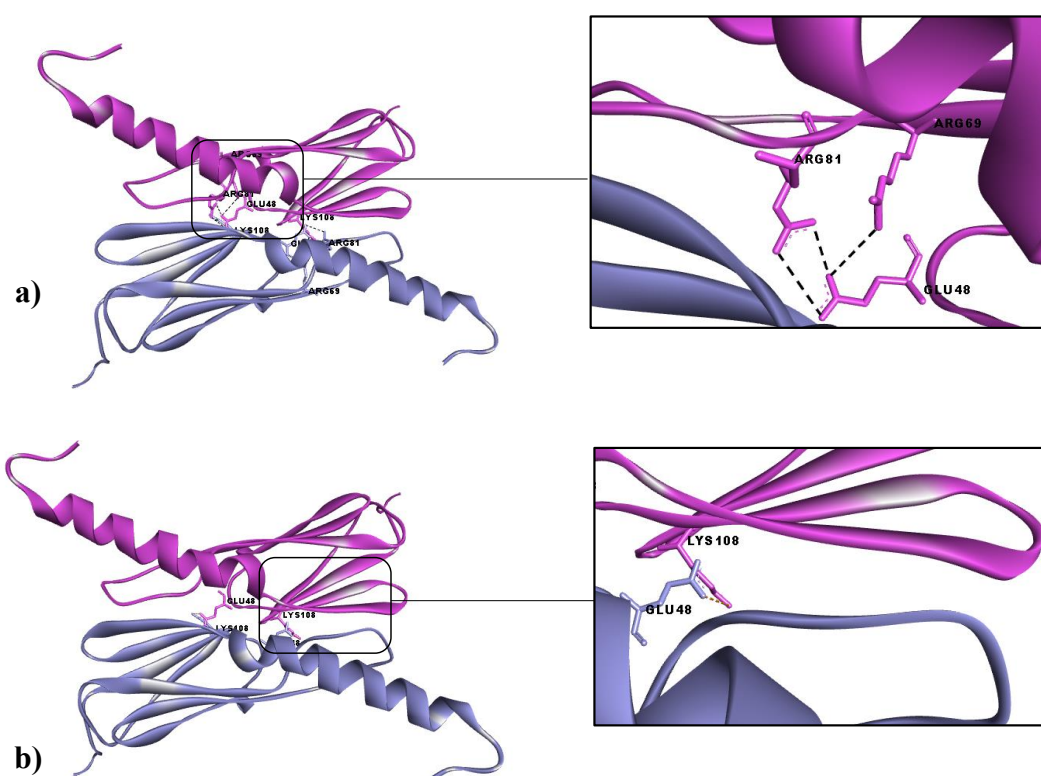


**Figure 3.135.** Intermolecular hydrogen bonds in the *Tpv* sHSP14.3 WT (a) and Y34FG48E double mutant (b). Pink indicating the one monomer and blue indicating the partner monomer.

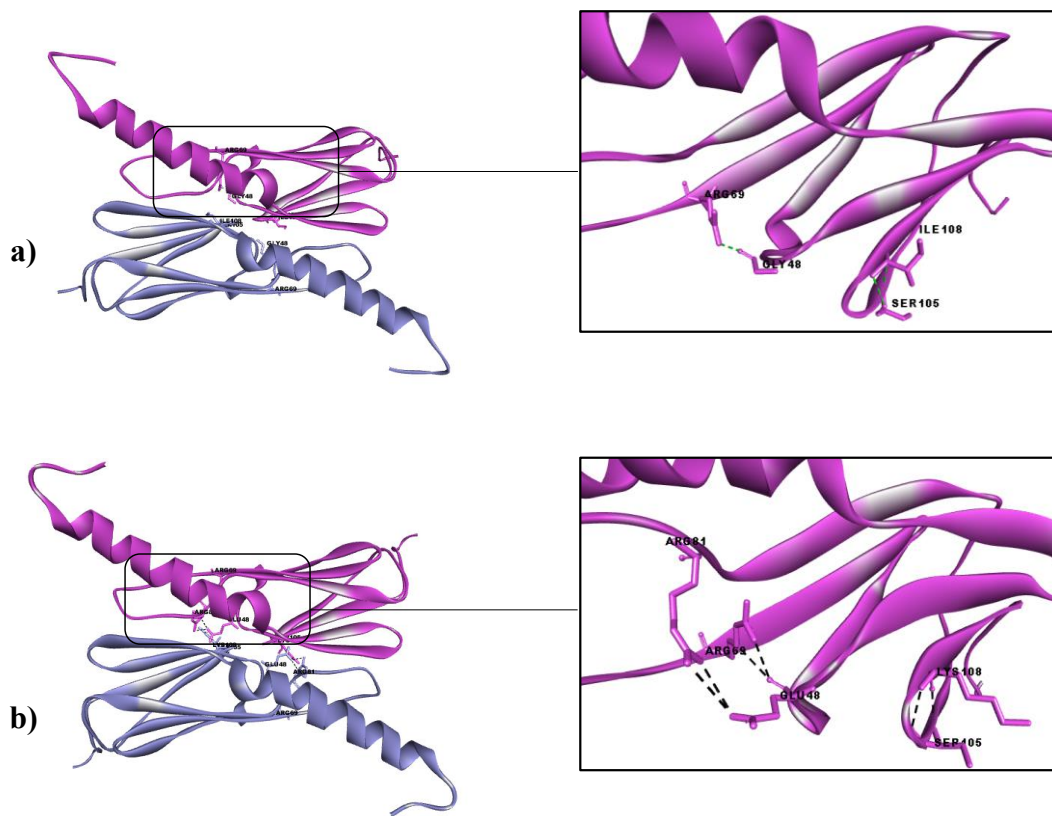
### 3.17.11 *Tpv* sHSP14.3 G48E1108K Double Mutant Variant

The double mutation G48E1108K, is combining two single mutations that introduce charged residues in place of neutral/hydrophobic residues. These mutations resulted in gain of new intra/intermolecular electrostatic interactions that are not present in the WT. Set of those interactions in both subunits include intramolecular Arg69:NH1-Glu48:OE1, Arg81:NH1-Glu48:OE1, Arg81:NH2-Glu48:OE2 and intermolecular Lys108:NZ-Glu48:OE1 bonds (Figure 3.136). Among these interactions, salt bridges between Arg81-Glu48 (intramolecular) and Lys108-Glu48

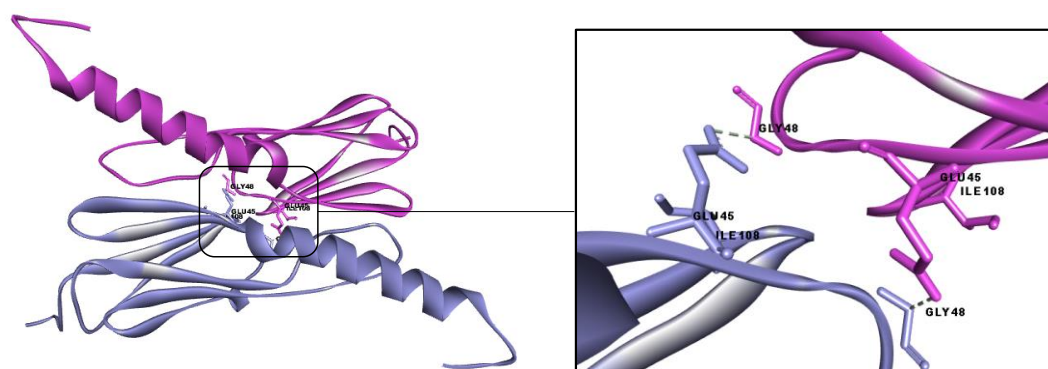
(intermolecular, attractive charge) were not observed in the related single mutant variants (*i.e.*, G48E, and I108K). Furthermore, unlike WT protein structure, double mutation increased the number of intramolecular hydrogen bonding between Arg69-Glu48 and Arg81-Glu48 (Figure 3.137). On the other hand, the intermolecular H-bond between Gly48 and Glu45 present in WT were disappeared in double mutant protein (Figure 3.138, Appendix G, Table G.11).



**Figure 3.136. Intramolecular (a) and intermolecular (b) electrostatic bonds in G48E108K double mutant. Pink indicating the one monomer and blue indicating the partner monomer.**



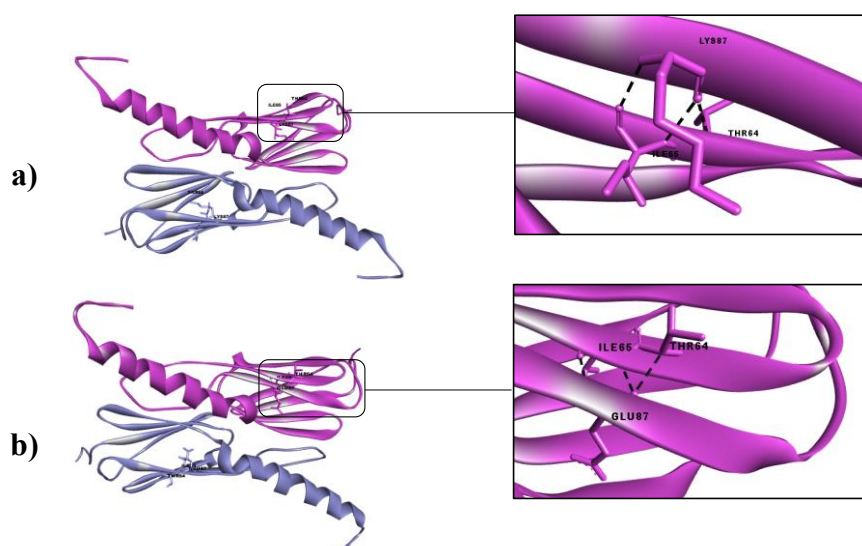
**Figure 3.137. Intramolecular hydrogen bond in WT (a) and G48E1108K (b).** Pink indicating the one monomer and blue indicating the partner monomer.



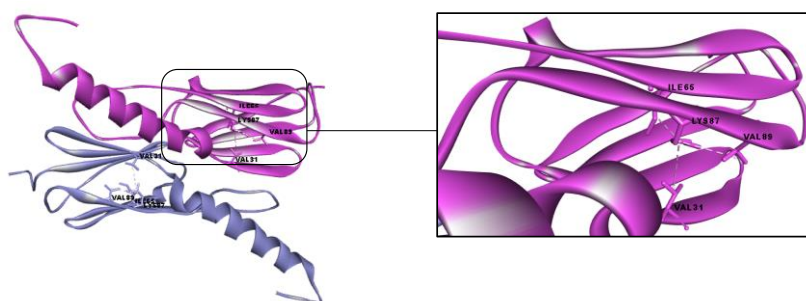
**Figure 3.138. Intermolecular hydrogen bond in WT at position Gly48.** Pink indicating the one monomer and blue indicating the partner monomer.

### 3.17.12 *Tpv* sHSP14.3 K87E Mutant Variant

Another mutation that caused the exchange of positive charge by negative ones at position 87 was K87E mutation. This reversal of charge at the position 87 (on the  $\beta 7$  strand) did not change the hydrogen bonds, but eliminated the hydrophobic interactions (Lys87-Ile65, Lys87-Val31, and Lys87-Val89) within the same subunits (Figure 3.139-3.140, Appendix G, Table G12). Those hydrophobic interactions present in the WT can be important for the structural integrity of the *Tpv* sHSP14.3 monomer, because they are involved extensive interactions between  $\beta$  strands of the ACD *i.e.*,  $\beta 2$  (Val31),  $\beta 5$  (Ile65) and  $\beta 7$  (Val89).



**Figure 3.139.** Intramolecular hydrogen bond in WT (a) and K87E mutant (b). Pink indicating the one monomer and blue indicating the partner monomer.

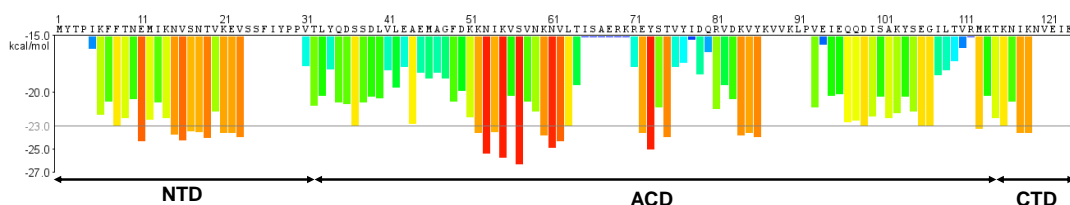


**Figure 3.140.** Intramolecular hydrophobic bond in the WT at position Lys87. Pink indicating the one monomer and blue indicating the partner monomer.

### 3.18 Prediction of the Fibril Forming Units

Zipper DB database was used for the predictions of fibril forming segments within sHSP proteins determined by the 3-D Profile method. Fibrillation propensity profile depends on the energy levels of hexapeptides whose energy threshold equal or below -23 kcal/mol are associated with the high fibrillation propensity. In addition to this, the shape complementarity (steric zipper interface) and area of interface (the area that is accessible to the solvent for each hexapeptide) was also calculated for each hexapeptide, before and after the mutation. Depending on this, the program gives a composite score (C-Score) to each hexapeptide, which determines the overall propensity of that hexa peptide towards the fibril formation ability, *i.e.*, higher numerical scores are associated with the increasing attitude of the hexapeptide to form fibril like structures.

Rosetta energy diagram of the WT *Tpv* sHSP14.3 showed that the ACD of the protein have higher tendency to form  $\beta$ -fibrils while the unstructured flanking regions NTD and CTD have less propensity to form fibrils (Figure 3.141).



**Figure 3.141. Rosetta Energy bar chart of *Tpv* sHSP 14.3 WT.** One hexapeptide is colored depending on its Rosetta energy values blue to red. Gray line representing the threshold of -23 kcal/mol.

Rosetta energy profile of the L33S mutant showed that introduced mutation slightly enhanced the fibrillation propensity at the segment of the protein (*e.g.*, SYQDSS). Although this segment did not pass the threshold level of -23 kcal/mol, both rosetta energy levels and C-scores (numerical value) were higher than the wild type (Figure 3.142, Appendix H, Table H.1). Other two segments in which the L33 is involved, show slight increase in Rosetta Energy following the replacement of Leu33 with hydrophilic residue Ser. On the other hand, it was noticed that there was an increase in the area of interface. All of the three hexapeptides which include S33, show an

increase in area of interface by a value of 40 Å<sup>2</sup> or more. This change led to a considerable change in the overall score, *i.e.*, the C-score of each segment, thus increasing the fibril formation propensity of the mutant protein.

In a mutation, where hydrophobicity is increased at position 34, by replacement of Y34 by highly hydrophobic Phe also resulted in increase in the area of interface around the position of mutation (Appendix H, Table H.1). The three hexa peptides including the position 34 showed increase in area of interface by 60 Å<sup>2</sup>. This increase in the area accessible to solvent directly affected the C-score which also increased (numerical value) while Rosetta Energy of three of the four hexa peptides decreased (numerical value) (Figure 3.142). The changes in the shape complementarity between the WT and Y34F mutant are not pronounced, where the highest change recorded is for the hexa peptide <sub>32</sub>TLYQDS<sub>37</sub>, that shows an increase in shape complementarity by a value of 0.14. The overall C-score of this hexapeptide that shows the maximum increase in the numerical value from 32.4 to 41.2. Besides, higher values of shape complementarity predict the hexapeptide to not to form fibrils. (Appendix H, Table H.1).

Replacing the charged residues at position 43 and 45 with hydrophobic one, in separate single mutant proteins, resulted in increase in the numerical values of Rosetta Energy levels of all the segments that involve these positions (Figure 3.142-3.143). This increase in Rosetta Energy is high for the E43V as compared E45G, and this might be due to the highly hydrophobic nature of the residue, valine. First five hexa peptides of E43V showed a drastic increase in Rosetta Energy (numerical value) (around 3 Kcal/mol), and two of them (<sub>39</sub>DLVLEA<sub>44</sub> and <sub>40</sub>LVLEAE<sub>45</sub>) passed the threshold of -23 kcal/mol (Appendix H, Table H.1). Therefore, they have high propensity to form fibril like structure. Besides, the fifth hexa peptide <sub>44</sub>AEMAGF<sub>49</sub>, all other hexapeptides of E45G showed a considerable increase in Rosetta Energy values. However, none of hexapeptides of this mutant possessed the energy to pass the threshold (Figure 3.143). There is a considerable increase in the area of interface in majority of the hexa peptides of both of the mutants as well as in their C-scores as compared to that of the WT. The highest change is noted for peptide

<sup>42</sup>LEAEMA<sub>47</sub> (WT sequence) which had a C-score of 0.0 in WT and 43.3 and 41.3 in the E43V and E45G mutant proteins, respectively. Another notable increase is for the hexapeptide VLEAEM in E43V, which is around 41.2 scores (see Table H.1 in Appendix H).

Introduction of hydrophilicity at position 47, A47D, resulted in considerable decrease in Rosetta Energy levels (numerical values) of all hexapeptide segments that include this position, except the last segment, that showed a slight increase of 0.3 kcal/mol, still did not pass the threshold. Similar results of Rosetta Energy were seen when negative charge was introduced at position 48 in place of the small hydrophobic amino acid glycine (G48E mutation). Besides, the hexapeptide <sup>45</sup>EMAGFD<sub>50</sub> (WT) that shows a slight increase in Rosetta Energy, all other segments showed decrease in Rosetta Energy (Figure 3.143). Not only this, also both mutations led to a noticeable increase in the solvent accessible area, except <sup>45</sup>EMAEFD<sub>50</sub> of G48E mutant. Although this increase in area of interface and decrease in Rosetta Energy, that leads to the protein in the favour of decreased fibril formation, both the mutants show an increase in C-score. This increase in C-score is highest (38 score) for the first hexapeptide segment that include A47D mutation point (Appendix H, Table H.1).

The two mutations, at position 107, where a highly conserved glycine is replaced with a hydrophobic alanine (G107A) and a charged residue, aspartic acid (G107D), both resulted in alteration of the Rosetta Energy of the hexapeptides near the mutation site. Introduction of alanine (G107A mutation) increased the Rosetta Energy of all the hexapeptides and three of them exceeded the threshold -23kcal/mol (Figure 3.143, and Table H.2 in Appendix H). The C-score of all segments also show an increase in value where the highest change (by 10 score) was recorded for <sup>105</sup>SEGILT<sub>110</sub> (in WT) segment. It also showed the highest increase in Rosetta energy and in area of interface. On the other hand, when charged residue (Asp) is introduced at position 107, the Rosetta Energy values are lowered for all hexapeptide segments except one (<sup>103</sup>KYSEGI<sub>108</sub>). In addition to this, the C-score also gives a compensating result, where three segments exhibit slight increase and the other three, slight

decrease in score, as compared to the WT. These results predict, that increasing hydrophobicity from glycine to alanine (G107A) might increase the propensity to form fibril formation, while the loss of hydrophobicity may reduce tendency for fibrillation.

Similar to the results of G107D, introducing charge at the position of 108 (I108K) resulted in decrease in Rosetta Energy of all the hexapeptides of I108K (Figure 3.144). The same hexapeptide, that showed increased in Rosetta Energy in G107D mutation, this time resulted in the maximum decrease in Rosetta Energy (1.4 kcal/mol), as a result of this mutation. Except two hexapeptides ( $_{105}\text{SEGIL}_{110}$  and  $_{108}\text{ILTVRM}_{113}$ ), all hexapeptides showed a decrease in area of interface also, as compare to WT (Appendix H, Table H.2).

Area of interface of these two peptides 0 Å in the WT, while the mutation resulted in such structural changes leading to the increase of the interface area to 78 and 104 in respective hexapeptides. This led to the high increase in the C-score of these hexapeptides, while all other hexapeptides showed a mild increase or decrease in the score. The results of shape complementarity show that the segment  $_{104}\text{YSEGIL}_{109}$  make highest change (0.17).

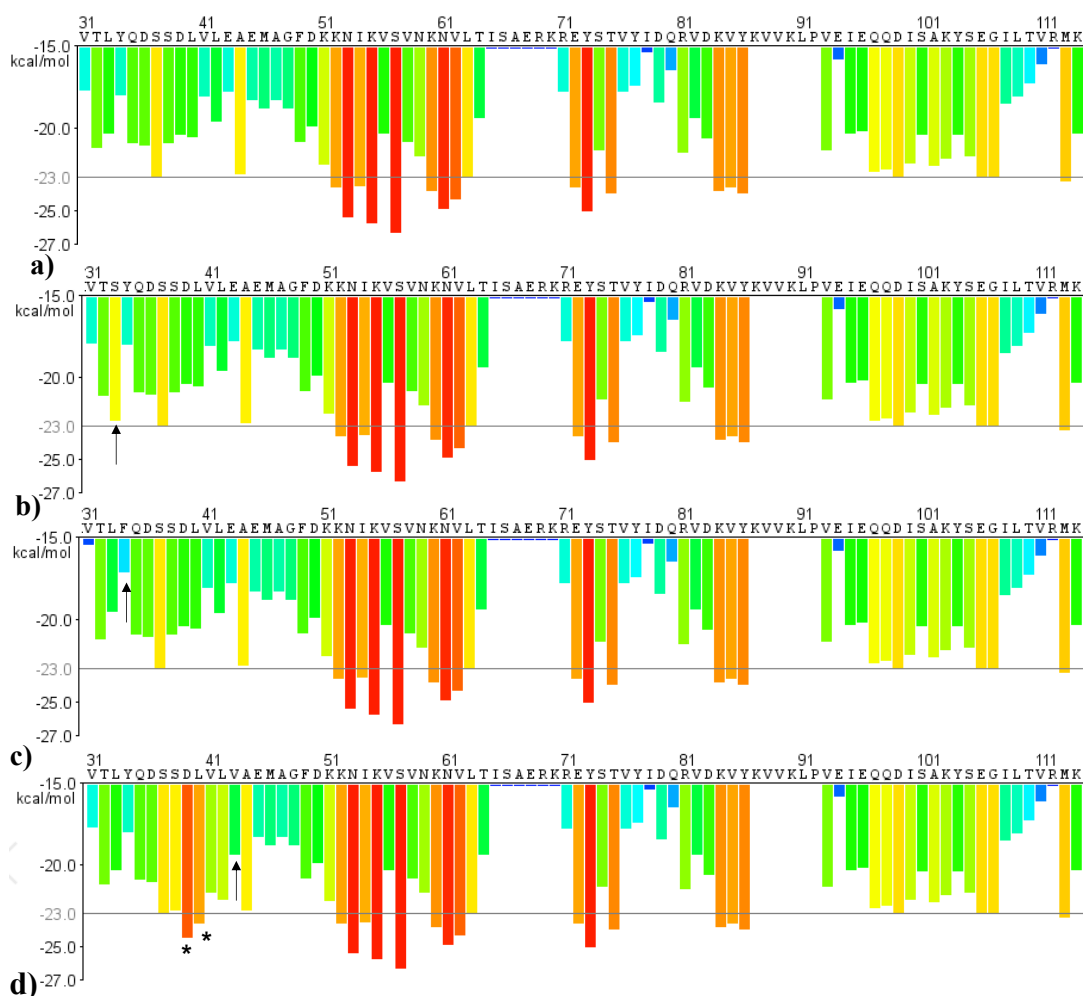
There is a mutation in this study, where reversing of charges takes place at highly conserved position 87 (K87E). The Rosetta Energy of all the hexapeptides, besides one ( $_{85}\text{VYK}_{90}$ ), showed a general trend of decreasing, where the maximum decrease was 1.3 kcal/mol for the segment  $_{83}\text{DKVYK}_{88}$  (Appendix H, Table H.2). However, three hexapeptides with Rosetta energy that passed the threshold still had energy levels below -23kcal/mol (Figure 3.144).

In addition to this, the shape complementarity of all hexapeptides, besides one ( $_{83}\text{DKVYK}_{88}$ ), showed a decrease or remained same following mutation. However, the mutation resulted in such changes in the hexapeptide structures that the area of interface increased for all of them, besides one ( $_{82}\text{VDKVYK}_{83}$ ). The C-score, thus resulted in decrease to 0.0 in the first hexapeptide, a mild decrease in the second one, and an increase (by a score of 4 or more) in the rest of the segments that involve K87.

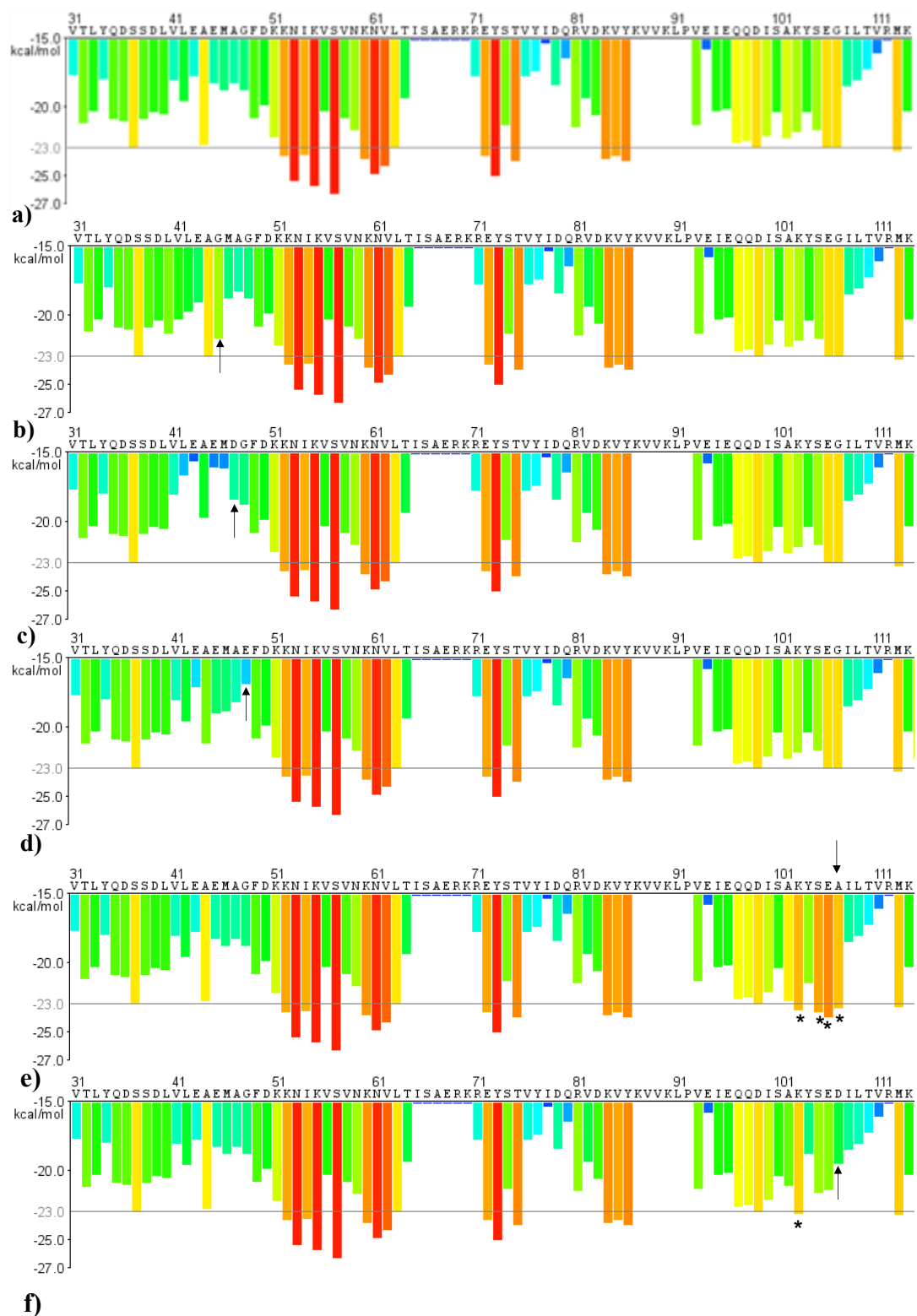


The overall results shows that introduction of charge, in place of hydrophobicity has a general trend of decreasing the Rosetta Energy of the hexapeptide segments, along with increasing the area available to the solvent, not only for the peptide.

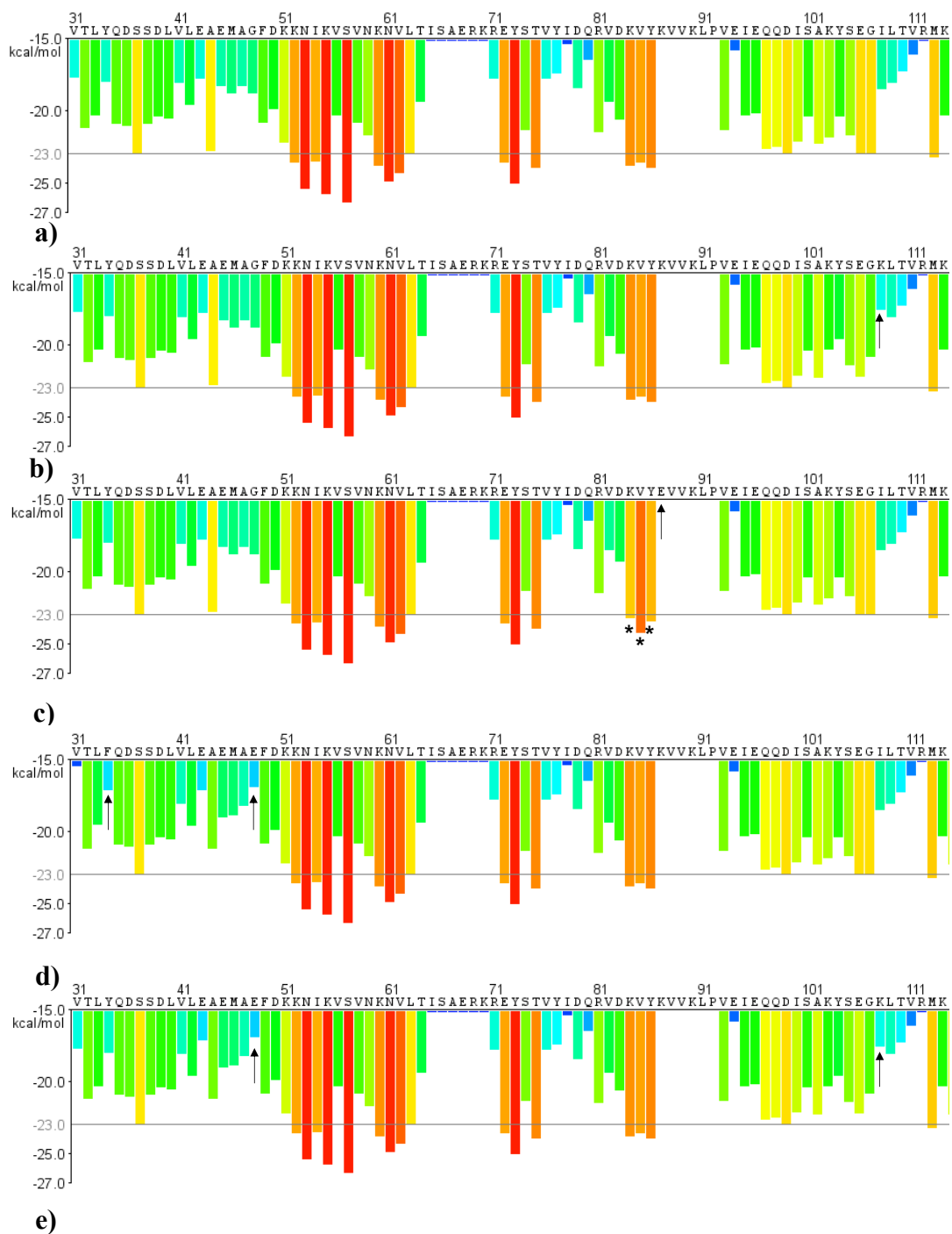
Except insignificant increase in Rosetta Energy for two hexapeptides ( $_{45}$ EMAGFD $_{50}$  and  $_{46}$ MAGFDK $_{51}$ , Y34FG48E and G48E1108K double mutants in general exhibited a decrease in the Rosetta Energy which might have shifted the equilibrium of the ACD, as a whole, towards decreasing fibrillation propensity. Although, these mutants mostly showed increased the area accessible to the solvent, the double mutant G48E1108K caused decrease at five hexapeptides around the second mutation (Figure 3.144, Appendix H, Table H.3).



**Figure 3.142. Rosetta Energy bar chart of the ACDs of *Tpv* sHSP 14.3 WT (a) and L33S (b), Y34F (c), E43V (d).** The black arrow indicated the point of mutation. Asteric indicating the hexapeptide passing the threshold of -23 kcal/mol.



**Figure 3.143.** Rosetta Energy bar chart of the ACDS of *Tpv* sHSP 14.3 WT (a) and E45G(b), A47D (c), G48E (d), G107A (e), G107D (f). The black arrow indicated the point of mutation. Asteric indicating the hexapeptide passing the threshold of -23 kcal/mol.



**Figure 3.144. Rosetta Energy bar chart of the ACDS of *Tpv* sHSP 14.3 WT (a) and I108K (b) K87E (c), Y34FG48E (d), G48E I108K (e). The black arrow indicated the point of mutation. Asteric indicating the hexapeptide passing the threshold of -23 kcal/mol.**

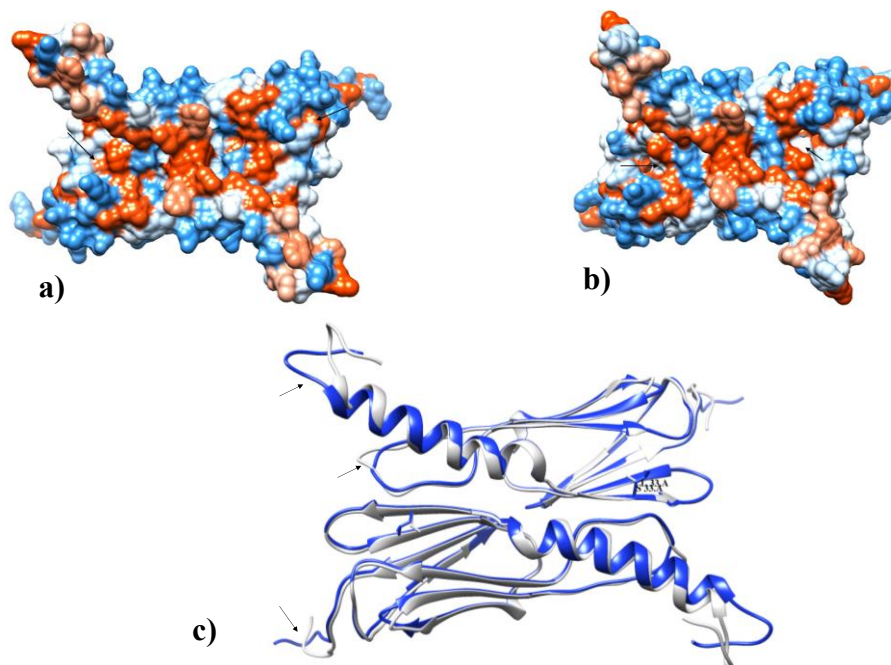
### 3.19 Comparison of the Hydrophobic Surface of the WT and Mutants

3-D hydrophobic surface models of *Tpv* sHSP14.3 WT and mutants were generated in order to understand the effect of the introduced mutations on the surface properties. CHIMERA program color the molecular surface by amino acid hydrophobicity on the Kyte-Doolittle scale such that min and max value are associated with the Dodger Blue (*i.e.*, bright blue) and orange red, respectively.

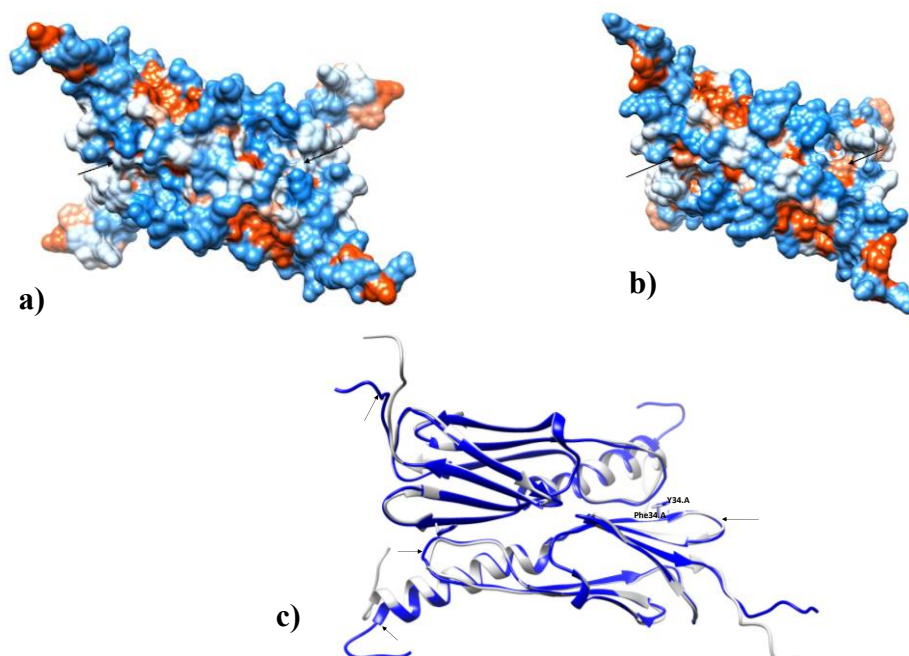
Introduction of the polar residue, serine, in the place of highly hydrophobic, leucine, decreased surface hydrophobicity at position 33, which was evident from the change of the color from red in the WT to blue in the L33S mutant. Differences in the surface properties of Leu33 in the *Tpv* sHSP14.3 WT and Ser33 in the mutant dimeric structure were marked by black arrow (Figure 3.145). Major differences were observed at the unstructured regions, NTD and CTD of the protein, as revealed by 3D superimpositions of the WT structure with that of the mutant. Differences in the coil structure of the proximal NTD of WT and mutant protein were also observed, where the Phe8 in the first coil moved apart between the two structures. Although replacement of the leucine with serine at position 33 did not induce significant structural changes at the beta strands of the ACD of the mutant protein, structural rearrangements at the  $\beta$ 4/ $\beta$ 5 and  $\beta$ 6 loop were observed (Figure 3.145). This could be due to the disappearance of the side chain interactions between  $\beta$ 2 and  $\beta$ 6 of the ACD after the mutation (*i.e.*, intermolecular hydrophobic interactions between Leu33-Ile78 and Leu33-Tyr77)

Besides, substitution of the Y by highly hydrophobic amino acid F at the position 34 to generate Y34F mutant increased surface hydrophobicity as compared to the WT. Superimposition of the structures showed that position of the initial N-terminal helices of the *Tpv* sHSP14.3 was found to be slightly shifted upon mutation. Besides, unstructured part of the NTD (Met1-Phe7) of the WT and the mutant moved in opposite directions. Also, positioning of the CTD of the protein was changed between the residues of Asn117 and Glu124, where they moved far apart from each other. The residue Y34 was located at the  $\beta$ 2 strand and mutating this residue by

phenylalanine shortened the  $\beta 2/\beta 3$  loop and affected the positioning of the  $\beta 6$  strand, which moved inward as compared to the WT (Figure 3.146).

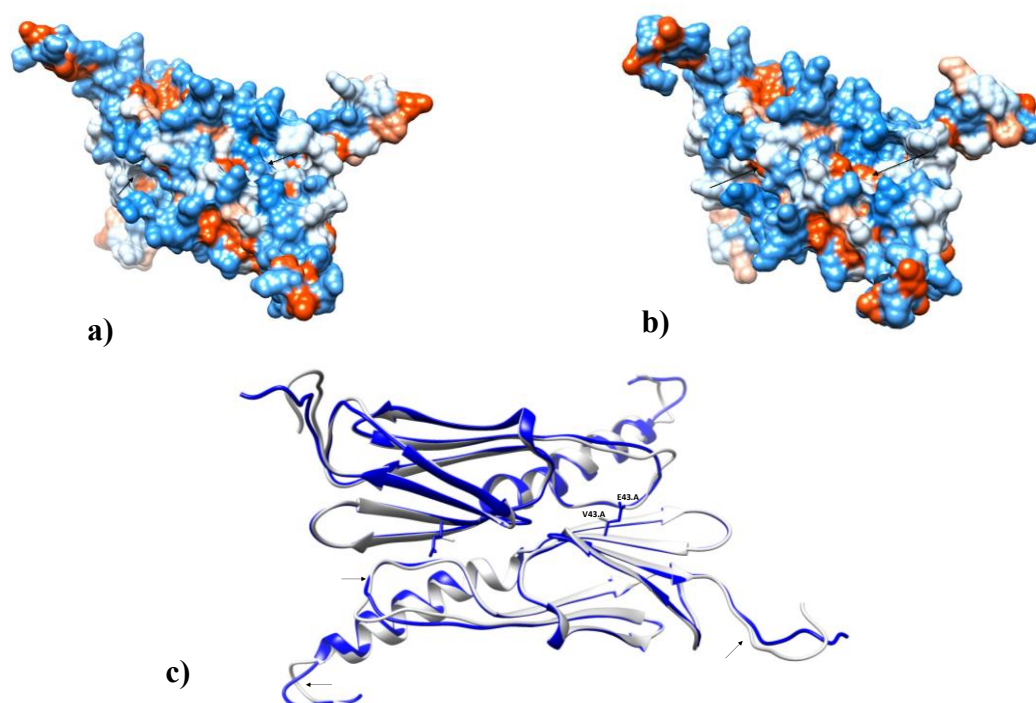


**Figure 3.145. Hydrophobic surface of *Tpv* sHSP14.3 WT and L33S mutant.** a) WT, b) L33S, c) Superimposition of the WT(Blue) and L33S (Gray) Structures. Arrow indicates the relative changes according to the WT.



**Figure 3.146. Hydrophobic surface of *Tpv* sHSP14.3 WT and Y34F mutant.** a) WT, b) Y34F, c) Superimposition of the WT(Blue) and Y34F (Gray) Structures. Arrow indicates the relative changes according to the WT.

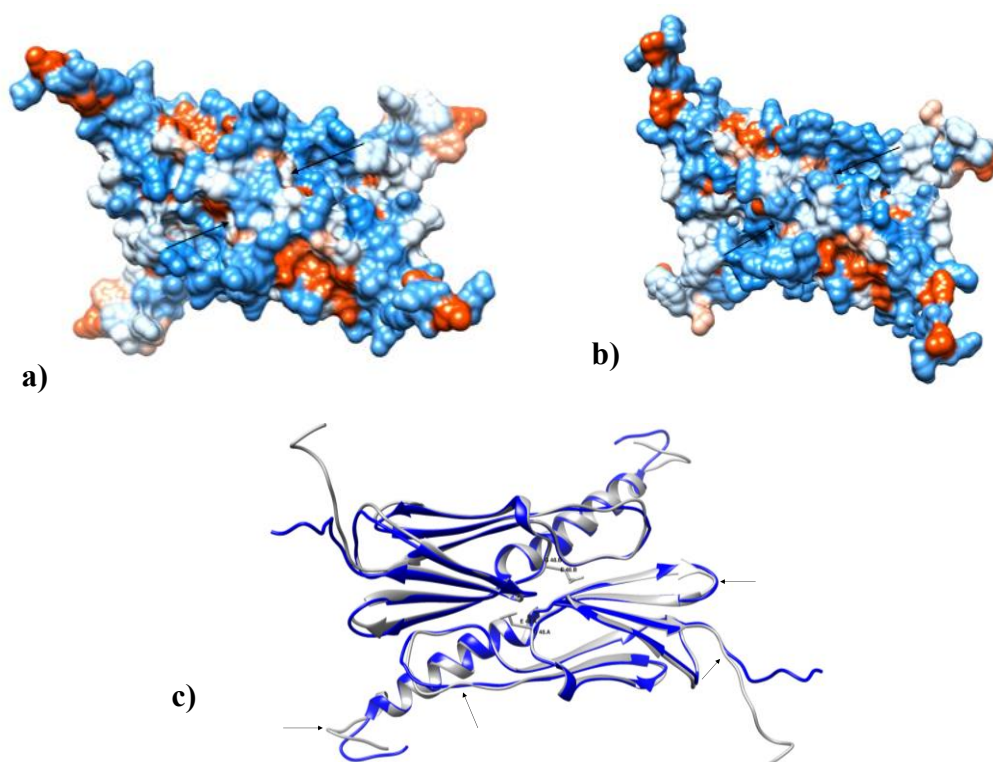
Similar to the Y34F mutation, E43V mutation by introducing hydrophobic amino acid at residue 43 enhanced hydrophobicity of the surface at that position (Figure 3.147). Structure comparison of the WT and E43V mutant identified differences mainly at the  $\beta 6$  strand of the ACD, where the strand belonging to the E43V mutant structure were pulled away. Change in the  $\beta 6$  strand of the mutant can be explained by the lost of the side chain interactions between the residue E43 from  $\beta 3$  strand and Arg81 from the  $\beta 6$  strand to form intermolecular salt bridges, which were present in the WT. Some minor differences were also observed for the unstructured NTD and CTD regions (Figure 3.147).



**Figure 3.147. Hydrophobic surface of *Tpv* sHSP14.3 WT and E43V mutant. a) WT, b) E43V, c) Superimposition of the WT(Blue) and E43V (Gray) Structures. Arrow indicates the relative changes according to the WT.**

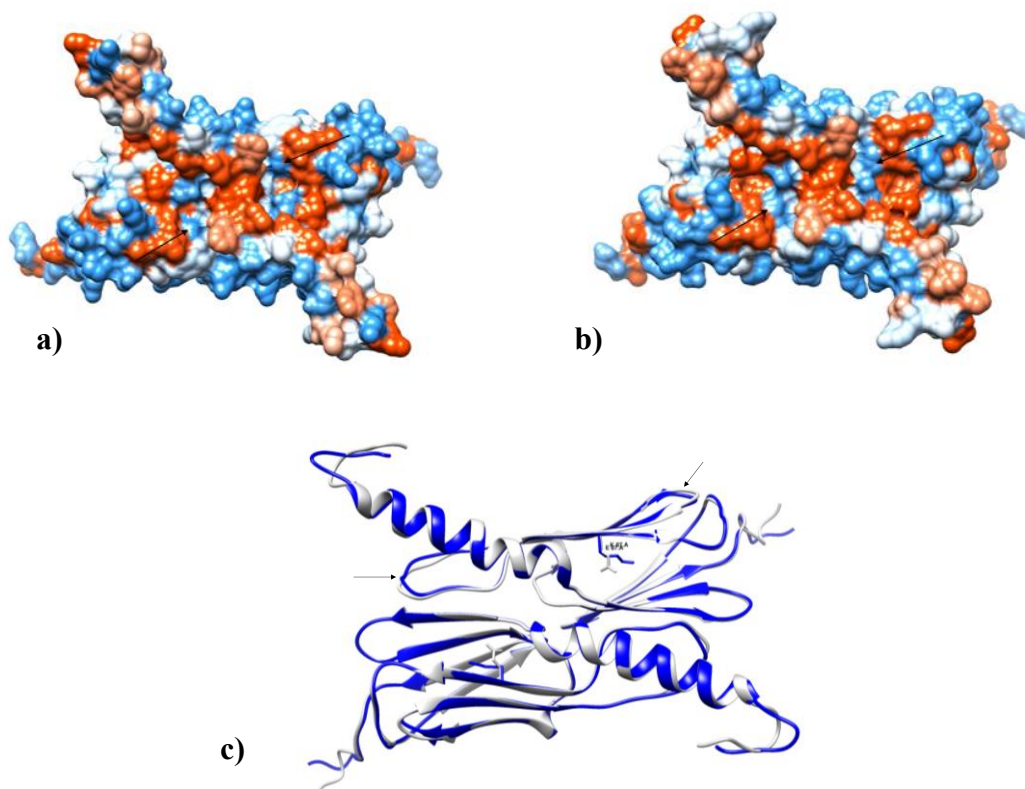
On the other hand, another mutation, G48E, altered the surface property towards hydrophilicity due to substituting charged residue glutamic acid in the place of neutral amino acid, glycine at the residue 48 (Figure 3.148). 3-D structure superimposition analysis of the WT structure with the that of the mutant identified

differences at the disordered region of the NTD (Met1-Phe7) and in the first helical coil part of it in which structures of the WT. Moreover, the CTD of the two structures moved opposite site. Besides, it was observed that in the G48E protein 3D structure,  $\beta 2/\beta 3$  loop and  $\beta 6$  strand of the ACD have changed by moving inward. Gain of the site chain interactions (*e.g.*, intramolecular hydrogen bond between the Arg69 from  $\beta 6$  strand and Glu48 from the  $\beta 3/\beta 4$  loop) may induce structural alterations in the loop (Figure 3.148).



**Figure 3.148. Hydrophobic surface of *Tpv* sHSP14.3 WT and G48E mutant. a) WT, b) G48E, c) Superimposition of the WT(Blue) and G48E (Gray) Structures. Arrow indicates the relative changes according to the WT.**

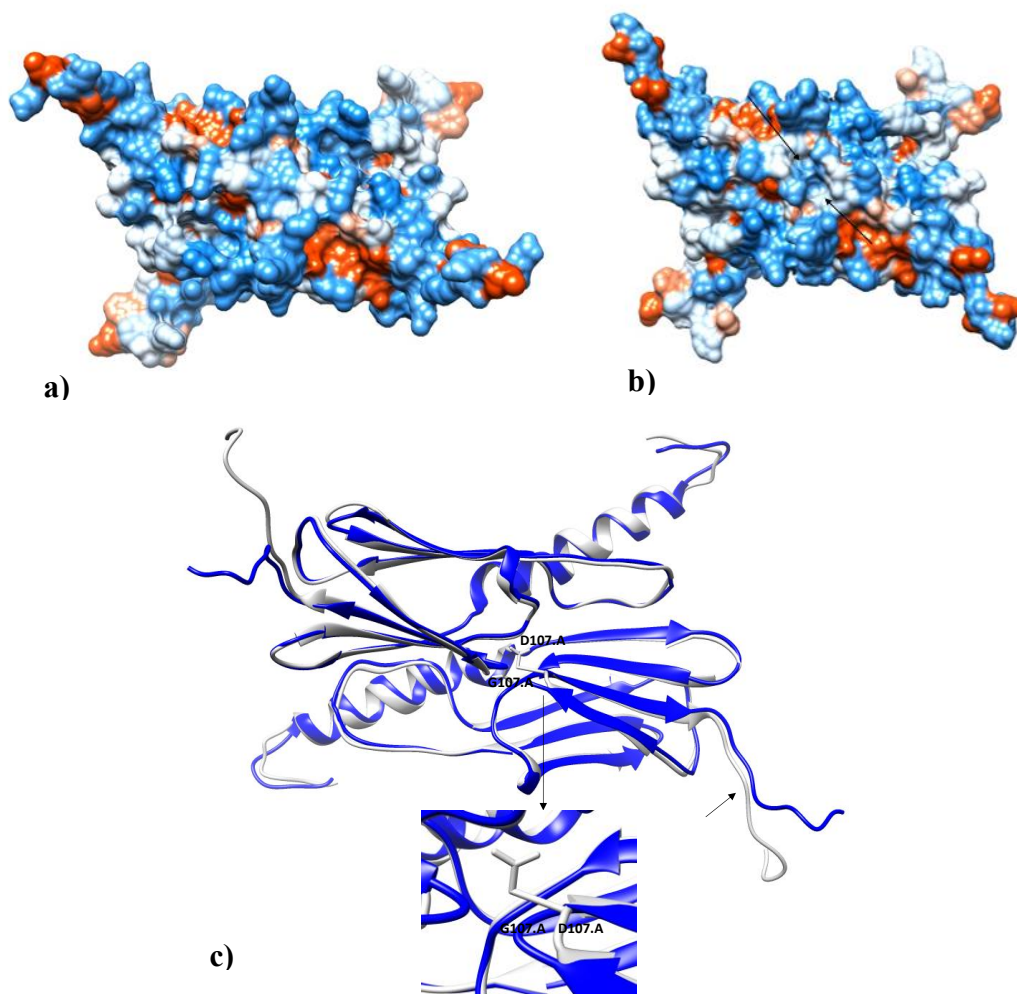
Besides, K87E mutation generated by reversing the charge at 87 did not alter surface hydrophobicity. Both WT and K87E mutant's surface have same hydrophilic character at that position, as understood from the color code. This mutation altered mainly changed the  $\beta 4/\beta 5$  loop and  $\beta 6$  loop in the ACD by moving them outward besides positioning of CTD and NTD (Figure 3.149).



**Figure 3.149. Hydrophobic surface of *Tpv* sHSP14.3 WT and K87E mutant. a) WT, b) K87E, c) Superimposition of the WT (Blue) and K87E (Gray) Structures. Arrow indicates the relative changes according to the WT.**

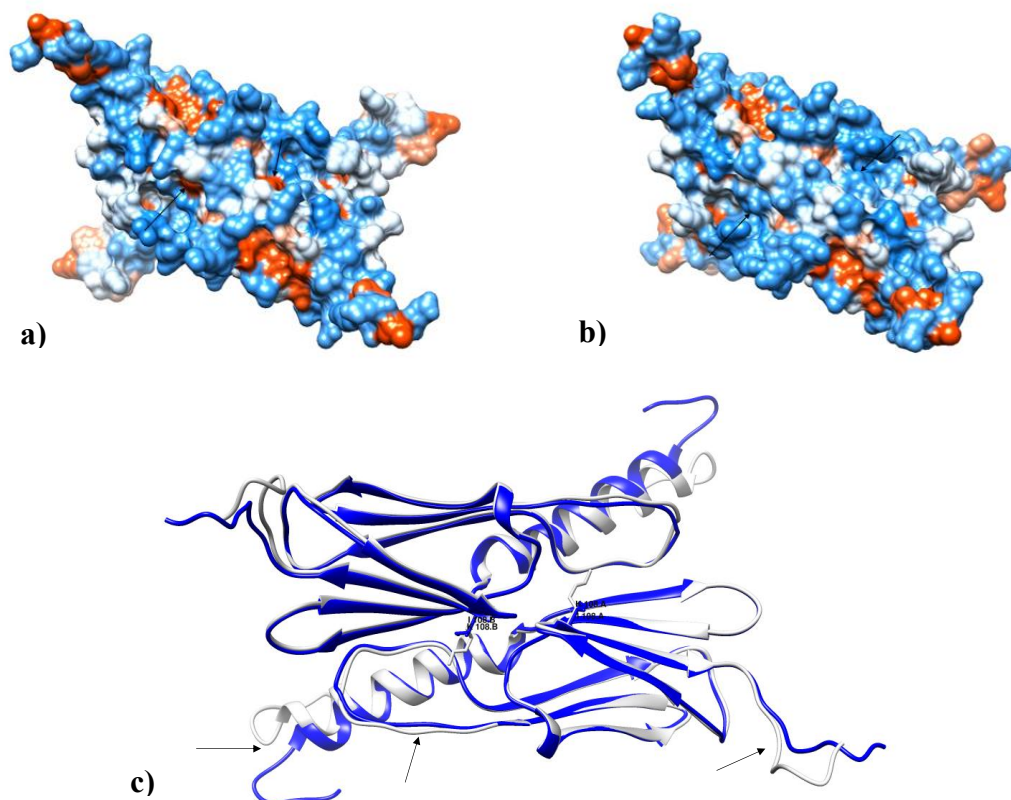
The Gly107 residue in the WT is a buried residue in the WT structure. Its substitution by aspartic acid changed residue character from buried to exposed ones. Also, hydrophilicity was enhanced at the position of 107 by G107D mutation, where the surface color became bright blue (Figure 3.150). When the mutant structure was superimposed with the WT structure, differences at the proximal of the unstructured NTD region were noticed, as well as the position of the CTD of the mutant sHSP shifted. Besides, the structural rearrangement at the the  $\beta 8/\beta 9$  loop of the G107D mutant protein occurred after the mutation. This loop become narrow (Figure 3.150).





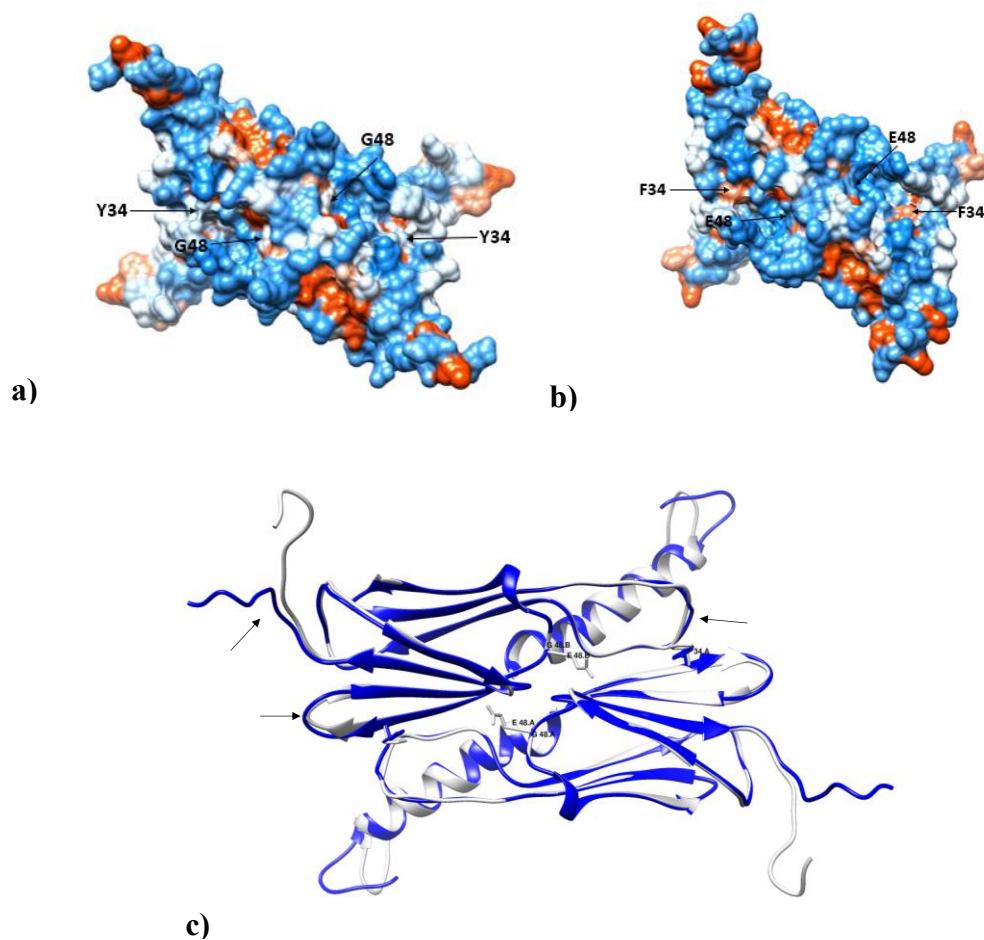
**Figure 3.150. Hydrophobic surface of *Tpv* sHSP14.3 WT and G107D mutant. a) WT, b) G107D, c) Superimposition of the WT (Blue) and G107D (Gray) Structures. Arrow indicates the relative changes according to the WT.**

In addition to this, surface analysis of I108K showed a change from hydrophobicity to hydrophilicity at residue 108. This was inferred from the color code where the hydrophobic isoleucine at position of 108 marked by red color, while substituted residue lysine marked by the blue color. The superimposition showed that initial part of the NTD coils towards the helix in the mutant. In addition to this, a slight change in the spatial arrangement of the CTD of this mutant was observed. Similar to the other *Tpv* ACD mutants, the  $\beta 6$  strand position was found to be changed after the I108K mutation (Figure 3.151).



**Figure 3.151. Hydrophobic surface of *Tpv* sHSP14.3 WT and I108K mutant. a) WT, b) I108K, c) Superimposition of the WT (Blue) and I108K (Gray) Structures. Arrow indicates the relative changes according to the WT.**

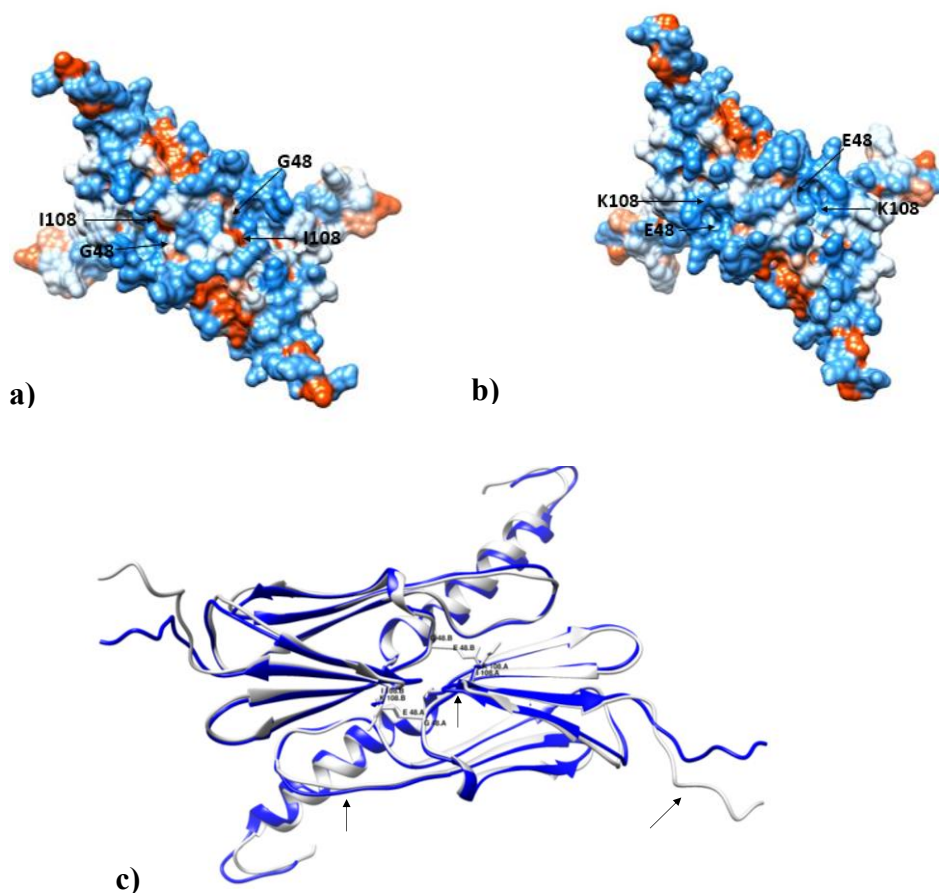
Changes in the surface properties as a result of mutation was also observed in the double mutants (Y34FG48E and G48EI108K). Similar to its single variants, double mutant Y34FG48E increased hydrophobicity at position 34 and reduced at position 48 (Figure 3.152). It was observed this double mutation induced some changes in the NTD, and  $\beta 2/\beta 3$  loop and  $\beta 6$  loop of the ACD. In the NTD part of the mutant protein, unstructured region between the Met1-Phe7 turned to form a loop at beginning of the helical coil. Besides, CTD of the Y34FG48E moved far apart from the WT structure and formation of the small coil observed between the Asn120-Glu124, which was towards the end of the CTD. Similar to their respective single mutant variants, the positioning of the  $\beta 2/\beta 3$  loop and  $\beta 6$  loop of the ACD was altered after the double mutation.



**Figure 3.152. Hydrophobic surface of *Tpv* sHSP14.3 WT and Y34FG48E mutant. a) WT, b) Y34FG48E, c) Superimposition of the WT (Blue) and Y34FG48E (Gray) Structures. Arrow indicates the relative changes according to the WT.**

Introduction of the charged residues Glu and Lys by G48E1108K mutation enhanced hydrophilicity at targeted mutation points, as understood from the light blue color in the double mutant indicating the strong hydrophilicity at those positions (Figure 3.153). 3-D structure superimposition of the mutant with WT showed that there were alterations mainly in the  $\beta 2/\beta 3$  loop and  $\beta 8/\beta 9$  loop in the ACD besides positioning of CTD. In the mutant structure, the mentioned loops moved inward since it became compact after the mutation.

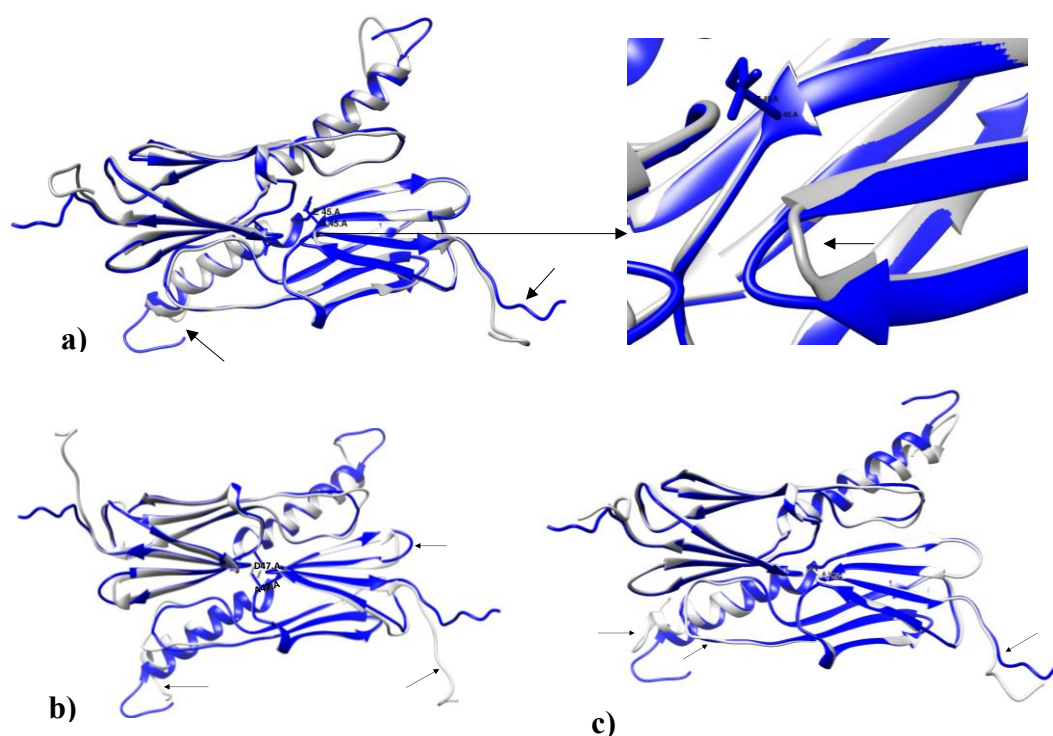
Our results suggest that even single amino acid replacement can change the surface properties, as well as the unstructured region of the protein (NTD, CTD).



**Figure 3.153. Hydrophobic surface of *Tpv* sHSP14.3 WT and G48E1108K mutant. a) WT, b) G48E1108K, c) Superimposition of the WT (Blue) and G48E1108K (Gray) Structures. Arrow indicates the relative changes according to the WT.**

Rest of the *Tpv* sHSP 14.3 ACD mutants, E45G, A47D and G107A, have all buried residues at targeted positions and they are not exposed to surface after mutations. The 3-D structural superimposition of E45G mutant sHSP structure with the that of WT identified the noticeable alterations in the NTD and CTD of the mutant sHSP. In the NTD part of the mutant protein, disordered region in the beginning (between the Met1-Phe7) turned towards the helical coil in the form of a loop. Similarly, CTD of this mutant protein also bended and formed coil towards the end of the CTD (between Lys119-Glu124). Although beta strands in the ACD of the mutant protein well superimposed with the WT protein,  $\beta 8/\beta 9$  loop of the mutant moved inward due to shortening of the size of the loop (Figure 3.154-a).

3-D superimposition analysis of the A47D mutant showed that CTD of the protein in the WT and in the mutant moved opposite site. Besides, in the first helical coil part of the NTD, structures of the WT and the mutant did not fit well. A shift towards the inside of the coil occurred after the A47D mutation. Moreover, the mutant structure became compact which was more evident at the  $\beta 2/\beta 3$  loop, whose size was reduced upon mutation (Figure 3.154-b). Furthermore, G107A mutation induced major structural changes in the NTD of the protein, where the first coil in the helix moved upward together with unstructured part. The CTD and  $\beta 6$  strand were slightly affected from the mutation. Different from the G107D mutation, structural rearrangement at the  $\beta 8/\beta 9$  loop was not observed for the G107A mutant, suggesting that substituting the residue with a similar biophysical characteristic of amino acid did not induce drastic alterations in that loop (Figure 3.154-c).



**Figure 3.154. Superimposition of the WT (Blue) and the E45G (a), A47D(b) and G107A(c) mutant structures. Arrow indicates the relative changes according to the WT.**

### 3.20 Prediction of *Tpv* sHSP14.3 WT and its Mutants Protein Stability After Mutation

Stability of proteins is a critical property which affects their proper function, activity and regulation. Incorrect folding and decreased stability are mainly arised from missense mutations (Khan & Vihinen, 2010). Protein thermodynamic stability is described by  $\Delta G$  that is equal to the difference between the free energy of folded and unfolded state (Bigman & Levy, 2020). MUpro web server was used for prediction of thermodynamic stability of our mutant proteins. Accuracy of the program was found more than 84%. If the relative stability change ( $\Delta\Delta G$ ) which is a difference between free energy of WT and mutant is positive, it is accepted that mutation increases stability, or vice versa. The method utilizes the support vector machine (SVM), sequence of the protein, and a confidence score between -1 to 1. A score less than 0 is attributed to the reduced the stability after mutation (J. Cheng *et al.*, 2005; Jianlin Cheng *et al.*, 2006).

According to the MUpro analysis, thermodynamic stability of *Tpv* sHSP14.3 ACD mutants were decreased, except thermodynamic stability of E43V mutation, which slightly increased with a high confidence score (Table 3.5). Among the *Tpv* sHSP14.3 ACD mutants, reduced thermodynamic stability was found more pronounced for the mutants, which was generated by substitution of charged residues in place of hydrophobic one (*e.g.*, I108K). In previous reports, it was stated that reduction in hydrophobic region, and over packing as a result of single point mutations are suggested as the factors that decrease protein stability (Khan & Vihinen, 2010). It is possible that reduced hydrophobicity in the L33S, A47D, I108K besides E43V could be a reason of a decline in the thermodynamic stability.

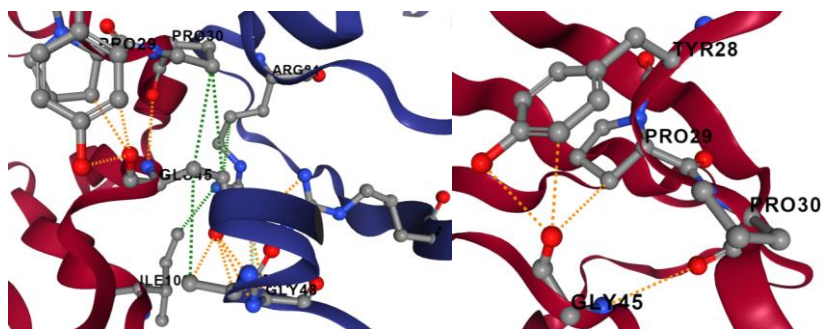
Two other single mutations, Y34F and K87E, whose  $\Delta\Delta G$  values are near to the zero, were found to have less effect on the thermodynamic stability as compared to the other mutants.

**Table 3.5. Thermodynamic stability of *Tpv* sHSP14.3 mutants**

| Samples      | $\Delta\Delta G$ | SVM Confidence Score | Comment |
|--------------|------------------|----------------------|---------|
| <b>L33S</b>  | -2.2027356       | -1                   | ↓       |
| <b>Y34F</b>  | -0.40968474      | -0.27927445          | ↓       |
| <b>E43V</b>  | 0.089916757      | 1                    | ↑       |
| <b>E45G</b>  | -2.1598572       | -1                   | ↓       |
| <b>A47D</b>  | -0.99872233      | 0.85089409           | ↓       |
| <b>G48E</b>  | -1.0470907       | -0.267276            | ↓       |
| <b>K87E</b>  | -0.47134194      | -0.1555863           | ↓       |
| <b>G107A</b> | -1.1048973       | 0.16524926           | ↓       |
| <b>G107D</b> | -0.81421305      | 0.27021835           | ↓       |
| <b>I108K</b> | -2.8952533       | -1                   | ↓       |

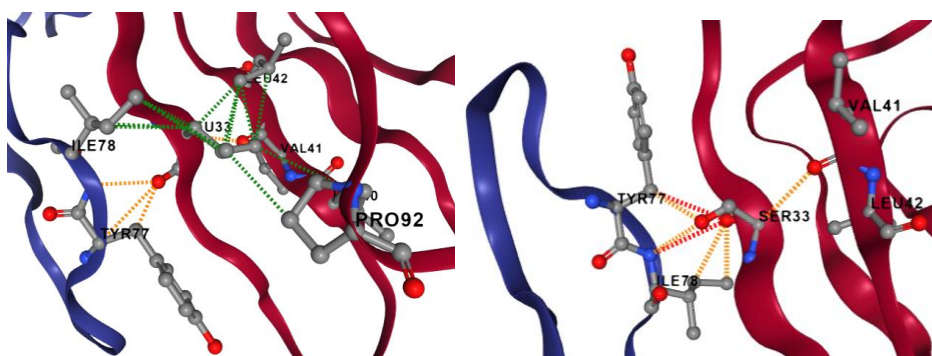
Secondly, Dynamut2 web server was also used for predicting the effects of single- and double-point mutations on protein stability and dynamics. Results were evaluated basen on the  $\Delta\Delta G$  values (*i.e.*,  $\Delta\Delta G < 0.0$  kcal/mol is destabilizing or vice verca), as in the MUpro programme. Also, the analysis of non-covalent interactions was given as an out put. Different from MUpro programme, this one enables us to study the effect of multiple mutations on the protein structure and thermodynamic stability. According to the Dynamut2 analysis, thermodynamic stability of the all *Tpv* sHSP14.3 ACD mutants were decreased (Table 3.6). Among the *Tpv* sHSP14.3 ACD mutants, the highest decrease in the thermodynamic stability of the protein was found for the E45G mutant, where the negative charged residue was replaced by the hydrophobic glycine. MUpro analysis also showed that E45G mutation drastically reduced thermodynamic stability. The Dynamut2 results showed that E45G mutation decreased the total number of non-covalent interactions (Figure 3.155). It is known that set of intramolecular interactions affect protein folding, stability, dynamics and consequently function (Rodrigues *et al.*, 2021). Therefore, low thermodynamic

stability of the E45G mutant protein could be due to the extensive loss interactions, which were present in the WT (Figure 3.155). This can also be the reason of decreased chaperone activity of the mutant protein against heat induced aggregation of the CS and low heat stability of the protein at 60°C.



**Figure 3.155. DynaMut2 result of the E45G mutant.** The non-covalent interactions of the WT(left) and the mutant structure (right).

According to Dynamut2 program, in L33S mutant reduced hydrophobicity was accompanied by a decrease in the thermodynamic stability of the sHSP protein, as well. The programme also showed that intermolecular (*e.g.*, Ile78-Leu33) and intramolecular hydrophobic interactions (*e.g.*, Leu33-Leu42, Leu33-Val41) were lost following the mutation (Figure 3.156). Altogether, this might be the reason for decrease in the thermal stability of the L33S mutant protein.

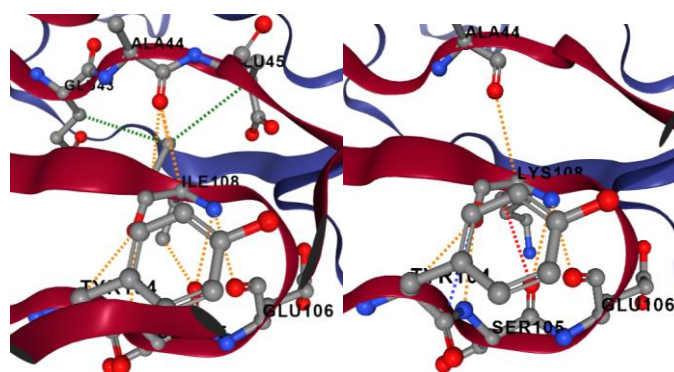


**Figure 3.156. DynaMut2 result of the L33S mutant.** The non-covalent interactions of the WT (left) and the mutant structure (right).

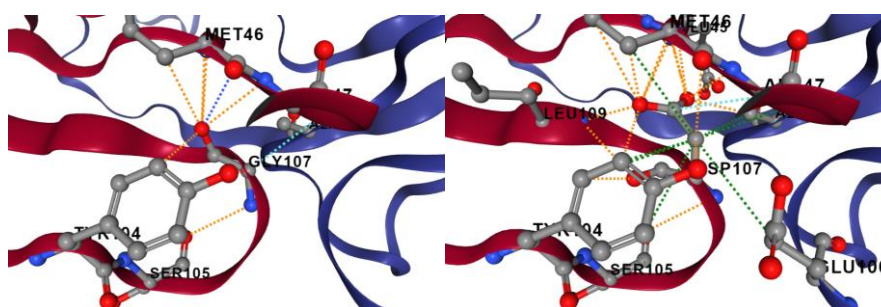
Similarly, I108K, and G107D single mutations involving decrease in hydrophobicity also reduced thermodynamic stability of the protein as found by Dynamut2 programme. The hydrophobic interaction between Glu45-Ile108 and Glu43-Ile108



present in the WT were disappeared by I108K mutation (Figure 3.157). Although G107D mutation increased the number of the interactions, thermodynamic stability of the protein was decreased (Figure 3.158). It might be possible that such new interactions in the G107D mutant can energetically destabilize the protein structure. Reduced thermodynamic stability of the G107D mutant protein may account for the inactivation of this mutant sHSP even at 60°C.

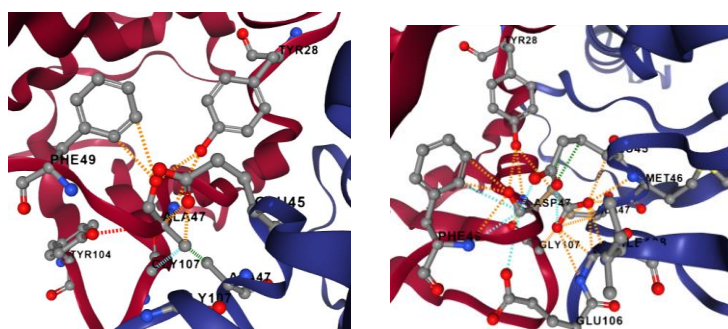


**Figure 3.157. DynaMut2 result of the I108K mutant.** The non-covalent interactions of the WT (left) and the mutant structure (right).

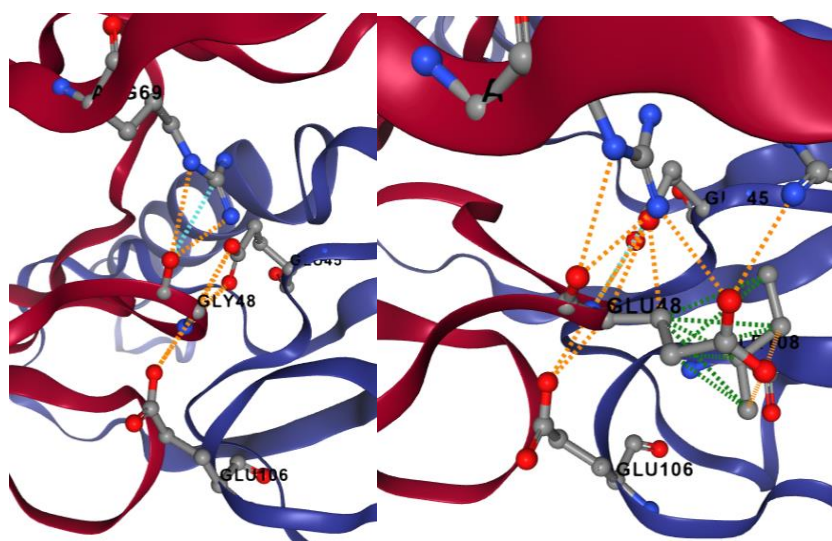


**Figure 3.158. DynaMut2 result of the G107D mutant.** The non-covalent interactions of the WT (left) and the mutant structure (right).

For the two other single mutations, A47D and G48E reduced protein stability can be discussed on the same bases. Their  $\Delta\Delta G$  values were almost same (-1.1). In both mutants, introduction of the charged residue in the place of the hydrophobic amino acid may cause perturbation of network of the interaction so that protein can be energetically destabilized (Figure 3.159-3.160).

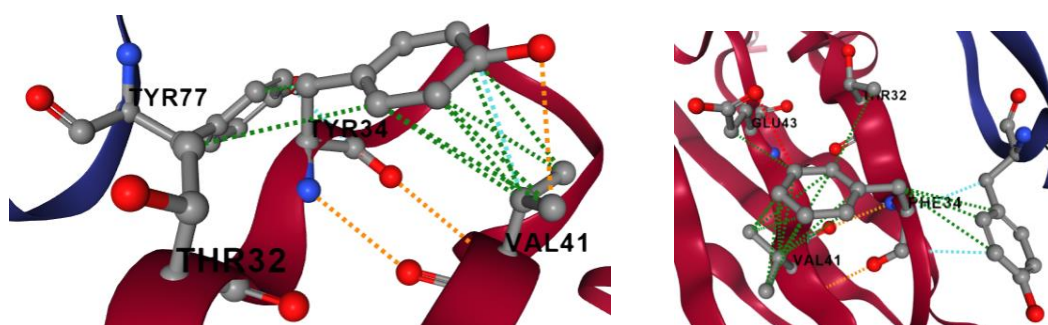


**Figure 3.159. DynaMut2 result of the A47D mutant.** The non-covalent interactions of the WT (left) and the mutant structure (right).

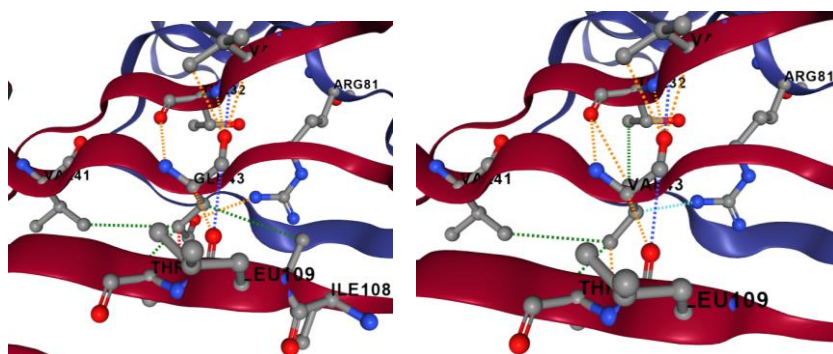


**Figure 3.160. DynaMut2 result of the G48E mutant.** The non-covalent interactions of the WT (left) and the mutant structure (right).

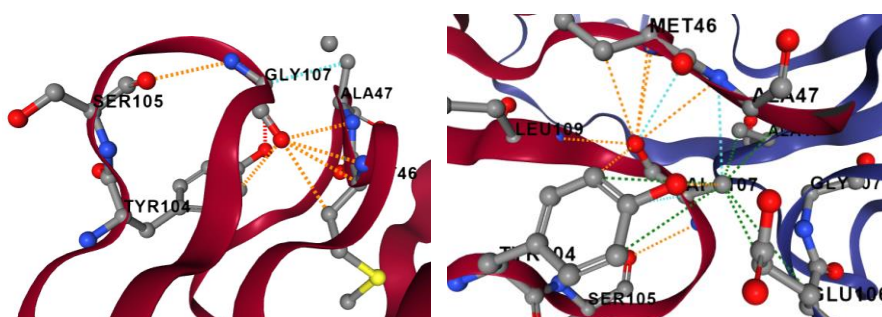
On the other hand, three single mutations, Y34F, E43V and G107A also decreased the thermodynamic stability as revealed by Dynamut analysis, but their effect was not so pronounced (Table 3.6). The non-covalent interactions present in the Y34F mutant and E43V were slightly higher than that of WT (Figure 3.161-3.162). On the other hand, G107A mutant formed more new contacts, which were not available in the WT (Figure 3.163).



**Figure 3.161.** DynaMut2 result of the Y34F mutant. The non-covalent interactions of the WT (left) and the mutant structure (right).

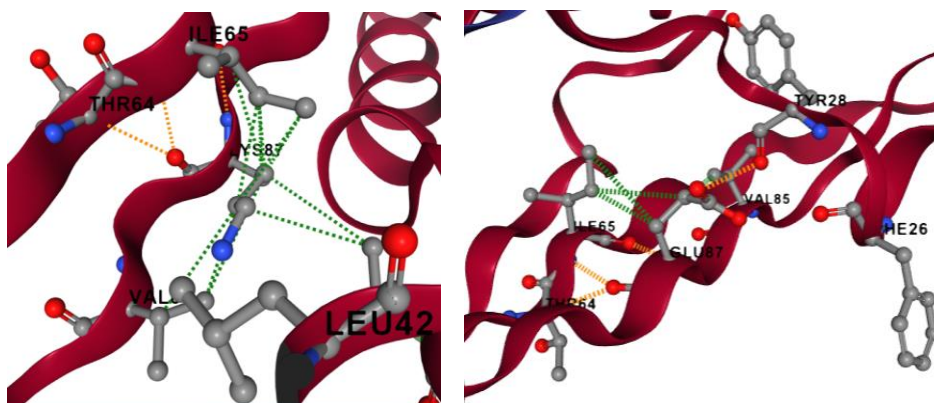


**Figure 3.162.** DynaMut2 result of the E43V mutant. The non-covalent interactions of the WT (left) and the mutant structure (right).



**Figure 3.163.** DynaMut2 result of the G107A mutant. The non-covalent interactions of the WT (left) and the mutant structure (right).

$\Delta\Delta G$  values of K87E mutant were found to be near zero, which could have less effect on the thermodynamic stability as compared to the other mutants according to the Dynamut programme. As a result of this mutation, hydrophobic interactions present in the WT were lost after the mutation (Figure 3.164).



**Figure 3.164. DynaMut2 result of the K87E mutant.** The non-covalent interactions of the WT (left) and the mutant structure (right).

Furthermore, thermodynamic stability was reduced by Y34FG48E and G48EI108K double mutations as revealed by Dynamut2 analysis. The latter had more pronounced effect on destabilizing the structure ( $\Delta\Delta G = -2.23$ ) than the former one ( $\Delta\Delta G = -0.23$ ). For the double mutations, this programme did not show non-covalent interactions, instead it gave the distance values between the mutation points such that Y34FG48E mutation has distance of 23.8 Å, whereas, G48EI108K has distance of 9.5 Å in the 3-D structure.

**Table 3.6. Thermodynamic stability of *Tpv* sHSP14.3 mutants using DynaMut2 Programme**

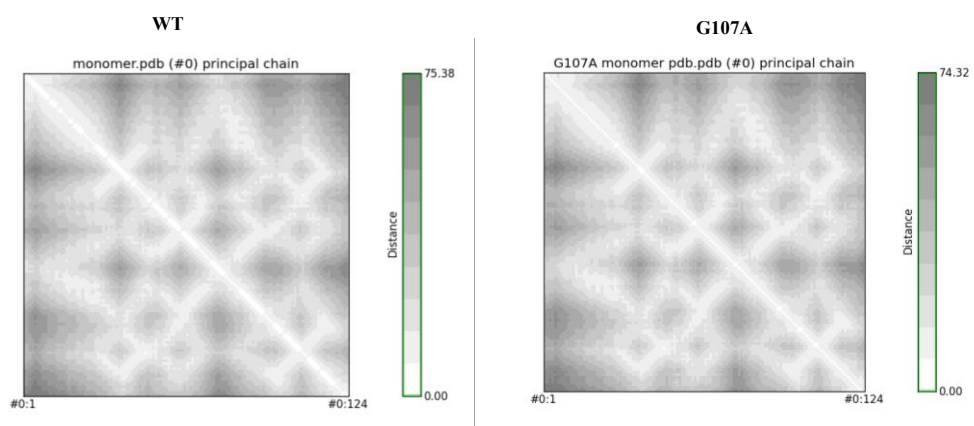
| Samples   | DynaMut2<br>$\Delta\Delta G$ |
|-----------|------------------------------|
| E45G      | -2.81                        |
| L33S      | -2.17                        |
| I108K     | -1.37                        |
| G107D     | -1.19                        |
| A47D      | -1.13                        |
| G48E      | -1.09                        |
| Y34F      | -1.15                        |
| E43V      | -1.06                        |
| G107A     | -0.53                        |
| K87E      | -0.26                        |
| Y34FG48E  | -0.21                        |
| G48EI108K | -2.23                        |

### 3.21 Generation of Residue-Residue Distance Map Using UCSF Chimera Tool

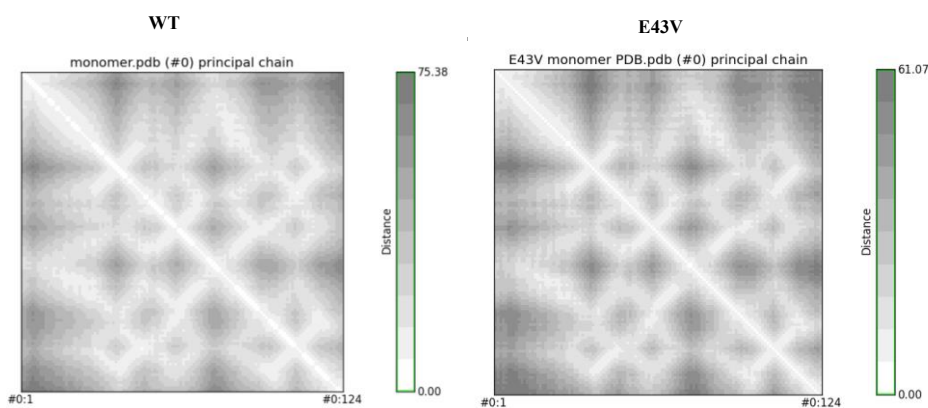
Residue-Residue (RR) Distance between alpha carbon atoms of amino acid pairs were computed using the RRDistMaps tools of UCSF Chimera. This analysis enables us to identify all pairwise inter-residue distances and to compare the 3-D structure of the individual proteins. Residue- residue distances are represented by a grayscale distance map together with a color coding legend in which white is associated with residues at short distances; whereas, dark gray is indicating the residues that are far apart from each other (Chen *et al.*, 2015).

All the mutations planned in this study resulted in decreasing the RR distance of the monomer molecule such that, indicating the mutant monomer to make a more compact configuration than the *Tpv* sHSP 14.3 WT. Overall results show that, the the mutation that resulted in the least affect on the RR distance was increasing hydrophobicity at position 107 (G107A). It resulted in a slight decrease of RR distance by 1 Å (WT = 75.38; G107A = 74.32) (Figure 3.165). Making an already hydrophobic position, more hydrophobic, by G107A mutation, although results in a slight movement of the CTD towards the centre of the monomer, the extension of the NTD away from the centre might have compensated such that the overall RR distance of this mutant was the most similar to WT (See Section 3.19, Figure 3.154-c)

On the other hand, introduction of hydrophobicity, valine, at a highly conserved position, in terms of charge, E43 resulted in the maximum compactness of the monomer molecule which was 14 Å less than WT (Figure 3.166). This can be correlated with the ribbon structure where, not only the NTD was prone to fold towards the  $\beta$ 6 loop but also the 3-D spatial configuration of the CTD flexible coil was affected and it lies near to the  $\beta$ -sandwich (See Section 3.19, Figure 3.147-c).

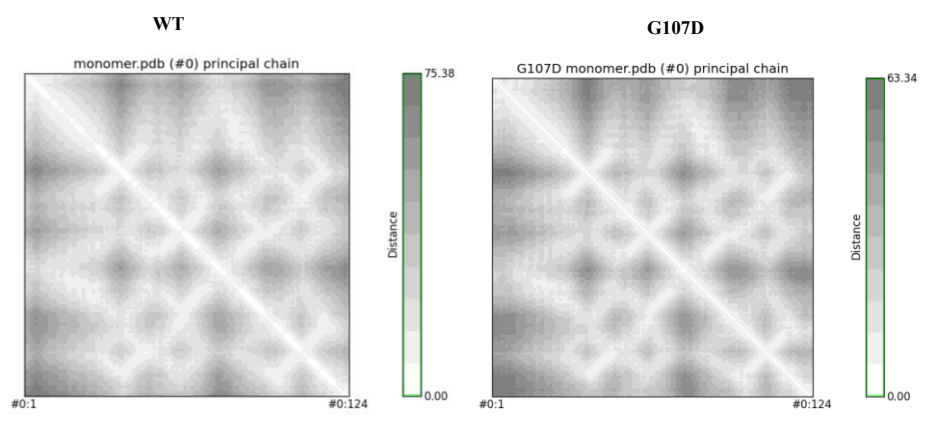


**Figure 3.165. RRDist map analysis of *Tpv* sHSP14.3 WT and G107A mutant**

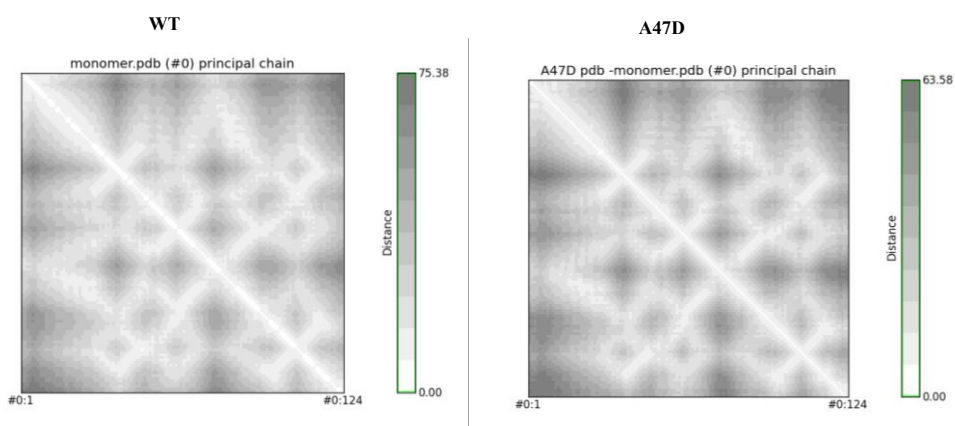


**Figure 3.166. RRDist map analysis of *Tpv* sHSP14.3 WT and E43V mutant**

Two mutations of similar nature, planned in this study, resulted in similar decrease in the RR distance of the generated mutants. The introduction of negative charge, (aspartate) in place of hydrophobic alanine and glycine at positions 47 and 107, respectively, resulted in the molecule having an average distance of 63 Å (Figure 3.167-3.168). Both the mutants showed a similar arrangement of the terminal regions. The CTD was less extended, away from the  $\beta$ -sandwich, in both the mutant monomers (See Section 3.19, Figure 3.150-c and Figure 3.154-b). In addition to this, the unstructured region of the NTD also showed a similar displacement which was opposite to that in the WT. Although, the  $\beta_6$  loop of both the monomers were oriented differently, where, in G107D, the loop was further away from the NTD in a monomeric structure. However, the effect of this might not be so much to affect the RR distance to an extent making it dissimilar to A47D.

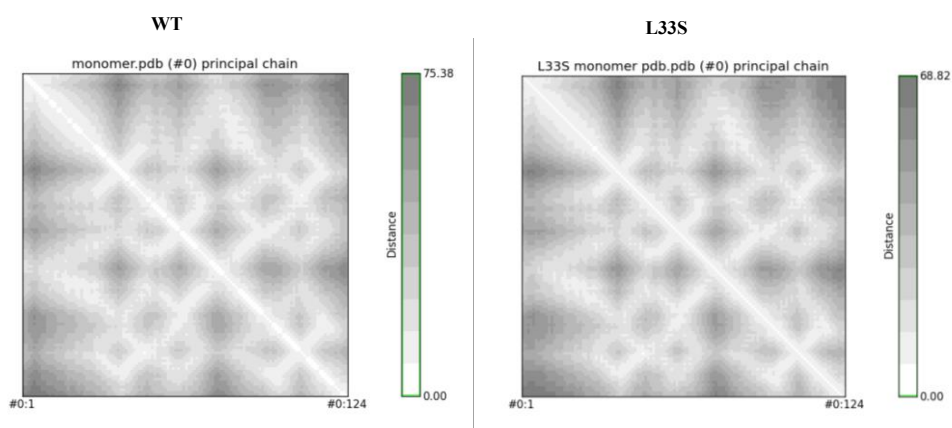


**Figure 3.167. RRDIST map analysis of *Tpv* sHSP14.3 WT and G107D mutant**



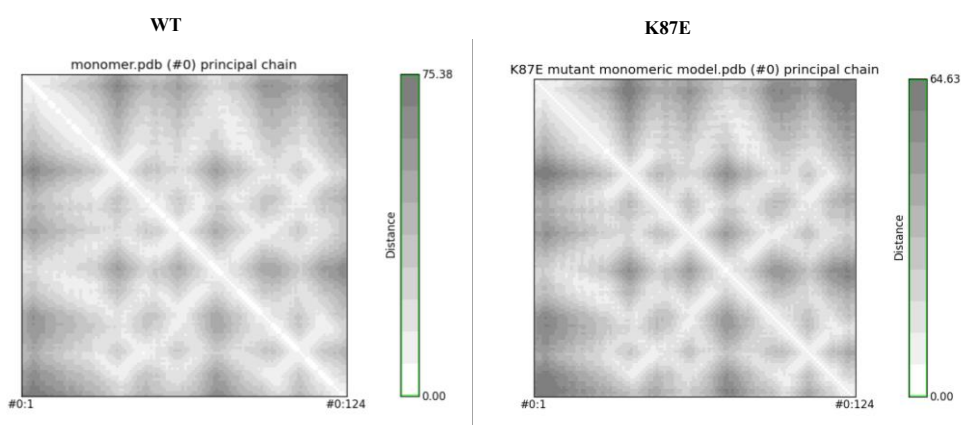
**Figure 3.168. RRDIST map analysis of *Tpv* sHSP14.3 WT and A47D mutant**

Another mutation, where hydrophobicity was reduced at a highly conserved position, 33, this time by replacing it with polar uncharged serine residue (L33S), also resulted in a similar compactness of the monomer, as above (RR distance = 68.82) (Figure 3.169). The NTD in this mutant monomer also showed similar changes in the placement as A47D and G107D. Not only this, the shape and arrangement of  $\beta_6$  loop of L33S mutant was also similar to the above two mutants (See Section 3.19, Figure 3.145-c). Thus, the similarity in the decrease in distance between  $C\alpha$  atoms was expectedly similar for all three, L33S, A47D and G107D.



**Figure 3.169. RRDist map analysis of *Tpv* sHSP14.3 WT and L33S mutant**

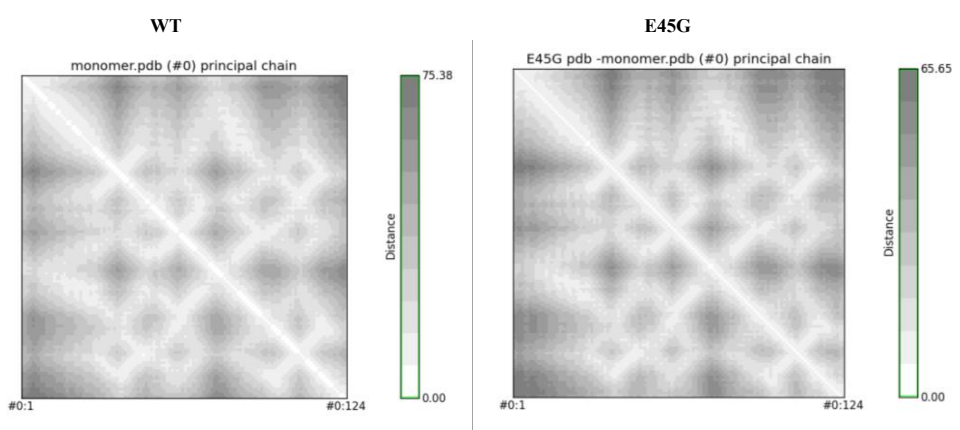
Although the mutation, that involves the reversing of charges at a crucial position of 87 (K87E), resulted in the CTD of the monomer to move further towards the  $\beta$  sandwich, the NTD moved away from the centre of the monomer (See Section 3.19, Figure 3.149-c). The compactness of the structure is more likely due to an 11 Å decrease in the RR distance (Figure 3.170).



**Figure 3.170. RRDist map analysis of *Tpv* sHSP14.3 WT and K87E mutant**

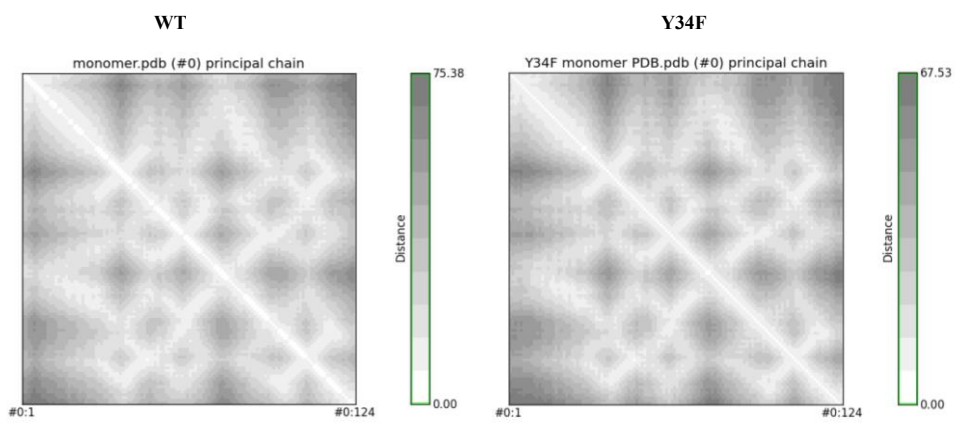
Abolishing charge and replacing it with hydrophobicity at a putative dimer interface position, the E45G mutation, resulted in a decrease of RR distance of almost 10 Å, as compared to the WT (Figure 3.171). This might be due to the reason that the terminal portion of the CTD, in this mutant monomer showed a complete bend towards the  $\beta$ -sandwich. Besides this, the NTD of the mutant also shows structural dissimilarities at the terminal position, as compared to the WT (See Section 3.19, Figure 3.154-a).





**Figure 3.171. RRDist map analysis of *Tpv* sHSP14.3 WT and E45G mutant**

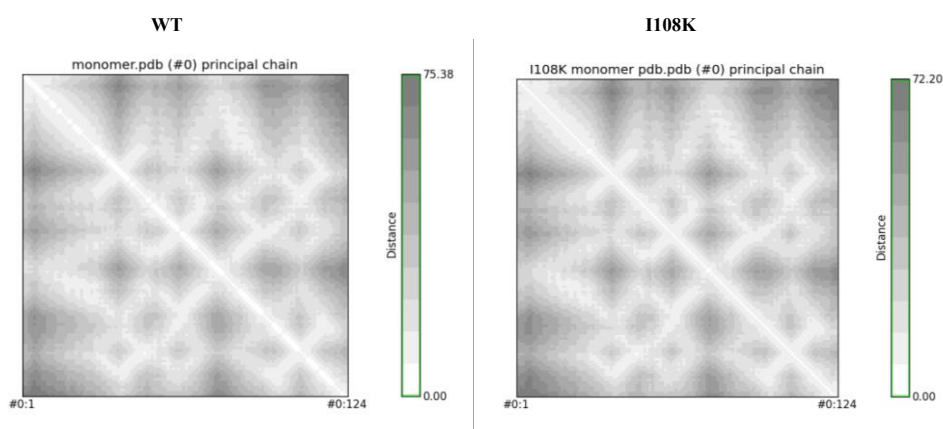
Distance map analysis of the Y34F mutant showed that overall RR distance in the monomeric structure was reduced when compared to WT. Color code difference was mostly evident at the beginning of the NTD. This result may indicate that Y34F mutation may have affected the 3-D structure of protein such that Y34F became more compact, especially at the proximal of unstructured region NTD region (See Section 3.19, Figure 3.146-c). In accordance with this result, there has been about 8 Å decrease of RR-distance (Figure 3.172).



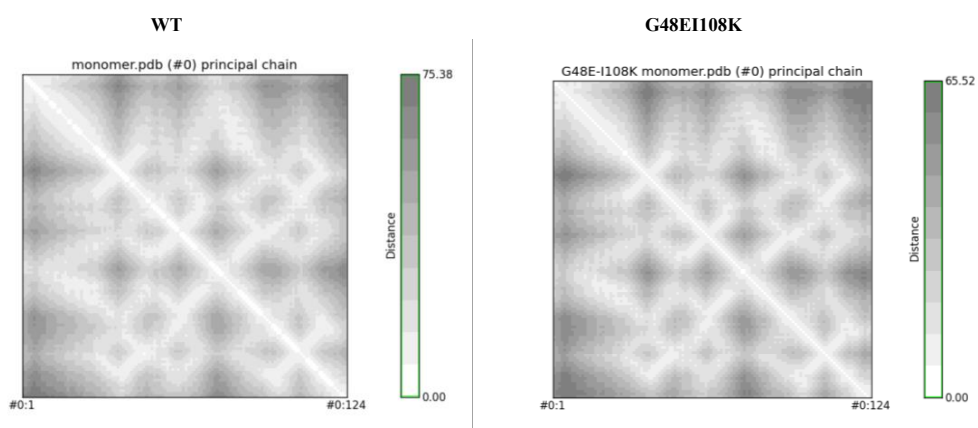
**Figure 3.172. RRDist map analysis of *Tpv* sHSP14.3 WT and Y34F mutant**

Although I108K point mutation had only a slight affect on overall residue distance, G48E/I108K double mutation significantly reduced the RRDist of the *Tpv* 14.3 sHSP (Figure 3.173-3.174). The latter, altered mainly the size of  $\beta$ 2/ $\beta$ 3 loop and  $\beta$ 8/ $\beta$ 9 loop in the ACD besides positioning of CTD. Relatively short loop size could be a

factor increasing compactness of the mutant protein (See Section 3.19, Figure 3.153-c).



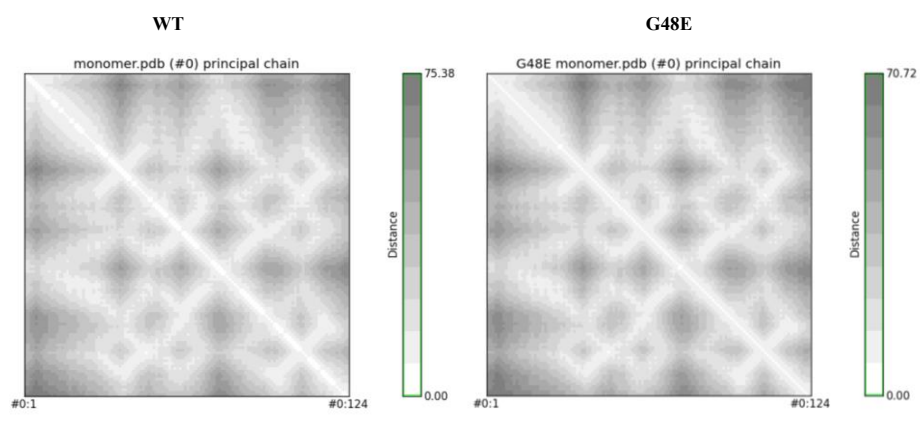
**Figure 3.173. RRDist map analysis of *Tpv* sHSP14.3 WT and I108K mutant**



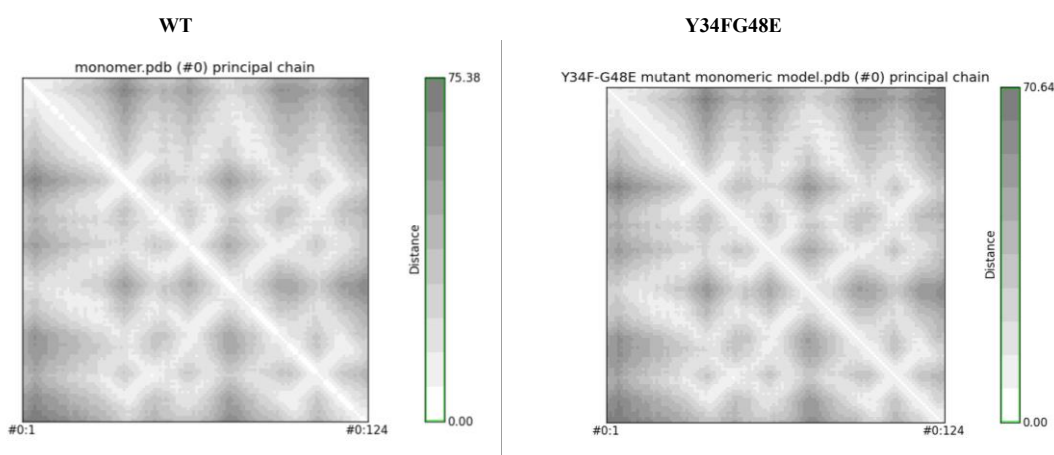
**Figure 3.174. RRDist map analysis of *Tpv* sHSP14.3 WT and G48E/I108K mutant**

G48E point mutation and Y34FG48E double mutation have reduced the overall RRDist of *Tpv* 14.3 sHSP (about 5 Å) at the same extend (Figure 3.175-3.176). It was observed that in the double mutant protein 3D structure, NTD,  $\beta$ 2/ $\beta$ 3 loop and  $\beta$ 6 loop of the ACD have changed (See Section 3.19, Figure 3.148 and 3.152-c).

Overall, it is seen that, the maximum changes have been occurred in the terminal regions, the NTD and CTD, as a result of mutation. Thus, it can be interpreted that, the substitutions occurring in the highly conserved position in a well structure ACD of a sHSP still impacts the flexible disordered regions in such a way so as to result in altered 3-D arrangements, especially of the unstructured regions.



**Figure 3.175. RRDist map analysis of *Tpv* sHSP14.3 WT and G48E mutant**



**Figure 3.176. RRDist map analysis of *Tpv* sHSP14.3 WT and Y34FG48E mutant**

### 3.22 Residue Contact Map Prediction of Mutants and WT

Intramolecular interactions in the polypeptide chain are crucial for formation of stable secondary structures and specific tertiary structures of proteins. The alpha carbon is the center in the backbone of every amino acid. The distance between the alpha carbon atom of each residue in a polypeptide chain can give information about thermal stability of proteins, the folding rate of proteins, and protein stability after mutations (Ponnuswamy & Michael Gromiha, 1994; Gromiha & Selvaraj, 2004). The rapid improvement in the field of computational biology gives us opportunity to understand interactions in the protein. A web server, SCRATCH (available at <http://scratch.proteomics.ics.uci.edu>), was used to predict secondary and tertiary

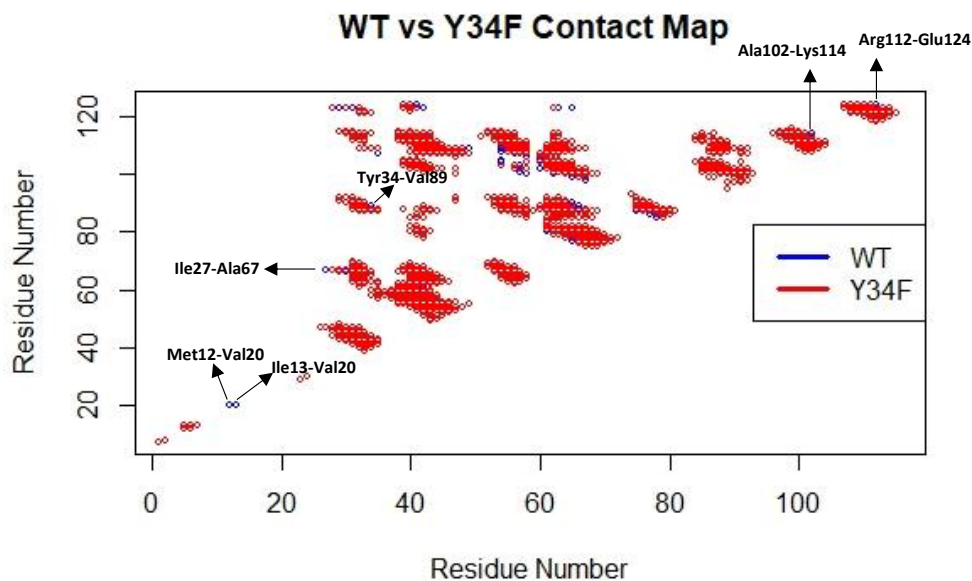
structural features of the *Tpv* sHSP 14.3 WT and its mutants. Detection of contacts between C $\alpha$  atoms of the amino acids within the protein was found by performing contact map analysis using SVMcon predictor in SCRATCH. This analysis generates a 2-D matrix consisting of residue-residue interactions within a distance at 8 Å between C $\alpha$  atoms of amino acids. Its accuracy is 4% higher than another contact map predictor tool, CMAPpro (Jianlin Cheng & Baldi, 2007). Our search using SVMcon predictor provided a contact file for the monomeric sHSP proteins that contained information about contacts for each residue, the cut off distance (Å) between contains atoms and probability scores for predicting contacts. Total number of contacts in the monomeric structures of the WT and mutant variants is given in Table 3.7. Furthermore, data visualization was processed by using R programming language (R Core Team, 2020). Simple scatter plot was generated in the R Programming in order to show contact points interacting with each residue. Each dot marked the residues in contact when the distance threshold between alpha carbon atom was 8 Å. Besides, comparative graphs were also plotted to distinguish contacts in WT and mutants. Contacts were evaluated depending on the separated residues such that short-range contacts are those separated by 6–11 residues in the sequence; medium-range contacts are those separated by 12–23 residues, and long-range contacts are those separated by at least 24 residues (Girirajan *et al.*, 2011).

**Table 3.7. Summary of contact prediction results of *Tpv* sHSP 14.3 and its mutants**

| Samples   | Total Number of contacts in the monomeric structure of <i>Tpv</i> sHSP14.3 | Common contacts with WT <i>Tpv</i> sHSP14.3 | Mutant specific contacts |
|-----------|--|---|--------------------------|
| WT        | 831  |   |                          |
| Y34F      | 841  | 793   | 48                       |
| G107A     | 817  | 776   | 41                       |
| L33S      | 839  | 794   | 45                       |
| G48E      | 820  | 783   | 37                       |
| I108K     | 839  | 794   | 45                       |
| G107D     | 813  | 774   | 39                       |
| A47D      | 814  | 774   | 40                       |
| E45G      | 821  | 782   | 39                       |
| E43V      | 788  | 734   | 54                       |
| G48EI108K | 839  | 798   | 41                       |
| Y34FG48E  | 832  | 782   | 50                       |
| K87E      | 818  | 785   | 33                       |

Introduction of the highly hydrophobic amino acid (phenylalanine) in the place of tyrosine at residue 34 (Y34F mutation) increased the total number of contacts as compared to the WT by formation of 48 new contacts (Table 3.7). On the other hand, this mutation resulted in loss of the 38 residue contacts within the monomeric structure. There are four new contacts formed between the residues 1 – 13 (*i.e.*, Met1-Phe7, Ile5-Met12, Lys6-Ile13, Lys6-Met12) in the proximal NTD (Appendix I, Table I.1). This mutation also resulted in the loss of the two contacts between the alpha carbon atom of the Met12-Val20 and Ile13-Val20 residues of the NTD of the sHSP protein. On the other hand, structural superimposition of Y34F with the WT showed that Y34F mutation caused alterations in the 3-D arrangement of the NTD such as a shift in the initial part of the N-terminal helix. The residue contacts alterations in the NTD and structural changes can be a reason of the slight enhancement of the chaperone activity of the Y34F mutant protein against heat induced inactivation of the model substrates, CS and ADH. When we compared two contact maps of WT and Y34F, most of the contacts which were different from WT were present in the alpha crystallin domain of the *Tpv* sHSP14.3 (residues lying between 33-114) as shown by two different colors and labelled (*e.g.*, Ile27-Ala67, Ala102-Lys104, Tyr34-Val89) (Figure 3.177). Besides, Y34 to F mutation specific three contact pairs (*i.e.*, F34-Ser 66, F34-Lys114 and F34-Lys60) were formed, which are not found in the WT. Most of the contacts at the position 34 were between hydrophobic residue pairs (*e.g.*, Leu40, Val41, Leu42, Ile65, Val88, Ala44) (Appendix I, Table I.2). Among those contacts with hydrophobic residues, the contacts between the alpha carbon atom of the Tyr34 in WT or Phe34 in the mutant and Val41 which have higher probability score as compared to the other contacts, also involved in the formation of intramolecular hydrophobic and hydrogen bonds, as we found by DSV programme.

Contact map analysis of the G107A mutant showed that total number of contacts decreased by 14 as compared to the WT. Out of 817 inter-residue contacts of G107A mutant protein, 41 contacts did not have their equivalent in the WT (Table 3.7).

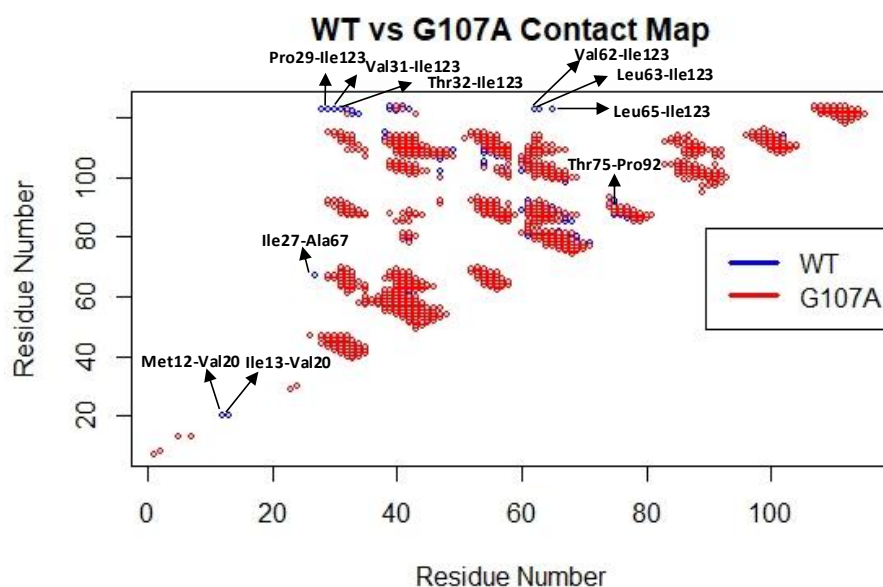


**Figure 3.177. Contact map analysis of WT and Y34F mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.

One of the mutant specific contact was between the Met1-Phe7, which can be related to the helical change in helical structure in the beginning of the NTD, as revealed by the 3-D superimpositions. In addition to these, NTD of the WT showed more contacts as compared to that of the mutant. For example, the hydrophobic contacts, Met12-Val20, Ile13-Val20, Ile27-Ala67, Tyr28-Ile123, Pro29-Ile123, Pro30-Ile123, Val31-Ile123, and Thr32-Ile123, were not found in the mutant sHSP (Appendix I, Table I.3). These changes in the hydrophobic contacts within the NTD could be a reason for decreased chaperone activity of the G107A mutant sHSP in preventing the CS thermal aggregation and the activity, considering involvement of the NTD in substrate recognition and protection.

When compared with the contact map of the WT, some contacts with the residue Ile123 at the CTD of the *Tpv* sHSP 14.3 WT (*i.e.*, Pro29-Ile123; Val31-Ile123; Thr32-Ile123; Val62-Ile123; Leu63-Ile123; Ile65-Ile123) are not present in the G107A mutant sHSP (Figure 3.178). Since the targeted mutation point is located towards the end of the ACD, it is not surprising that the residue contacts in the CTD region would be more affected. Also, the residue contact differences between WT

and G107A were more pronounced for hydrophobic-hydrophobic contacts and hydrophobic-polar contacts. However, the contacts of the alpha carbon Gly107 or Ala107 remained same (Appendix I, Table I.4). Both of them made contacts with the Ile123 and Glu124. Thermophilic proteins mostly prefer to form residue contacts between hydrophobic and polar residues, which is important for their stability (Gromiha, 2001). Therefore, decreased number of total contacts and losing some hydrophobic contacts within the G107A monomer structure could be responsible for reduced thermal stability of the mutant protein as compared to the WT.



**Figure 3.178. Contact map analysis of WT and G107A mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.

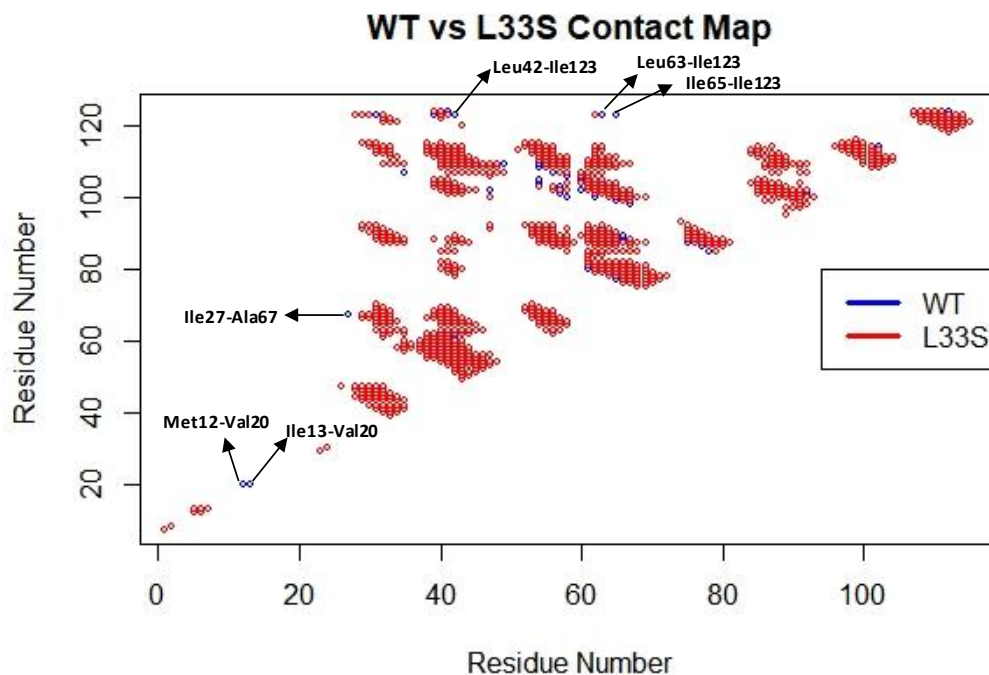
Replacement of the hydrophobic amino acid, Leu with the polar amino acid, Ser, at the position of 33 (L33S mutation) increased total number of contacts, by eight, as compared to the WT (Table 3.7). Although contacts with the residue Leu33 in WT and Ser33 in the mutant did not change, 45 mutant specific contacts, which are not present in the WT structure, appeared in the monomeric structure of the L33S mutant protein (Appendix I, Table I.5). The L33S mutation in the  $\beta$ 2 strand, although did not affect the  $\beta$ -sheet itself as observed by the 3-D structure superimposition it

imparted slight disposition of the adjacent loop ( $\beta 2/\beta 3$  loop) and surrounding residues. This structural change can be a result of increased probability of forming residue contacts among the alpha carbon atoms within the 8 Å°. Similar results are also seen in the R-R distance map, which shows the L33S mutant monomer to be more compact than the WT. In the initial part of the NTD, new contacts between the alpha carbon atom of Met1-Phe7, Ile5-Met12, Lys6-Met12, and Lys6-Ile13 were formed after the L33 to S mutation. This could be related to the differences in the coil structure of NTD of WT and mutant protein, as shown by superimpositions. Furthermore, contact map analysis showed that most of the contacts at the position 33 were found between hydrophobic residue pairs (*e.g.*, Leu, Val, Ile, Ala). In addition to alpha carbon atom contacts between Leu33-Val41 and Leu33-Leu42 in the WT and Ser33-Val41 and Ser33-Val42 in the mutant, additional contacts were found between these residues; alpha carbon of the former and oxygen of the side chain of the latter residue as revealed by DSV analysis.

When we compared two contact maps of WT and L33S, contacts between hydrophobic amino acids, Met12-Val20, Ile13-Val20 at the NTD of the *Tpv* sHSP14.3 WT protein and Ile27-Ala67, Leu42-Ile123, Leu63-Ile123, and Ile65-Ile123 at the ACD-CTD of the protein were lost after the L33S mutation (Figure 3.179). Since hydrophobic amino acids in the core of thermophilic protein are crucial for adapting to higher temperatures (Sælensminde *et al.*, 2009), this might be the reason of a slight decrease in the thermal stability of the L33S protein. Moreover, these results can suggest that single amino acid substitution can change short, medium and long-range interactions.

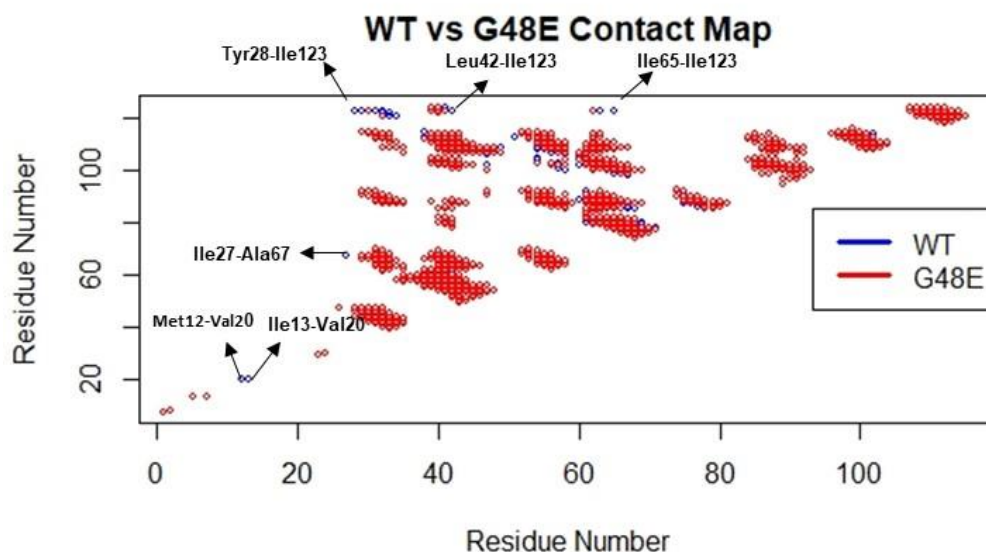
By G48E single mutation, introduction of negatively charged residue reduced the total number of contacts from 831 to 820 within the monomeric structure (Table 3.7). Different from WT, few contacts between residues lying at the NTE and at the ACD (*e.g.*, Met12-Val20, Ile13-Val20 and Ile27-Ala67) together with some contacts of the Ile123 at the CTD (*i.e.*, Tyr28-Ile123, Pro29-Ile123, Leu42-Ile123 and Ile65-Ile123) were not found in G48E mutant sHSP (Figure 3.180, Appendix I, Table I.7).





**Figure 3.179. Contact map analysis of WT and L33S mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.

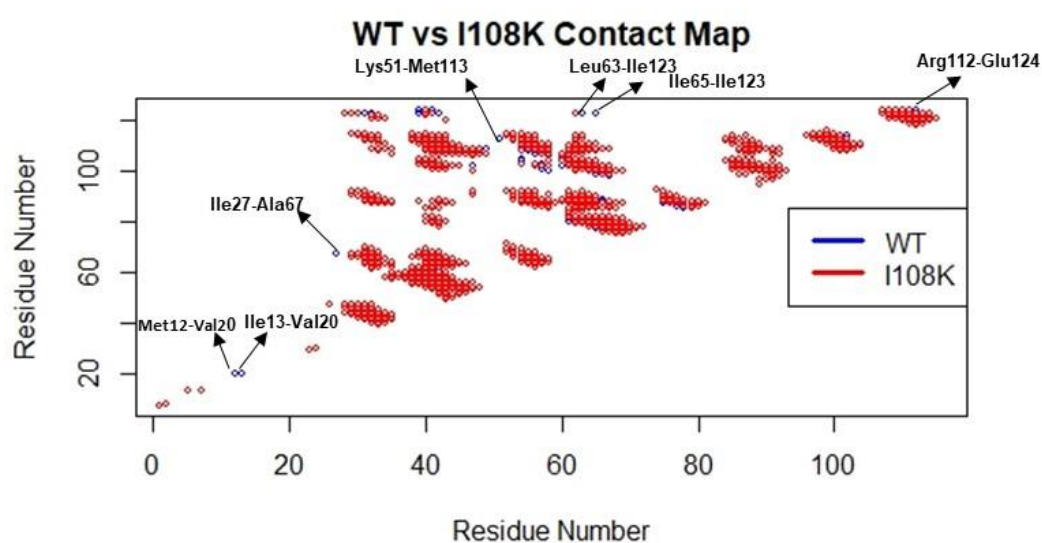
This may imply that G48 to E mutation could have reduced the long range interactions between the alpha carbon atom of the residues. Indeed, 48 WT specific contacts, 37 of them involved in long range interactions; while, G48E mutant formed 18 long range interactions out of 35 mutant specific contacts (Appendix I, Table I.7). For example, Tyr28 in the WT made long range contact with the Ile123, but after mutation, Tyr 28 form a medium range contact with the Ala44 and Met46. In the literature, it was mentioned that long range interactions are important for the protein thermodynamic stability (Bigman & Levy, 2020). On the other hand, at the position 48, WT and the mutant have same interactions, but the probability scores have changed. Although specific contacts vary, the number of the contacts in the NTD of both WT and the mutant were similar. Therefore, 3-D spatial arrangement of NTD of mutant seemed less affected by the mutation, as revealed by comparison of mutant structure with the WT.



**Figure 3.180. Contact map analysis of WT and G48E mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.

Moreover, contact map analysis of I108K mutant showed that this mutation increased the total number of contacts by nine within the monomeric structure (Table 3.7). Out of 839 contacts in the I108K mutant, 45 of them were not found in WT (Appendix I, Table I.9). Comparative analysis of two contact maps showed the differences between WT and I108K mutant. Some of the specific contacts for WT were marked in the Figure 3.181. Besides, Leu33, Tyr86 and Val121 made contacts with the residue K108 in mutant and not in the WT, while long range inter-residue interaction between the alpha carbon atom of the Lys55 and Ile108 present in the WT disappeared after the mutation (Appendix I, Table I.10). It was noted that I108K mutation decreased the long-range interactions such that 50% of the mutant specific contacts were long range interactions; while, WT structure has 70% long range inter-residue contacts. Similar to the G48E mutation, replacement of hydrophobicity by charge (I108K) had similar consequences that was an increase in the medium and short-range contacts and decrease in the number of the long-range contacts. This leads to decreased thermodynamic stability. The superimposition showed that initial part of the NTD coils towards the helix in the mutant that could be due to newly

formed alpha carbon atom interaction between, residue 1 and 7. On the other hand, the mutant specific interaction between Lys108 and Val121 residues might have a pulling effect on the CTD, thus changing its spatial arrangement. Also, as a different from WT, in the mutant the residues Arg81 and Asn120 interact with residues Leu42 and Glu43, respectively. On the other hand, the short-range interaction between residue Ile13 and Val20, in NTD of WT was absent in the mutant (Figure 3.181). Such changes in the contact interactions may be the reason of the slight shift of the initial NTD part towards the dimer interface.



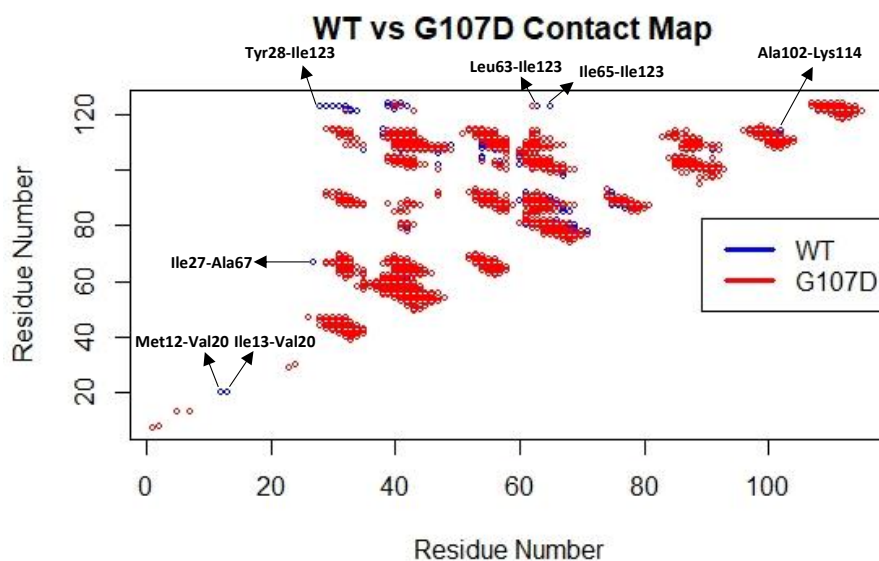
**Figure 3.181. Contact map analysis of WT and I108K mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.

Besides, introduction of the negative charged residue at the highly conserved position 107 (G107D mutation) resulted in the decrease in the number of the contacts by 17 (Table 3.7) and 39 of the alpha carbon atom contacts were only found in the G107D mutant (Appendix I, Table I.11). On the other hand, the WT structure had 57 different contacts, which were absent in the mutant sHSP. Similar to the other *Tpv* ACD mutants with gaining of charge (G48E and I108K), short range contact between the Met1 and Phe7 was found only for the mutant, which was positioned in the beginning of the NTD domain. On the other hand, the WT structure had two specific

short-range interactions in the middle part of the NTD (*i.e.*, Met12-Val20 and Ile13-Val20). When the mutant structure was superimposed with the WT structure, differences at the proximal of the unstructured NTD region were found pronounced (Figure 3.182). Although the mutation did not change the contacts between the residue 107 and Ile123 and Glu124, three long range contacts *i.e.*, Gln35-Gly107, Val41-Gly107 and Val56-Gly107 and one medium range contact, Leu91-Gly107, were lost upon the mutation (Appendix I, Table I.11). Elimination of those polar-hydrophobic and hydrophobic-hydrophobic contacts at the targeted mutation point could be reason of the inactivation of this mutant sHSP even at 60°C. Furthermore, 3-D superimpositions showed structural rearrangement at the the  $\beta 8/\beta 9$  loop which become narrow by the G107D mutation. It is possible that change in the residue contacts may alter the structure. Furthermore, the WT monomeric structure had more long-range contacts (19%) than that of the mutant sHSP. The thermodynamic stability programmes that we used revealed a decrease in the thermodynamic stability of the protein. Therefore, long range contacts could be important for the stability of the protein.

It was also noticed that some long-range contacts at the NTD-CTD border of the protein were lost after the mutation. These following contacts were found in the WT between the alpha carbon atom of the residues Ile123 and Tyr28, Pro29, Pro30, Val31, Thr32. Such residue contact alterations could be responsible for the change in the orientation of CTD, as revealed by the 3-D structure superimpositions. Moreover, RRdist map analysis by chimera also showed that the mutant protein became more compact.

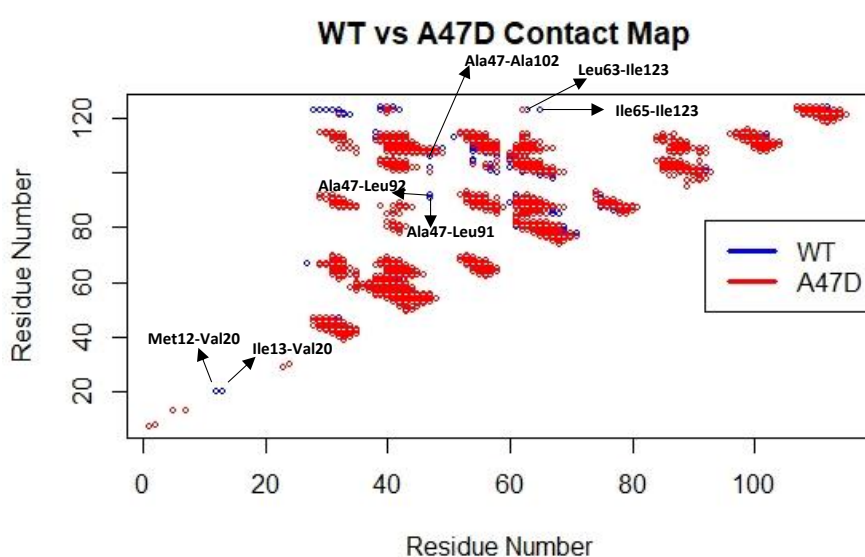
Introduction of the negatively charged residue in the place of hydrophobic amino acid alanine reduced the total number of contacts from 831 to 814 within the monomeric structure (Table 3.7). 40 specific contacts were present in the A47D mutant protein but not in the WT. Half of those contacts were long range, 35% of them were medium range contacts and 15% of them formed short range contacts; while, WT had 57 specific contacts and 70% of them were found to be in the long-range interactions (Appendix I, Table I.13).



**Figure 3.182. Contact map analysis of WT and G107D mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.

Therefore, A47D mutation caused a decline in the long-range interactions. In the WT sHSP, Ile123 which is located at the CTD makes contacts with the amino acid clusters (*i.e.*, Tyr28, Pro29, Pro30, Val31 and Thr32) in the NTD, as well as the residues lying at the ACD (*i.e.*, Leu63 and Ile65). These contacts are not found in the A47D mutant sHSP. Some of these contact pairs shown in the Figure 3.183. Structural alterations noticed at the CTD of the protein besides, in the first helical coil part of the NTD, could be consequences of the new contacts formed between the Met1 and Phe7 and the residue contacts lost between Met12-Val20 and Ile13-Val20 after the mutation. Among the 11 long range interactions that involves Ala47 in the WT, four made interactions with the alpha carbon atom of the Ala102, Pro92, Leu91 and Glu106 were lost after the mutation. The Asp47-Lys55 contact was a new mutant specific interaction (Appendix I, Table I.14). Inter-residue interactions are known to be important for protein folding and to maintain the stability of the folded proteins, among which hydrophobic interactions are the major forces dominating the stability of the protein (Gromiha & Selvaraj, 2004). Therefore, lose of long-range hydrophobic contacts could be a reason for low heat stability of the mutant protein

at 60 °C as compared to the WT. Also, thermodynamic stability of this mutant was found to be decreased by Dynamut2 programme. By the loss of long range interactions, decrease in the protein stability may be due to increased entropy of unfolded (Bigman & Levy, 2020). Increased compactness of the sHSP protein structure by mutation as revealed by RRdistance map analysis should be closely related to decreased distances between residues. All these changes in the structure induced by mutation may account for a decrease in chaperone activity of the mutant sHSP as compared to WT protein.



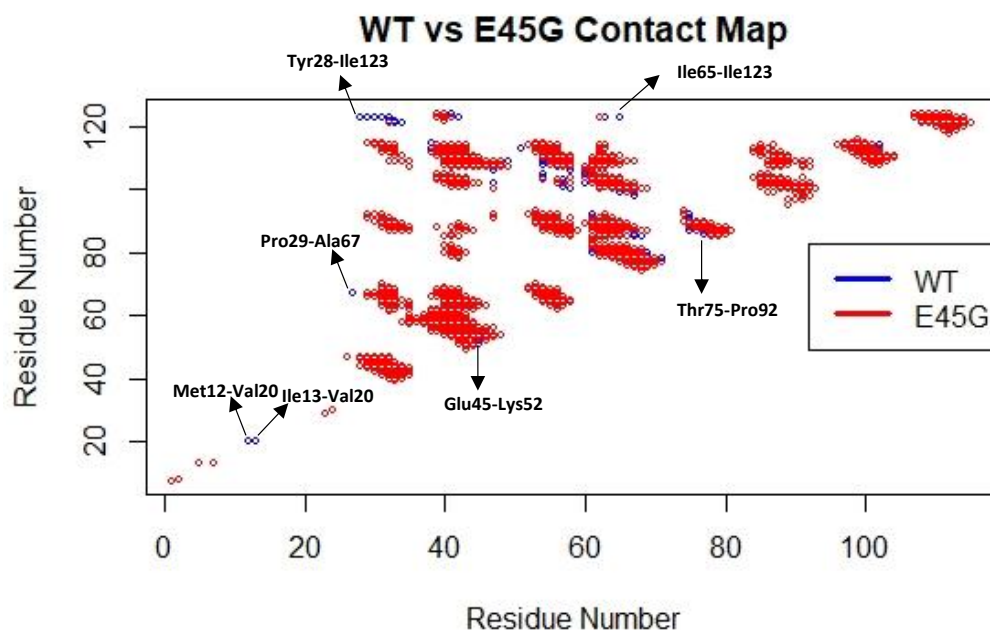
**Figure 3.183. Contact map analysis of WT and A47D mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.

Ten residue contacts were lost as a result of replacement of the negatively charged residue by neutral glycine (E45G mutation) (Table 3.7). This exchange between amino acids with different physicochemical property altered the contact preferences within the monomeric structure. As a result, 39 new contact pairs were formed, which were not found in the WT structure (Appendix I, Table I.15). At the targeted mutation position, the residue 45 formed contact between the alpha carbon atom of the following residues: Lys51, Asn53, Lys55, Val56, Leu63, Thr64, Ala102, Gly107, and Ile108 both in the WT and the E45G mutant sHSP. The contact with

Lys52 available in the WT-sHSP disappeared after the mutation (Appendix I, Table I.16). According to the 3-D model structure analysis, this point mutation did not affect the structure at the targeted position (at the end of the  $\beta$ 3), but alterations in the NTD and CTD of the protein were noticed. Beyond the mutation site, Met12-Val20, Ile13-Val20 at the NTD and several contacts of the Ile123 with Tyr28, Pro29, Pro30, Val31 at the CTD of the sHSP protein were also lost after the mutation (Figure 3.184). But, majority of the mutant specific contacts were formed between the residues located at the ACD of the protein (*e.g.*, Tyr34-Val111, Gln35-Thr64, Ser38-Ser66, Leu40-Gln80) (Appendix I, Table I.15). Therefore, the changes in the positioning of the NTD and CTD could be related with change in the amino acids that form contacts with each other. Some of WT specific contacts were labelled and shown in the Figure 3.184.

Moreover, E45G mutation reduced the long-range contacts by 20% as compared to the WT. As mentioned for the other mutants (*e.g.*, G48E, G107D, I108K), a decrease of the long-range contacts is accompanied by the decrease in the thermodynamic stability, as found by Dynamut and MUpro analysis. RRDist map analysis also indicated the reduction in the overall distance of the mutant protein. It is likely that such alterations can have correlation with decreased chaperone activity of the mutant protein and its low heat stability.

Among the all mutations, E43V mutation reduced the total number of contacts the most in the monomeric structure of the *Tpv* sHSP 14.3 (Table 3.7). Out of total 788 contacts in the E43V mutant monomeric structure, 54 of them were not present in WT. List of the residues forming contacts with E43 in the WT sHSP and V43 in the mutant are given in the Appendix I (Table I.17-I.18). Different from WT *Tpv* sHSP14.3, at the position of 43, E substituted residue Valine formed contacts with the alpha carbon atom of the Asn59, Asp79, and Tyr104 while Glu43-Phe49 and Glu43-Lys60 contacts were available only in the WT structure. When comparing the inter-residue contacts of the WT and the mutant protein, 97 contacts available in the WT were not found in the mutant sHSP (Appendix I, Table I.17).

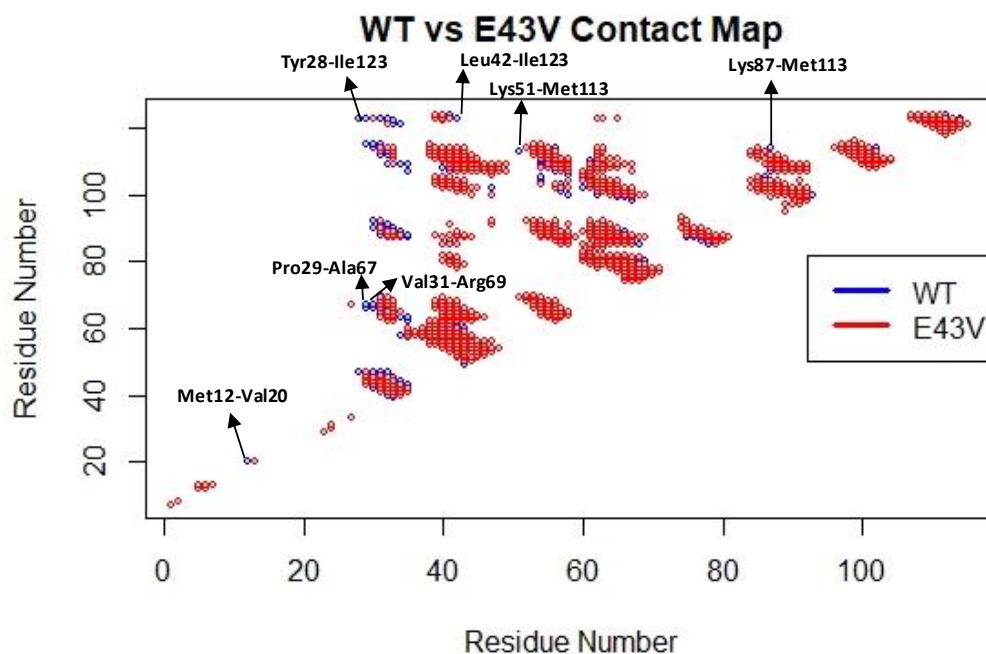


**Figure 3.184. Contact map analysis of WT and E45G mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.

Majority of the disappeared contacts included the long-range interactions between hydrophobic, charged or polar residues, which was counted as 67 contacts that was equivalent to the 70% of the WT specific contacts. Different contacts were found to be distributed at residues lying between NTD-CTD of the protein (*e.g.*, Tyr28-Ile123, Pro29-Ile123, Pro29-Thr115, Pro30-Thr115, Val31-Ile123), and between NTD-ACD of the protein (*e.g.*, Pro29-Ala67, Pro29-Ser66, Pro30-Lys90, Val31-Lys88, Val31-Lys70), and between ACD-CTD of the protein (*e.g.*, Leu33-Glu122, Leu33-Val121, Val41-Glu124, Leu42-Ile123) (Figure 3.185). Such different number of contacts between the WT and E43V mutant may induce structural alterations. 3-D superimposition of the E43V mutant structure with the WT showed that positioning of unstructured part of the NTD and the CTD changed, while the point mutation site remained unchanged ( $\beta 3$  strand). RRDist map analysis predicted that the mutation reduced the overall distance within the monomeric structure. The loss of long-range interactions by mutation also should have impact on thermodynamic stability of the protein through destabilization of its structure ( $\Delta\Delta G$  value, -1.06



as revealed by Dynamut2 programme). Besides, the mutant protein has not been efficient in protection of the model substrates, CS and ADH against heat inactivation.

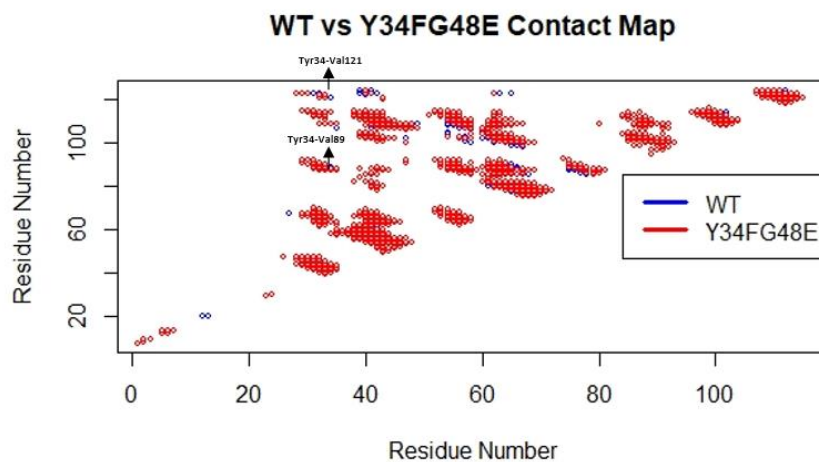


**Figure 3.185. Contact map analysis of WT and E43V mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.

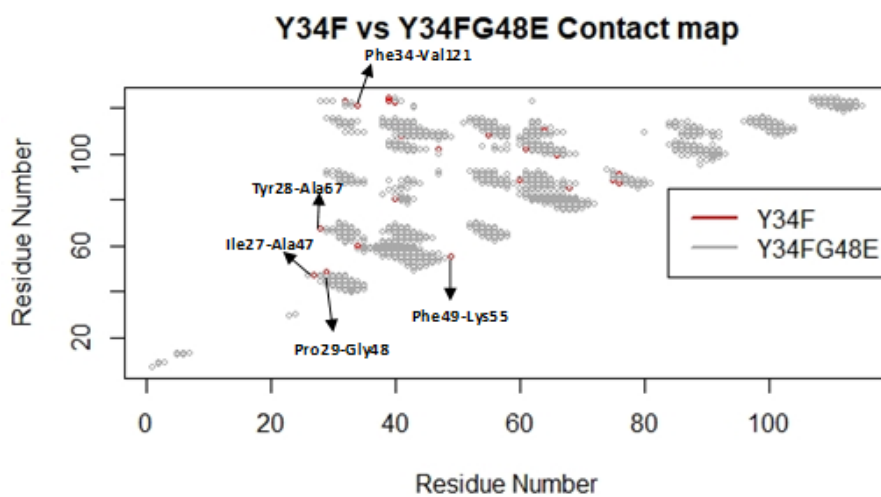
Although double mutation Y34FG48E did not change the total number of contacts in the structure, type of interactions and long-range interactions were changed and the long-range interactions decreased between the residues. The WT protein had 49 specific contacts and 34 of them (70%) participated in the formation of the long-range interactions, which were not present in the double mutant (*e.g.*, Ile27-Ala67, Val31-Ile123, Thr32-Ile123, Asp39-Glu124, Asp39-Ile123, Val41-Gly107) (Appendix I, Table I.19). On the other hand, out of total 50 specific contacts of the double mutant, half of them formed long range contacts (*e.g.*, Pro29-Leu91, Pro29-Pro92, Ser38-Ser66, Ser38-Val82, Asp39-Lys84, Glu43-Gln80, Gln80-Leu109). While the contact pairs between Tyr34-Val89 and Tyr34-Val121 disappeared (Figure 3.186), a new contact of Phe34-Ser66 pair formed in the Y34FG48E double

mutant (Appendix I, Table I20). Therefore, decreased number of long-range contacts including hydrophobic ones within the Y34FG48E monomer could be the reason of reduced thermal stability of the mutant protein as compared to the WT. As a result of this mutation, 3D structure of the double mutant protein, NTD,  $\beta 2/\beta 3$  loop and  $\beta 6$  loop of the ACD were changed. Furthermore, Y34FG48E double mutation reduced the overall RRDist of *Tpv* 14.3 sHSP. Such structural alteration can be attributed to the changed inter-residue contacts occurred in the NTD (*i.e.*, Met1-Phe7, Tyr2-Thr9, Thr3-Thr9, Ile5-Met12, Lys6-Met12, Lys6-Met12, Lys6-Ile13), and in the ACD (contacts of Ser38 on the  $\beta 2/\beta 3$  loop with Ser66, Val82, Lys 114) of the double mutant protein. Also, some contacts specific to the Y34FG48E mutant sHSP were found in the  $\beta 6$  loop of the ACD *i.e.*, Arg69-Arg81, Lys70-Asp79, Arg71-Asp79, Glu72-Ile78, Gln80-Val88, and Gln80-Leu109. As reported previously, changes in the loops could have impact on the thermodynamic and kinetic properties of the proteins (Gavrilov *et al.*, 2015; Bigman & Levy, 2020).

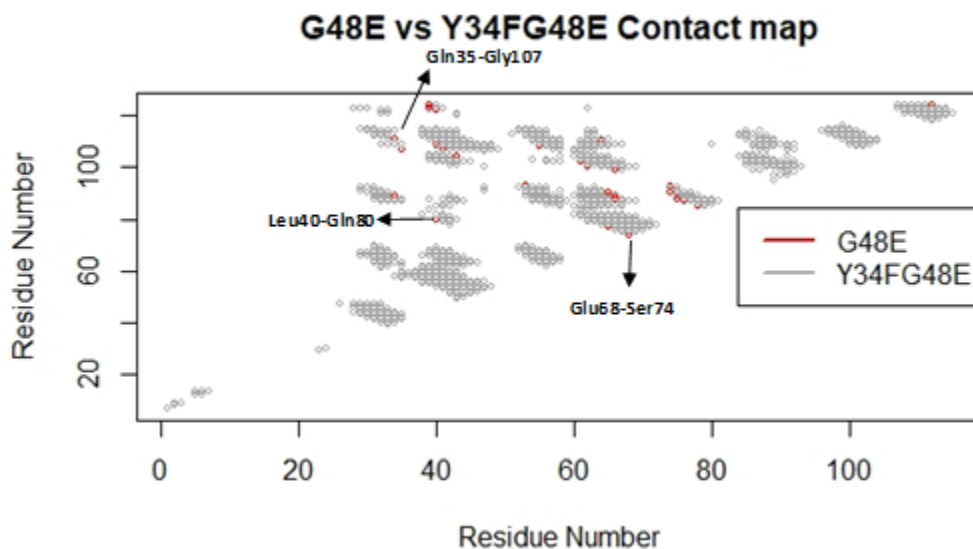
Comparing double mutant with its single mutant variants (*i.e.*, Y34F and G48E) showed that their amino acid preferences in formation of residue contacts were different from each other. A number of residue contacts were found in the single mutants which were not present in the double mutant as shown in the graph with different color where few are also labelled (Figure 3.187). The mutations were positioned in the proximal part of the ACD and the residue contact differences were also seen in the nearby region. Our contact map analysis showed that long range contacts (*e.g.*, Phe34-Val121, Ile27-Ala47, and Tyr28-Ala67) which were seen in the single mutant Y34F disappeared upon double mutation. Furthermore, different from G48E, double mutation resulted in the elimination of the polar-polar or polar charged interactions (Figure 3.188).



**Figure 3.186. Contact map analysis of WT and Y34FG48E mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.



**Figure 3.187. Contact map analysis of Y34F and Y34FG48E mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.



**Figure 3.188. Contact map analysis of G48E and Y34FG48E mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.

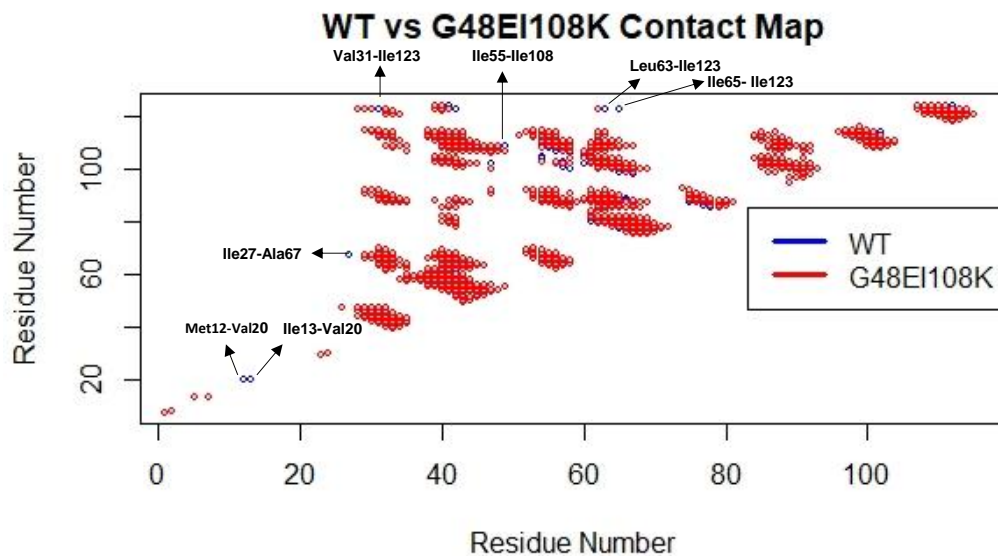
In other double mutant where the mutated residues are far apart from each other (G48EI108K), specified mutations increased the total number of contacts from 831 to 839 in the monomeric structure but there has been reduction in the number of the long-range interactions from 64 % to 44 %. The double mutant sHSP had 15% more medium and 5 % more short range interactions than the WT structure (Appendix I, Table I.21). This may be due to the low propensity of the Glu and Lys for long range contact. As reported before, charged residues formed minimum number of long range contacts in other proteins, as well (Gromiha & Selvaraj, 1997).

Contacts between C $\alpha$  atoms of the amino acids at the residue 48 did not change in both WT and G48EI108K mutant. On the other hand, at the position 108, few differences were observed between WT and double mutant. Residue-residue contacts between Leu33-Lys108 and Val121-Lys108 occurred only in the double mutant, but contacts seen in WT between Ile55-Ile108 were lost after double mutation (Appendix I, Table I.22). When we compare the contact maps of WT and G48EI108K, contacts between hydrophobic amino acids present in the WT were lost after the double mutation (*i.e.*, Met12-Val20, Ile13-Val20 in the NTD and Val31-Ile123 at the NTD-

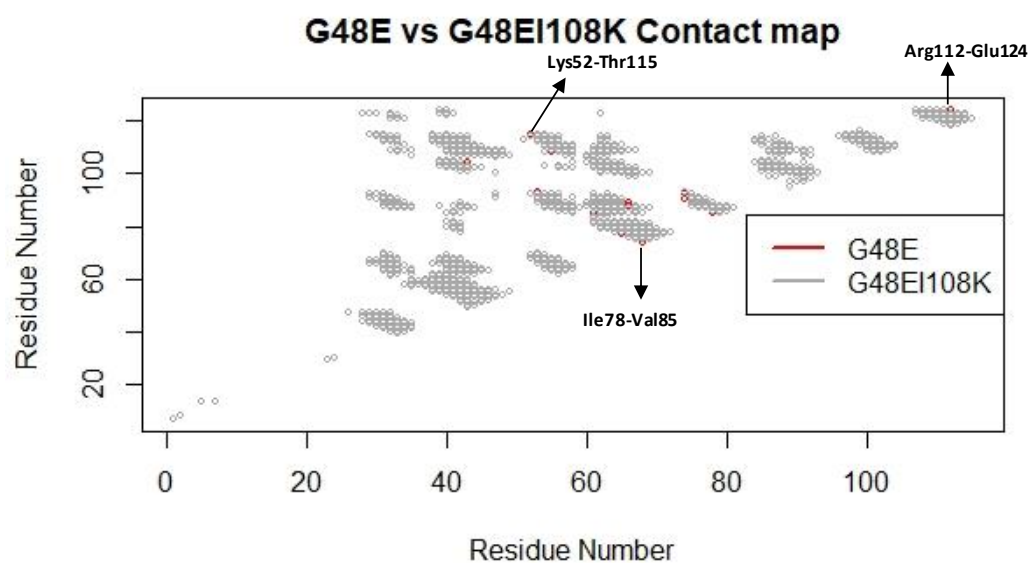
CTD interface, and Leu42-Ile123, Leu63-Ile123, and Ile65-Ile123 at the ACD-CTD interface of the protein) (Figure 3.189). In thermophilic proteins, the strength of the contacts between hydrophobic residues and also between polar-charged residues are thought to be one of the factors contributing to the thermal stability of these protein (Goldstein, 2007). Therefore, low heat stability of the double mutant as compared to the WT at 70°C could be due to the loss of the essential hydrophobic contacts, which might be involved in the thermal stability of the protein.

G48E I108K double mutation significantly reduced the RRDist of the mutant protein and there were alterations mainly in the size of  $\beta 2/\beta 3$  loop and  $\beta 8/\beta 9$  loop in the ACD besides positioning of CTD. These structural alterations could be arisen from the changes between the inter-residue contacts occurred in the ACD of the mutant protein. The residues found in the  $\beta 2/\beta 3$  loop of the ACD formed contacts between Gln35-Ser57, Gln35-Thr64, Ser38-Ser66, Ser38-Lys114, which were not found in the WT. Three contacts, Ile54-Ser105, Val56-Gly107, Val58-Glu106 that were found in the  $\beta 8/\beta 9$  loop were lost after the double mutation and this may be a reason of structural change in that loop. The increased compactness of the G48E I108K mutant protein could be related to reduced loop size. It is known that thermodynamic and kinetic properties of the proteins can be affected from the loop size (Gavrilov *et al.*, 2015; Bigman & Levy, 2020). Such compact structure may account for the better chaperone activity against thermal aggregation of ADH as compared to the WT and its single variants.

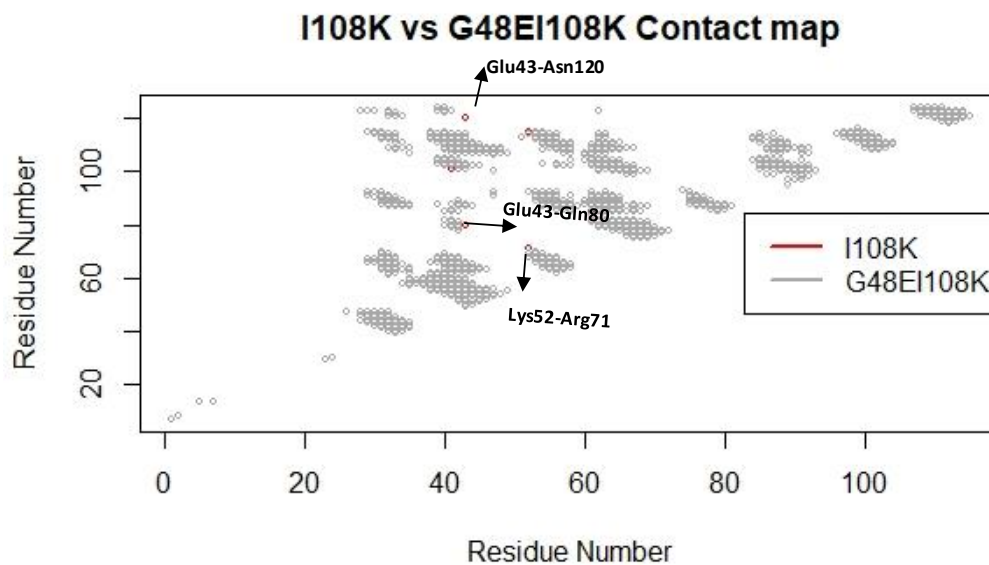
Comparison of contact map of the double mutant G48E I108K and G48E single variant showed small difference in forming contacts (Figure 3.190). Additive effect of second mutation at position 108 caused loss of some more contact pairs which are mostly located in the ACD of the G48E protein (*i.e.*, Lys52-Thr115, Ile78-Val85, Arg112-Glu124). Furthermore, few contact differences were observed between the single mutant I108K and double mutant G48E I108K. In addition, polar-charged interactions around position 48 such as Glu43-Asn120, Glu43-Gln80, and Lys52-Arg71 were not found for double mutant G48E I108K (Figure 3.191).



**Figure 3.189. Contact map analysis of WT and G48E108K mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.



**Figure 3.190. Contact map analysis of G48E and G48E108K mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.

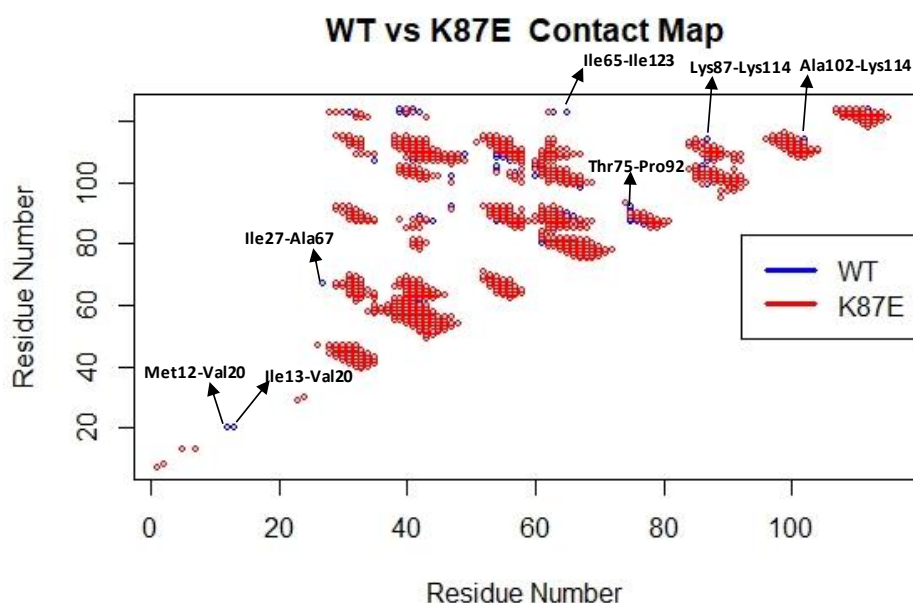


**Figure 3.191. Contact map analysis of I108K and G48E I108K mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.

The contact map analysis of K87E mutant showed that this mutation decreased the total number of the contact from 831 to 818 within monomeric structure (Table 3.7). The mutation led to the formation of the 33 new contacts, which was not found in the WT structure (Appendix I, Table I.23). The contacts specific to the K87E mutation included 45% long range interactions, 40% medium range contacts and 15% short range contacts. In the WT sHSP, among 46 WT specific contacts, long range contacts and medium range contacts constituted the 63% and 28% of the interactions, respectively. This result can imply that substituting positively charged residue by a negatively charged glutamic acid decreased the total number of the long-range interactions. Some of the lost long-range contacts were at the CTD of the *Tpv* sHSP 14.3 with the residue Ile123 (*i.e.*, Val31-Ile123; Asp39-Ile123; Leu42-Ile123; Leu63-Ile123; Leu63-Ile123; Ile65-Ile123). This could be responsible for the alteration in the positioning of the CTD of the mutant protein, as revealed by the 3-D structure superimpositions. Besides, Thr75 on the  $\beta_6$  strand of the WT *Tpv* sHSP14.3 formed three polar-hydrophobic interactions between Pro92, Lys87 and

Val88 residues, and Val76 and Tyr77 on the same strand made contacts with the Lys87, and Tyr86 respectively. Those interactions were lost after the mutation. Changes of this type amino acid contacts on the  $\beta_6$  strand may account for the alterations in the loop, as revealed by the structure superimpositions.

Few contacts between residues lying at the NTD and at the ACD (e.g., Met12-Val20, Ile13-Val20 and Ile27-Ala67) were lost after K87E mutation (Figure 3.192). At the position 87, long range interactions with residues Ile108, Leu109, Thr110, Val111 and Arg112 were observed both in the WT and K87E mutant sHSP. However, long-range interaction of the residue K87 and other residues with K114, Asp99, Gly107 available in the WT *Tpv* sHSP14.3 were not found in the mutant protein. As in G48E mutant, a decrease in the number of the long-range contacts in the K87E mutant protein led to the decrease in the overall RRdist of the sHSP protein and in its thermodynamic stability.



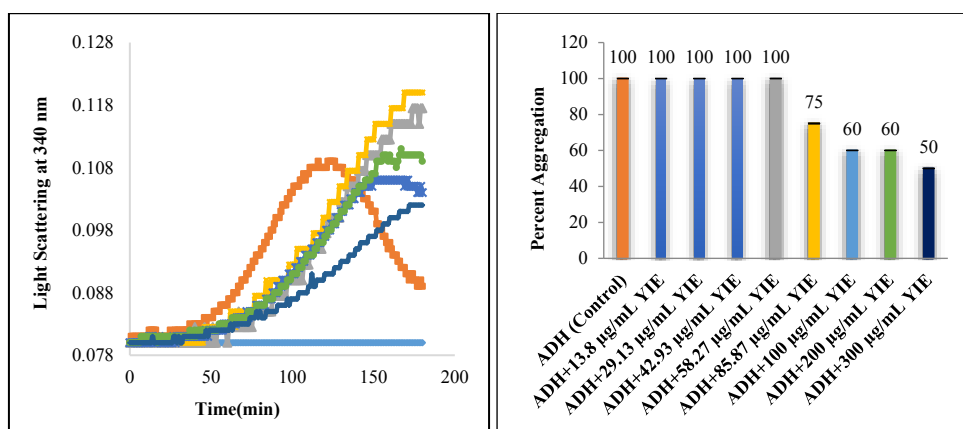
**Figure 3.192. Contact map analysis of WT and K87E mutant.** x- and y- axes indicate the residue numbers (1 to 124) in primary sequence of *Tpv* sHSP14.3. Each dot marks the residues in contact when the distance threshold between alpha carbon atom was 8 Å.



### 3.23 Chaperone Activity of *Tpv* ACD Derived Mini Chaperones

Function of the mini chaperone peptides to prevent protein aggregation was investigated by using the model substrate proteins, yeast ADH and pig heart Citrate Synthase as described in the Material and Methods. Four synthetic peptides named as YIE, GIK, LVD and SDS, which were designed based on different regions of the ACD, were used in the experiments.

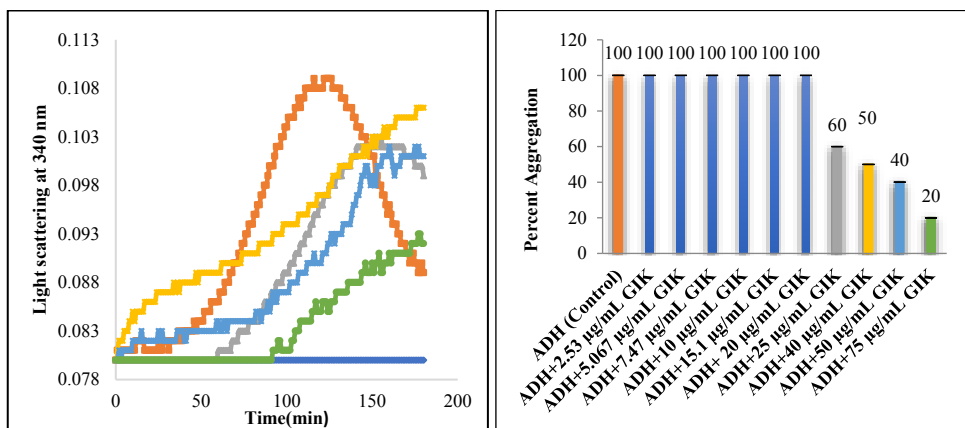
The ability of the four peptides to protect the heat induced aggregation of ADH was measured by thermal aggregation assay at 43°C. Results showed that YIE mini peptide prevented the ADH aggregation in a concentration dependent manner. At the concentrations between 14 µg/mL to 58 µg/mL, the YIE peptide, could not prevent ADH aggregation. Above 85.27 µg/mL concentrations, this peptide protect the ADH enzyme against heat induced aggregation significantly. The highest protection (50 %) was observed in the presence of 300 µg/mL of the YIE peptide (Figure 3.193).



**Figure 3.193. Suppression of ADH aggregation in the absence or presence of YIE mini-chaperone.** ADH aggregation was monitored by measuring light scattering at 340 nm at 43°C in the absence or presence of 58-300 µg/mL YIE. (—◆—) Blank, (—■—) ADH control, (—▲—) ADH+58.27 µg/mL YIE, (—×—) ADH+85.87 µg/mL YIE, (—\*—) ADH+100 µg/mL YIE, (—●—) ADH+200 µg/mL YIE, (—+—) ADH+300 µg/mL YIE. Percent aggregation when the aggregation of ADH alone (Control) was taken as 100 %. Each experiment was repeated at least three times.

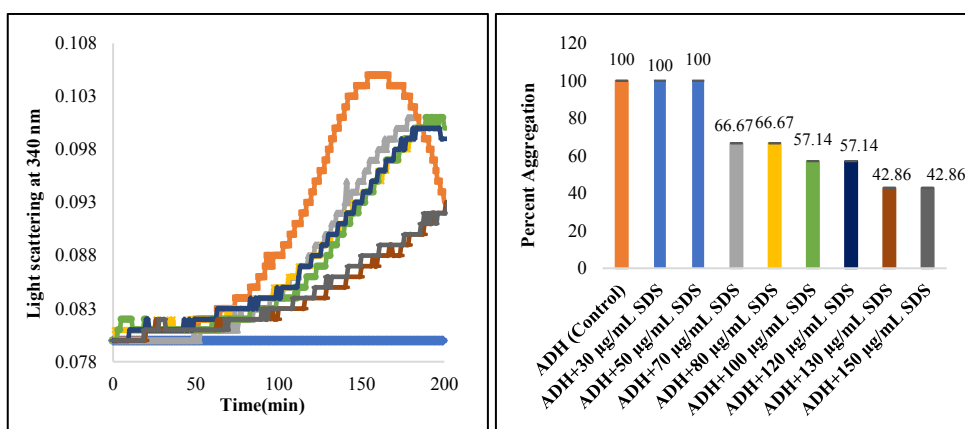
Like YIE mini peptide, GIK peptide also showed good chaperone activity at higher concentrations. The GIK peptide did not protect the ADH from heat aggregation at concentrations varying from 3 to 20 µg/mL. However, at concentrations  $\geq 25$  µg/mL

(i.e., 25-75  $\mu\text{g/mL}$ ) supplementation of the GIK peptide suppressed the ADH aggregation to 80% (Figure 3.194).



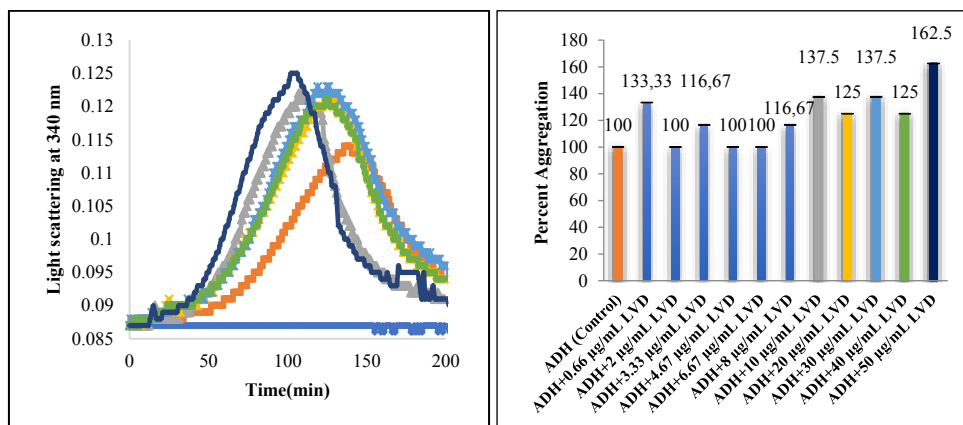
**Figure 3.194. Suppression of ADH aggregation in the absence or presence of GIK mini-chaperone.** ADH aggregation was monitored by measuring light scattering at 340 nm at 43°C in the absence and presence of 25-75  $\mu\text{g/mL}$  GIK mini-peptide. (—◆—) Blank, (—■—) ADH control, (—▲—) ADH+25  $\mu\text{g/mL}$  GIK, (—✕—) ADH+40  $\mu\text{g/mL}$  GIK, (—✱—) ADH+50  $\mu\text{g/mL}$  GIK, (—●—) ADH+75  $\mu\text{g/mL}$  GIK. Percent aggregation when the aggregation of ADH alone (Control) was taken as 100 %. Each experiment was repeated at least three times.

When the dose dependent chaperone activity of SDS peptide was assessed, it was found that heat induced aggregation of ADH was not prevented at concentrations 30-50  $\mu\text{g/mL}$  of the mini peptide. At concentrations  $\geq 70$   $\mu\text{g/mL}$ , SDS peptide displayed more efficient chaperone activity. The ADH aggregation could be suppressed by 58 % in the presence of 149  $\mu\text{g/mL}$  of this peptide (Figure 3.195).



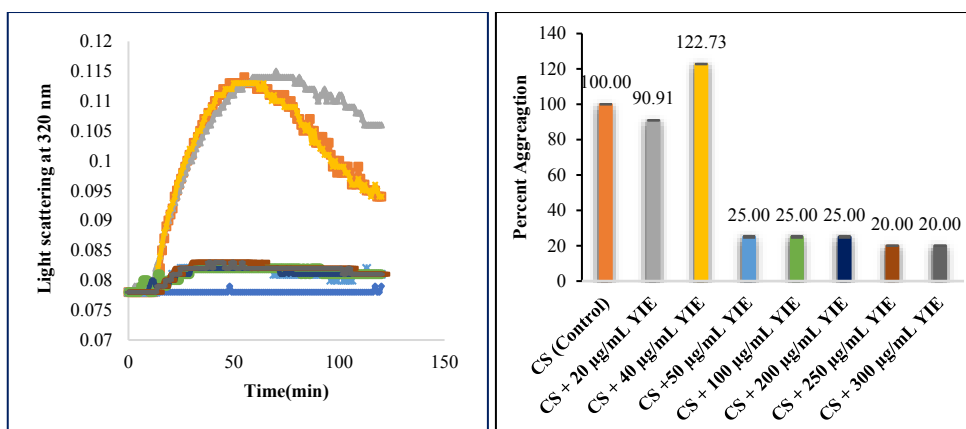
**Figure 3.195. Suppression of ADH aggregation in the absence or presence of SDS mini-chaperone.** ADH aggregation was monitored by measuring light scattering at 340 nm at 43°C in the absence and presence of 70-150  $\mu\text{g/mL}$  SDS mini-peptide. (—◆—) Blank, (—■—) ADH control, (—▲—) ADH+70  $\mu\text{g/mL}$  SDS, (—✕—) ADH+80  $\mu\text{g/mL}$  SDS, (—●—) ADH+100  $\mu\text{g/mL}$  SDS, (—✱—) ADH+120  $\mu\text{g/mL}$  SDS, (—■—) ADH+130  $\mu\text{g/mL}$  SDS, (—■—) ADH+150  $\mu\text{g/mL}$  SDS. Percent aggregation when the aggregation of ADH alone (Control) was taken as 100 %. Each experiment was repeated at least three times.

Among the four mini-peptides, LVD peptide was the only one that induced aggregation of ADH over the time of incubation at 43°C. The extent of the ADH aggregation was increased more with high dose of the LVD peptide. Neither low amount nor high amount of peptide had an effect on reducing aggregation of ADH. This may imply that peptide itself aggregated with the substrate (Figure 3.196).



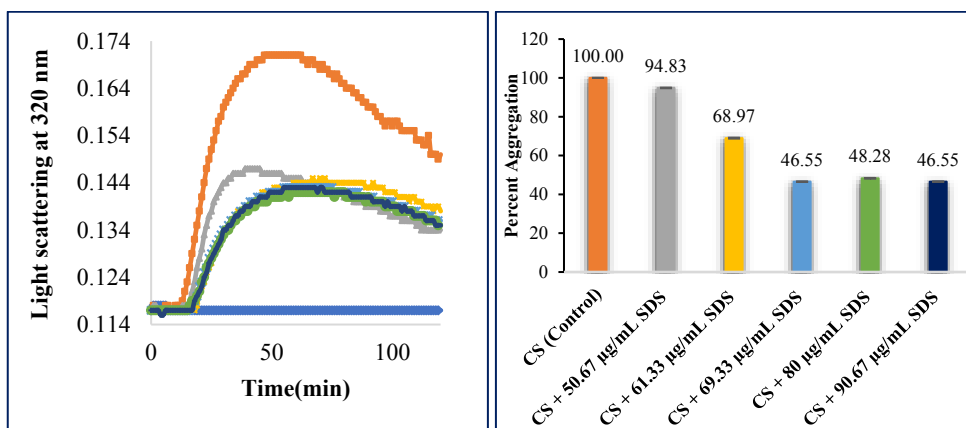
**Figure 3.196. Suppression of ADH aggregation in the absence or presence of LVD mini-chaperone.** ADH aggregation was monitored by measuring light scattering at 340 nm at 43°C in the absence and presence of 10-50 µg/mL LVD mini-peptide. (—◆—) Blank, (—■—) ADH control, (—▲—) ADH+10 µg/mL LVD, (—×—) ADH+20 µg/mL LVD, (—\*—) ADH+30 µg/mL LVD, (—●—) ADH+40 µg/mL LVD, (—+—) ADH+50µg/mL LVD. Percent aggregation when the aggregation of ADH alone (Control) was taken as 100 %. Each experiment was repeated at least three times.

Thermal aggregation assay using four synthetic ACD peptides were also performed with citrate synthase as the model substrate. Our results showed that YIE peptide efficiently suppressed the aggregation of CS at concentrations 50-300 µg/mL and almost 75-80 % protection was achieved within this range of the YIE concentration. When the amount of the peptide used was less than the 50 µg/mL, it did not display an effective chaperone function (Figure 3.197).



**Figure 3.197. Suppression of CS aggregation in the absence or presence of YIE mini-chaperone.** CS aggregation was monitored by measuring light scattering at 320 nm at 45°C in the absence and presence of 20-250 µg/mL YIE mini-peptide. (—◆—) Blank, (—■—) CS control, (—▲—) CS+20 µg/mL YIE, (—×—) CS+40 µg/mL YIE, (—\*—) CS+50 µg/mL YIE, (—●—) CS+100 µg/mL YIE, (—+—) CS+200 µg/mL YIE, (—○—) CS+250 µg/mL YIE, (—◊—) CS+300 µg/mL YIE. Percent aggregation when the aggregation of CS alone (Control) was taken as 100 %. Each experiment was repeated at least three times.

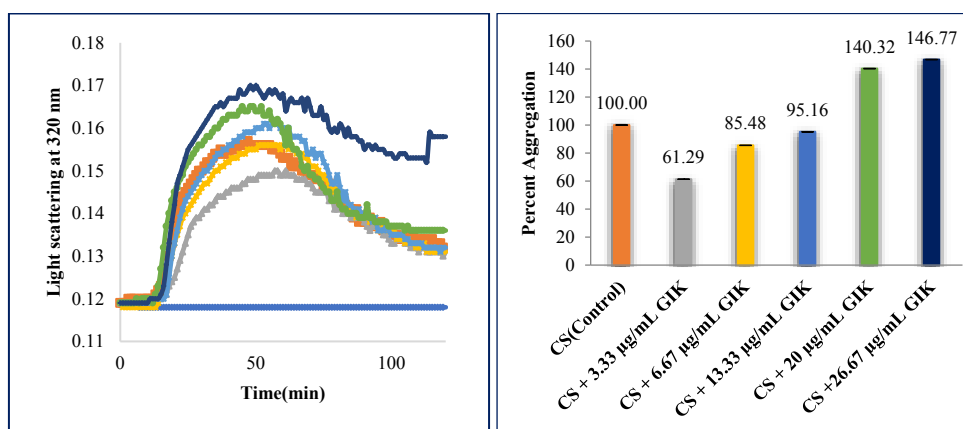
Another mini peptide, SDS, also prevented thermal aggregation of CS more than 50 % at concentrations 70-90 µg/mL (Figure 3.198).



**Figure 3.198. Suppression of CS aggregation in the absence or presence of SDS mini-chaperone.** CS aggregation was monitored by measuring light scattering at 320 nm at 45°C in the absence and presence of 50-90 µg/mL SDS mini-peptide. (—◆—) Blank, (—■—) CS control, (—▲—) CS+50.67µg/mL SDS, (—×—)CS+61.33 µg/mL SDS, (—\*—) CS+69.33 µg/mL SDS, (—●—) CS+80 µg/mL SDS, (—+—) CS+90.67 µg/mL SDS. Percent aggregation when the aggregation of CS alone (Control) was taken as 100 %. Each experiment was repeated at least three times.

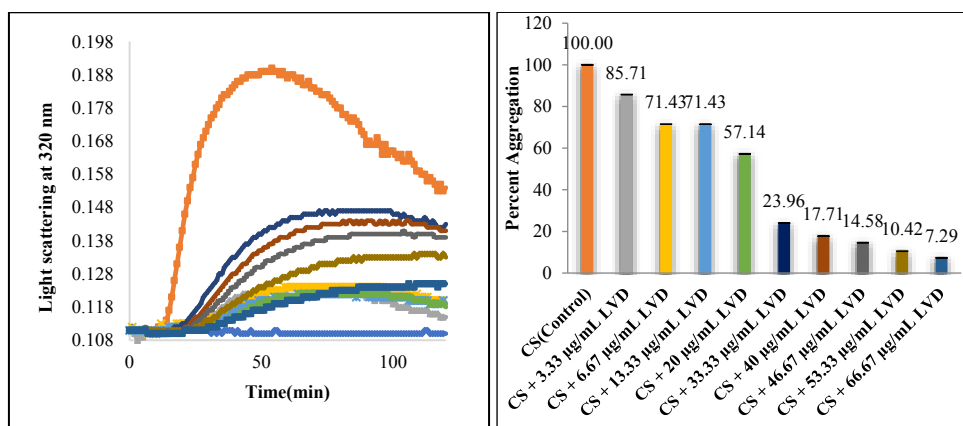
Interestingly, the mini peptide GIK performed better chaperone activity at lower concentrations (< 7 µg/mL). At most 40 % protection efficiency can be attained. At

concentrations  $>13 \mu\text{g/mL}$ , GIK peptide accelerated the heat induced aggregation of CS up to 146 % (at  $27 \mu\text{g/mL}$  concentration) (Figure 3.199).



**Figure 3.199. Suppression of CS aggregation in the absence or presence of GIK mini-chaperone.** CS aggregation was monitored by measuring light scattering at 320 nm at  $45^\circ\text{C}$  in the absence and presence of 3-27  $\mu\text{g/mL}$  GIK mini-peptide. (—◆—) Blank, (—■—) CS control, (—▲—) CS+3.33  $\mu\text{g/mL}$  GIK, (—×—) CS+6.67  $\mu\text{g/mL}$  GIK, (—\*—) CS+13.33  $\mu\text{g/mL}$  GIK, (—●—) CS+20  $\mu\text{g/mL}$  GIK, (—+—) CS+26.67  $\mu\text{g/mL}$  GIK. Percent aggregation when the aggregation of CS alone (Control) was taken as 100 %. Each experiment was repeated at least three times.

On the other hand, the LVD synthetic peptide was found potent chaperone among the others against aggregation of the substrate protein, CS. Its chaperone activity was concentration dependent (Figure 3.200). More than 80 % inhibition of the heat aggregation of the CS was achieved at concentrations of  $\geq 33 \mu\text{g/mL}$ .



**Figure 3.200. Suppression of CS aggregation in the absence or presence of LVD mini-chaperone.** CS aggregation was monitored by measuring light scattering at 320 nm at  $45^\circ\text{C}$  in the absence and presence of 10-50  $\mu\text{g/mL}$  LVD mini-peptide. (—◆—) Blank, (—■—) CS control, (—▲—) CS+13.33  $\mu\text{g/mL}$  LVD, (—×—) CS+13.33  $\mu\text{g/mL}$  LVD, (—\*—) CS+20  $\mu\text{g/mL}$  LVD, (—●—) CS+57.14  $\mu\text{g/mL}$  LVD, (—+—) CS+33.33  $\mu\text{g/mL}$  LVD, (—■—) CS+40  $\mu\text{g/mL}$  LVD, (—■—) CS+46.67  $\mu\text{g/mL}$  LVD, (—◆—) CS+53.33  $\mu\text{g/mL}$  LVD, (—■—) CS+66.67  $\mu\text{g/mL}$  LVD. Percent aggregation when the aggregation of CS alone (Control) was taken as 100 %. Each experiment was repeated at least three times.



## CHAPTER 4

### DISCUSSION

Small heat shock protein *Tpv* sHSP 14.3 from *Thermoplasma volcanium* ( $T_{op}$  60°C; pHop 2.0) was heterologously over expressed in *E.coli* in order to study Alpha Crystallin Domain(ACD) of this protein in maintaining structural integrity of the oligomers and its preferential interactions with client proteins. Our model organism is the one of the most representatives of the thermoacidophilic archaea that sustain their life in extreme environments characterized by low pH and high temperatures (Bertoldo *et al.*, 2004). Therefore, their proteins are both heat and acid stable. In this study, recombinant *Tpv* sHSP 14.3 was subjected to heat treatment at high temperatures (60 °C -80 °C ) and found to be stable at temperatures up to 70°C for 30 min. Similar observations have been also reported for other archaeal sHSP proteins (*i.e.*, *Thermococcus* sp. KS-1 sHSP, *Sulfolobolus solfataricus* sHSP, SsHsp 20.1 and *Sulfolobolus tokodaii* sHSP, StHSP14.0 which could resist the heat treatment at 75°C for 30 min (Usui, 2004; Usui *et al.*, 2001 ; Liu *et al.*, 2015).

To assess the effect of targeted mutations on the chaperone action of the *Tpv* sHSP 14.3 anti-aggregation assays and chaperone activity assays were performed using two model substrate proteins: CS (citrate synthase) and Yeast Alcohol Dehydrogenase (ADH).

According to the CS aggregation assays using WT *Tpv* sHSP 14.3, it was noticed that suppression of aggregation was more effective at lower CS/Chaperone w/w ratios (*i.e.*, 1/2). In literature, it is implicated that inhibition of substrate aggregation by sHSPs occurs at different client:sHSP ratio, depending on the source of the small heat shock proteins (Basha, 2012). Similar to our findings, Hsp42 from *Saccharomyces cerevisiae* exhibited a concentration dependent CS aggregation up to a molar ratio of 1:1 (Hsp42 monomer: CS monomer) and above this ratio suppression of CS was not better (Haslbeck, Braun, *et al.*, 2004). The study

conducted using *Sulfolobus solfataricus* sHSP, StHsp14.0 showed that complete suppression of the CS was achieved in the presence of a 32-fold excess chaperone (Keisuke Usui, Ishii, *et al.*, 2004).

On the other hand, CS activity assay results showed that excess amount of the WT *Tpv* sHSP14.3 (*i.e.*, CS: chaperone w/w ratios of 1/500) were required to restore the CS activity after heat induced inactivation comparable to the positive control. Our research group previously also proved that high amount of *Tpv* sHSP14.3 were required for protection of mesophilic enzymes against thermal inactivation at temperatures far below the optimum growth temperatures of *T.volcanium* (Kocabiyik *et al.*, 2012). Consistently, small heat shock protein, MjHSP16.5 of *Methanococcus jannaschii* kept the single chain monellin (SCM) soluble at 80 °C for 20 min at a monomer molar ratio of 1:1 (SCM: MjHSP16.5). This substrate was heat stable at 60 °C, but it precipitated at 80 °C in the absence of the MjHSP16.5. Furthermore, the MjHSP16.5 protected the CS, which is a mesophilic enzyme, at 40 °C at a monomer molar ratio of 1:40 of CS: MjHSP16.9 ( Kim *et al.*, 1998). These results indicate that high amount of the chaperone will be required at elevated temperatures for full protection of the activity of the client protein.

Protection of ADH from heat induced aggregation by WT *Tpv* sHSP14.3 was not as effective as the protection CS. Only 20 % protection was found at low ADH/Chaperone w/w ratio (*i.e.*, 1/2).

When the concentration of the chaperone was increased, it further enhanced aggregation of the ADH up to 120 %. Our results are well agree with the data of human sHSP Hsp20 which promoted the ADH aggregation. This was attributed to the formation of the insoluble complexes between Hsp20 and denaturated ADH (Bukach *et al.*, 2004). This situation was more pronounced in our case at high chaperone concentrations. On the other hand,  $\alpha\beta$  Crystallin of human suppressed ADH aggregation in a concentration dependent way and ADH aggregation was prevented 50% at 1:17 w/w chaperone/ADH ratio at 37°C (P. Santhoshkumar & Sharma, 2006), while heat induced aggregation of ADH in the presence of HSP22 of human was completely prevented at high chaperone concentration (at 2:1 w/w ratio



of HSP22/ADH)(M. V Kim *et al.*, 2006). These results indicate that chaperone efficiencies vary according to the substrate proteins and type of the chaperones.

Our activity assays showed that *Tpv* sHSP14.3 WT had ability to protect CS (up to 100%) and ADH (up to 64 %) from thermal inactivation considerably and to keep them in a partially folded state at high temperature. On the other hand, small heat shock proteins from *Mycobacterium tuberculosis* Hsp16.3, *Pisum sativum* HSP18.1 and HSP17.7, and murine small heat shock protein HSP25 could not protect CS activity at high temperatures (Lee *et al.*, 1995; Chang *et al.*, 1996; Ehrnsperger *et al.*, 1997) possibly by formation of very stable complexes of the enzyme and the chaperone. On the other hand, Hsp42 from *Saccharomyces cerevisiae* protected CS from thermal inactivation by binding to it transiently (Haslbeck, Braun, *et al.*, 2004). In the light of this information, although the mechanism of substrate recognition by sHSPs is not well understood, it is possible that WT *Tpv* sHSP14.3 prevented the CS and ADH thermal inactivation by binding to their active site which might be released upon onset of normal conditions (*i.e.*, temperature).

The analysis of oligomeric distribution by BN-PAGE showed that WT *Tpv* sHSP 14.3 protein formed both higher oligomers (>25-mer) and small species (<24-mer). Among other forms, the band of 24-mer species is particularly clear. Oligomeric form of the same size is the typical for archaeal sHSPs from *Methanococcus jannaschii*, *Sulfolobolus tokodaii*, and *Sulfolobolus solfataricus* (K. K. Kim *et al.*, 1998; Hanazono *et al.*, 2012; Liu *et al.*, 2015). Besides, 12 mer species were mostly formed by plant heat shock protein, Hsp16.9 form wheat and Hsp21 of the *Arabidopsis thaliana* (Rutsdottir *et al.*, 2017; R. L. van Montfort *et al.*, 2001). It is generally accepted that large oligomers of sHSPs are inactive, storage form of the sHSPs. When they are exposed to the stress (*e.g.*, high temperature), they are dissociated into smaller species, mainly dimers which are the active form of the sHSPs (Mogk *et al.*, 2019). In this respect, smaller species present in the *Tpv* sHSP14.3 can be active form that interact with the unfolded form of the model substrates (CS and ADH) under non-permissive conditions.

Furthermore, Rosetta energy diagram of the WT *Tpv* sHSP14.3 showed that the ACD of the protein have higher tendency to form  $\beta$ -fibrils while the unstructured flanking regions NTD and CTD have less propensity to form a fibril. Similarly, the ten human sHSPs (*i.e.*, HspB1-HspB10) were found to form amyloid fibrils. For this propensity, the most dominant region was found as their ACD while flanking regions could be involved in the prevention of the central region of the protein from aggregation. This high tendency of the ACD to form fibrils was linked to the its immunoglobulin like structure, which was rich in the anti-parallel beta strands (Carver *et al.*, 2017).

Analysis of the 3-D model structure of the *Tpv* sHSP14.3 predicted salt bridges and hydrogen cores in this protein related to dynamic nature of the oligomers (dissociation/association), which is proposed to be critical for chaperone activity (Wen *et al.*, 2010). Furthermore, surfaces of the *Tpv* sHSP14.3 and other small heat shock proteins contain several hydrophobic patches that recognize and interact with partially denatured proteins through hydrophobic interactions (Nandi *et al.*, 2015). In these respect, multi network (-inter and -intra molecular interactions) of the WT *Tpv* sHSP14.3 and its ACD mutants and their surface properties were comparatively discussed considering their structural and functional importance in the following sessions.

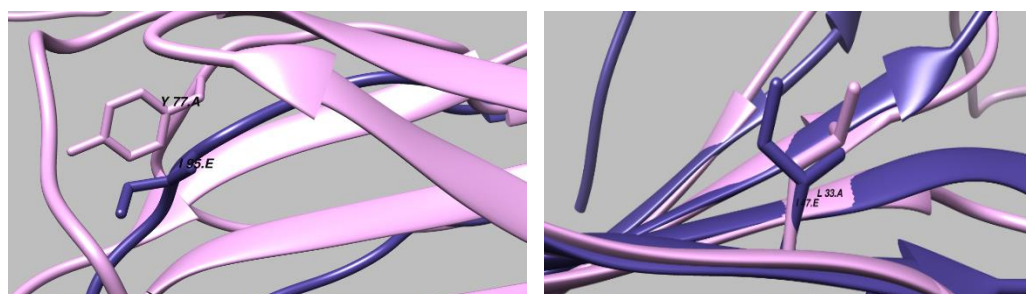
#### **4.1 Effect of the *Tpv* sHSP14.3 Mutants with Altered Hydrophobicity on Structure and Function**

##### **Group 1 (L33S, Y34F, G107A Mutants)**

The proposed and most widely accepted mechanism of chaperone action for sHSPs is that they recognize and interact with their substrate through exposed hydrophobic surfaces. Therefore, it is proposed that the hydrophobic residues present in the NTD and ACD of the sHSPs can be responsible for this function (Haslbeck *et al.*, 2019; Mogk *et al.*, 2019). Studies have supported the need of hydrophobicity in the proximal part of NTD in various sHSPs (Stromer *et al.*, 2004; Haslbeck *et al.*, 2004). Hydrophobicity at position 68 (M68) of CRYAB of mouse has been shown to play

an important role in its chaperone function. When hydrophobicity was increased by substituting methionine with highly hydrophobic isoleucine or valine, its chaperone activity also increased and decreasing the hydrophobicity by substituting with polar residue, threonine, decreased the chaperone function of the respective sHSP (N. P. Shroff *et al.*, 2001). Similar results were found in this study when the hydrophobicity at the corresponding residue L33 in *Tpv* sHSP 14.3 was reduced drastically by replacing it with serine. This mutation resulted in 1.21% decrease in the hydrophobicity of the ACD. As a result, there has been decrease (5% less than WT) in its efficiency to protect substrate (CS) against heat induced inactivation at lower concentrations (1:147 substrate: chaperone w/w ratio) (Figure 3.82). However, its protection efficiency also increased as the concentration of the mutant sHSP increased (*i.e.*, 1:250 and 1:500 substrate: chaperone w/w ratio) (Figure 3.62 & 3.74). Our research group previously showed that excess amount of *Tpv* sHSP14.3 were required for protection of mesophilic enzymes against thermal inactivation at temperatures above the optimum temperature for their activity (Kocabıyık *et al.*, 2012). The heat induced aggregation of CS was well protected by L33S mutant chaperone at low concentration (1:2 substrate: chaperone w/w ratio) (Figure 3.50). However, its slightly enhanced fibrillation propensity regarding one of the segments in L33S (*e.g.*, SYQDSS) as found by Rosetta energy analysis could make the protein more aggregation prone at 1/10 CS/chaperone w/w ratio due to molecular crowding. On the other hand, the chaperone activity of this mutant sHSP in protection against heat induced inactivation was observed to be similar to that of WT when ADH was used as the substrate (Figure 3.102). Substrate dependent change in the chaperone efficacy of the sHSP proteins, as we observed in this study, also reported for a mutant (M68T) of the CRYAB sHSP from mouse (N. P. Shroff *et al.*, 2001). In addition to this, the mutation L33S, slightly decreased the thermal stability of the protein which is similar to but less drastic than the decrease that in thermal stability that occurs in M68T mutation of CRYAB of *Mus musculus* (mouse) (N. P. Shroff *et al.*, 2001). The residue M68 was equivalent to L33 in *Tpv* sHSP14.3 as we found by MSA (Figure 3.13).

This decrease could be due to loss of hydrophobic interactions in the dimer as a result of mutation since hydrophobic residues are known to be involved in stabilization of the sHSP structure (dimer, since ACD has main role in dimer formation) by interacting with the counter residue of the adjacent monomer in a dimer (K. K. Kim *et al.*, 1998). Hydrophobic interactions are also important in formation of stable sHSP-substrate complex during chaperoning process as reported for putative substrate binding site (Trp48) of sHSP16.9 from wheat (R. L. M. Van Montfort *et al.*, 2001). The equivalent residue of the Trp48 in sHSP16.9 is Leu33 in *Tpv* sHSP 14.3 (Figure 3.14). Therefore, decrease in chaperone function that we observed in CS activity protection assay could be caused by the loss of the hydrophobic interactions formed by virtue of leucine. 3-D model structure analysis of WT showed that L33 displays intramolecular hydrophobic interactions between Leu33-Leu42, as well as the intermolecular interactions between Leu33-Tyr77 and Leu33-Ile78. Similar interactions were previously reported in crystal structure of *Methanococcus jannaschii* HSP16.5 where the equivalent of L33, Ile47 in the subunit contacts forms the hydrophobic bonds with Ile95 (Kim *et al.*, 1998). Ile95 of *M.jannaschii* corresponds to Tyr77 in the *Tpv* sHSP14.3 (Figure 4.1).



**Figure 4.1. Superimposition of the *Methanococcus jannaschii* HSP16.5 (blue, PDB:1shs) and *Thermoplasma volcanium* sHSP14.3 (Pink).** The residues of L33, and Tyr77 were found to equivalent to the I47 and I95 of the Mja HSP16.5, respectively.

In addition to the loss of hydrophobic interaction, introduction of serine decreased the surface hydrophobicity of the dimer. Although the area of interface around the mutation increased (evident from Zipper DB results), the hydrophobicity level of the exposed surface may not be sufficient to recognize and interact with the partially folded substrate. This decrease in hydrophobicity was found to favour formation of

medium and short-range contact while long-range contacts were quite less in the mutant.

Similar to the previous reports, decreasing the hydrophobicity at position 33 of *Tpv* sHSP14.3 in this study decreased its chaperone activity (Shroff *et al.*, 2001). Our next attempt was to analyse the effects of increasing the hydrophobicity at the adjacent position, 34 by replacing tyrosine with phenylalanine (Y34F mutation). This drastic increase in hydrophobicity at position 34 resulted in increase in the hydrophobicity of the ACD by 1.22%. As revealed by functional assays, this mutant showed appreciable chaperone activity in reducing light scattering of CS that was comparable to that of WT. In contrast to L33S mutant, at high CS: sHSP w/w ratio (1:10), Y34F shows better chaperone activity (Figure 3.51). This could be due to the low fibrillation propensity of the Y34F mutant protein as revealed by Zipper DB analysis that might have decreased the aggregation of the sHSP in a crowded environment, thus assisting the mutant protein in exerting its effect in the presence of the 5-fold-higher amount of the chaperone. Similar to our findings, increased hydrophobicity by Tyr6 to Val6 exchange in the Yeast Prion Sup 35 fragment (NNVVNV) resulted in low aggregation tendency of the mutant protein was attributed to the loss of tyrosine  $\pi$ -stacking, which is the side chain important for interactions within the interface (Do *et al.*, 2013). Furthermore, Y34F mutant was found to have 7.7% higher heat protection activity against CS thermal inactivation than WT at 1:500 CS: sHSP w/w ratio. Similar to the CS aggregation results, heat protection efficiency of the Y34F mutant protein was concentration dependent (Figure 3.63). When ADH was used as a substrate in heat induced inactivation assays, it was observed that ADH protection from heat inactivation was better than WT (Figure 3.99). The data available in literature have shown that enhanced chaperone activity of sHSPs at elevated temperature depends on the hydrophobic interactions which is the driving force to interact with the aggregation prone substrate proteins (Das & Surewicz 1995; Kim *et al.*, 2003; Van Montfort *et al.*, 2001 ; Bova *et al.* 2002). Y34F mutation did not change the putative intramolecular hydrogen bonds and the intramolecular hydrophobic interactions. On the other hand, hydrophobic surface

analysis revealed that substitution of the hydrophobic amino acid at the residue 34 to generate Y34F mutant caused slight increment in the hydrophobicity of its surface as compared to the WT. This could contribute to the improvement in its chaperone activity (as compared to WT), at higher concentrations. Alteration of the surface properties was also reported for small heat shock protein Hsp27 from *Drosophila melanogaster* by a mutation in the its ACD (DmHspR131G) which was accompanied by increased the chaperone activity as compared to the WT (Moutaoufik *et al.*, 2017). Also, in human small heat shock protein CRYAB, when H83 residue was replaced with highly hydrophobic alanine, a 30% increase in bis-ANS binding and 15%-20% increase in chaperone activity of the mutant sHSP was found as compared to the wild type (P. Santhoshkumar & Sharma, 2006). Our MSA and detailed 3-D structure analysis showed that the Y34 in *Tpv* sHSP14.3 is corresponding to the residues involved in subunit contacts in other sHSPs, *e.g.*, residue I48 of *M. jannaschii* HSP16.5, and K49 of *Xanthomonas axonopodis* XaHspA (Hilario *et al.*, 2011; Kim *et al.*, 1998). Therefore, it is likely that, point mutation at this position might have affected the dimerization and in turn the chaperone activity of the *Tpv* sHSP14.3.

Differences in the oligomeric status of the Y34F mutant and WT *Tpv* sHSP14.3 was shown by BN-PAGE analysis. This mutant formed 36-mer species instead of 24 mer species that was seen in the WT as also documented for engineered variants of the Hsp16.5 from *Methanococcus jannaschii*, Hsp20.2 of *Deinococcus radiodurans*, and XaHspA of *Xanthomonas axonopodis* that all formed 36-mer oligomers. This form was suggested to be storage form and dissociated into smaller assemblies when interacting with substrates (Bepperling *et al.*, 2012; Hilario *et al.*, 2011; Mishra *et al.*, 2018). Besides, both WT and Y34F mutant had octamer and dimer. In these respects, 36 mer in Y34F can represent the inactive form of the oligomer for chaperone function, but its activation may occur by structural rearrangements towards the smaller species.

On the other hand, the enhanced chaperone activity of Y34F mutant and its thermal stability can also be attributed to the increase in total number of contacts as compared to the WT. It is known that high number of inter-residue contacts between residues

in thermophilic proteins can increase their proteins H-bonding ability, which is important for thermal stability (Gromiha & Selvaraj, 2004). Our contact map analysis showed that most of the contacts at the position 34 were found between hydrophobic residue pairs (*e.g.*, Leu, Val, Ile, Ala). It is known that hydrophobic amino acids in the core of thermophilic protein are crucial for adapting to higher temperatures (Sælensminde *et al.*, 2009). Therefore, those contacts in the WT *Tpv* sHSP14.3 and Y34F mutant could be also important for thermal stability of the protein. As reported for other proteins, the distance between the alpha carbon atom of residues in a polypeptide chain can give information about the folding rate of proteins, protein stability upon mutations and thermal stability of proteins (Ponnuswamy & Michael Gromiha, 1994; Gromiha & Selvaraj, 2004). In our study, Chimera's distance map analysis showed that the overall distance in the monomeric structure of Y34F was reduced when compared the WT. This result may indicate that as a result of Y34F mutation, the *Tpv* sHSP14.3 protein became more compact, especially at the proximal of unstructured NTD region.

However, the prediction of mechanism of action of sHSPs is not so straight forward and does not follow a single rule means increase in hydrophobicity does not always result in an increase in chaperone activity. For example, in HspH from *Bradyrhizobium japonicum*, increasing hydrophobicity at a hydrophobic and highly conserved glycine at position 114 by replacing it with alanine suppressed its chaperone activity (Lentze *et al.*, 2003b). The equivalent residue in *Tpv* sHSP 14.3 is G107 (Figure 3.12), and replacing it with alanine resulted in similar results, *i.e.*, less potent chaperone activity. This mutation decreased protection ability of the mutant protein against thermal aggregation and inactivation of the CS. Although this mutation also increased the hydrophobicity of the ACD to a similar extent like Y34F (1.22%), the resultant mutant was less efficient as compared to the WT. This may be due to the fact that G107 residue is not surface exposed (in a dimer), therefore, any increase in the hydrophobicity might not cause any improvement in its recognition and interaction with substrate. Considering the oligomeric distribution, the mutant G107A behaved similar to WT indicating that this mutation did not affect the

oligomeric assembly. This decrease in chaperone activity can be explained by the results of the Zipper DB analysis. The increase in hydrophobicity was accompanied by an increase in the Rosetta energy and C-score of the mutation in the surrounding region, suggesting a disturbance in the equilibrium of fibril formation propensity by shifting it towards increased fibrillation capacity. Structural analysis demonstrated that the G107A point mutation caused formation of intramolecular hydrophobic interactions but disappearance of the intramolecular hydrogen bonds presents in the wild type sHSP. Since the G107 is located at a highly conserved G-x-L motif, change in its bonding property after point mutation may cause important structural perturbations that can be critical for its chaperone activity. Also, G107A mutation increased the putative fibrillation property. Due to this effect, this mutant protein can probably co-aggregate with CS in the thermal aggregation measurements when used 5-fold more chaperone (Figure 3.52).

Similar results were reported in another study where the corresponding residue in human CRYAB, G141, when substituted by a charged residue (arginine) also displayed oligomerization behaviour similar to WT, but it failed to exhibit efficient chaperone activity (Muchowski *et al.*, 1999) for both CS and ADH substrates. Similarly, the equivalent Glycine mutation G114A in *Bradyrhizobium japonicum* HspH showed decrease in chaperone activity and this mutant did not retain the CS in the soluble fractions, indicating a perturbation in substrate binding (Lentze *et al.*, 2003a). Besides, changes (decrease) in the hydrophobic contacts between the alpha carbon atoms within the NTD as a consequence of the mutation could be the reason of decreased chaperone activity of this mutant in preventing the CS thermal aggregation and the activity since the proposed mechanism of the sHSP action involvement of the NTD in substrate recognition and protection.

Overall, it can be stated that dimer interface residues are important for chaperone activity of the *Tpv* sHSP 14.3, regardless of whether and how much role they play in maintaining the structure integrity.



## 4.2 Effect of the *Tpv* sHSP14.3 Mutations that Remove Charge but Introduce Hydrophobicity

### Group 2 (E43V and E45G Mutants)

In this category, we checked the effect of substituting charged residues by hydrophobic amino acids; E43V and E45G mutations.

The resolved crystal structures of the ACD dimer of archaeal small heat shock protein Hsp20.1 of *Sulfolobus solfataricus*, Hsp14.0 of *Sulfolobus tokodaii* and Hsp16.5 of *Methanococcus jannaschii*, as well as, CRYAB of human showed that salt bridges, inter-molecular ionic interactions and hydrogen bonds in the dimer interface can be critical for subunit-subunit interactions for formation of dimeric and the oligomeric species, which in turn can affect function of the small heat shock proteins (Clark *et al.*, 2011; L. Liu, Chen, Yang, & Wang, 2015; Quinlan *et al.*, 2013). In our study, structure analysis showed that bonding pattern of the E43V variant changed after the mutation such that intermolecular H-bonding and salt bridges between Arg81-Glu43 were lost. These bonds could be important for stabilizing the dimerization loop  $\beta$ 6 by interaction of Arg81 on  $\beta$ 6 strand with the Glu43 on the partner monomer. It is known that intersubunit salt bridges play role in dimerization, stabilization of the  $\beta$ -sheet, and in turn affect chaperone function (Mani *et al.*, 2016). Structure comparison showed differences mainly at the  $\beta$ 6 strand of the ACD. It is highly possible that these changes can affect the dimerization of the *Tpv* sHSP14.3 and consequently its function in protecting model substrates at higher temperatures.

At lower concentrations of the E43V mutant protein, the protection of citrate synthase against heat induced aggregation and restoring its activity was lower than WT could be due to the loss of electrostatic interactions. It has been previously reported for human  $\alpha\beta$  crystallin and SsHsp 14.1 from *Sulfolobus solfataricus* that loss of salt bridges by mutations of glutamic acid in ACD impaired chaperone activity of the sHSPs (Clark *et al.*, 2011; Wen *et al.*, 2010). However, the chaperone activity of the E43V mutant was concentration dependent, such that when the

concentration of the mutant protein increased, its chaperone activity also increased becoming similar to (CS activity assay) or even more than (CS aggregation assay) WT.

Our results also indicated that E43V mutant exerts its chaperone activity in a substrate specific manner. Although it could protect CS from heat induced aggregation (>80%), it did not protect the ADH from thermal aggregation as same efficiently as CS, but still better protection as compared to the WT was found at high concentrations. ADH activity assay showed that the extend of the thermal inactivation of the ADH was decreased in the presence of E43V mutant and this effect was similar to the wild type. There are reports in literature about difference in chaperone efficiency of sHSPs according to the substrates (Basha, 2012).

Our MSA and 3-D structure analysis revealed that residue 43 could be important for subunit-subunit interactions. The residue Y58 of XaHspA from *Xanthomonas axonopodis* which has important role in dimerization was corresponded to E43 in *Tpv* sHSP14.3 (Hilario *et al.*, 2011). The residue N78 of the human CRYAB, which have role in the subunit-subunit interactions, also is topological equivalent to E43 in *Tpv* sHSP14.3 (Figure 3.16). The N78D mutation, resulted in formation of mostly larger oligomeric assemblies (Ghosh & Clark, 2005). Therefore, it is expected that E43V mutation can affect the structural dynamics (dimerization/oligomerization) of the *Tpv* sHSP14.3 protein.

Besides, E43V mutation resulted in the compactness of the monomer molecule as revealed by the RRdist map analysis. Correlated with this result, contact map analysis also a showed that total number of contacts (mostly long-range contacts) reduced within the monomeric structure of the *Tpv* sHSP14.3. Thermodynamic stability of this mutant was found to be stabilizing with a high confidence score, as revealed by the MUpro analysis. Considering the functional assays using the model substrates CS and ADH, further improvement compared to the WT in protection against the thermal inactivation of these enzymes was not found.

Although surface analysis showed that introduction of the highly hydrophobic amino acid valine enlarged the hydrophobic surface of the protein, its chaperone activity

was not enhanced. According to the functional assay results of E43V, charged residues in the dimer interface could be more important in substrate binding than the hydrophobic ones.

Another glutamic acid mutation, at position 45 on  $\beta$ 3 strand was produced by replacing it with glycine (E45G). This negatively charged residue was highly conserved in eukaryotic, eubacteria and archaeal sHSPs. When it was replaced by hydrophobic G on  $\beta$ 3 strand, its molecular chaperone function was adversely affected. This mutant sHSP was found to be less effective in prevention of the citrate synthase aggregation than the WT. Besides its low chaperone function, thermal instability was also noticed for this mutant. Although there was appreciable amount of E45G protein in the cell free extract before heat treatment, it has been completely lost at higher temperature (*i.e.*, 60°C) because of low thermal stability. Since electrostatic interactions, ion-pairs or hydrogen bonds are crucial for maintenance of protein stability and function (Nandi *et al.*, 2015), the disappearance of intermolecular salt bridges (*e.g.*, Glu45-Arg69), intermolecular hydrogen bonds (*e.g.*, Arg69-Glu45 and Gly48-Glu45), attractive charge interactions (*e.g.*, Arg81-Glu45) and pi-anion bond between the charged and aromatic amino acids (Glu45-Phe49) could be reason of the reduced thermal stability of the mutant protein. The interactions Arg69-Glu45, and Arg81-Glu45 could be also crucial for the stabilization of dimer interface because Arg69 and Arg81 were positioned on the dimerization loop,  $\beta$ 6. It is possible that losing the negative charge at the position of 45 can destabilize the dimeric structure and in turn oligomeric structure.

In another study, compensating effect of the negative charge, D109 against two positively charged residues (R107-R120) and one negatively charged residue (D80) in the dimer interface of human CRYAB was eliminated after the D109A mutation. Because of the loss of associated hydrogen bonds and ionic interactions not only the monomer of the mutant but also the whole oligomer became less stable and aggregation prone. This D109A mutation was associated with the myofibrillar myopathy in human due to inability to prevent aggregation of the proteins. The resultant mutant formed large aggregated fibrils (Fichna *et al.*, 2017). Similar to

these results, the E45G mutant protein of the *Tpv* sHSP14.3 also showed an enhanced tendency to form fibrils, as revealed by Rosetta energy levels. This could be arise from the loss of a cluster of ionic interactions present in the WT *Tpv* sHSP14.3 and could account for the decreased heat stability and chaperone activity of the mutant protein. Our MSA analysis and 3-D structure superimposition of the *Tpv* sHSP14.3 with human CRYAB showed that E45 is equivalent to the D80 in CRYAB, which was suggested to be crucial in stabilization of the dimer. Therefore, it is quite possible that E45 could also be critical residue for proper assembly of the oligomers. In another study, D140N cataract causing mutation in human CRYAB, consistent with our results, led to a decline in thermal stability and chaperone activity of this sHSP. Their results showed D140 forms a salt bridge with H83 inside the  $\beta$ 3 strands that is essential for stabilization of the intra and intersubunit contacts in the structure of  $\alpha\beta$ -crystallin (Datskevich *et al.*, 2012; Y. Liu *et al.*, 2006).

This mutation also led to loss of long-range contacts and decline in RR overall distance, which are accompanied by the decrease in the thermodynamic stability of the mutant protein.

Overall, our structural analyses and functional assay results, we suggested that E43 and E45 are critical residues for intermolecular ionic interactions between the monomers of sHSP14.3 dimer. Their replacement by hydrophobic amino acids may result in pronounced disruption of monomer-monomer association that would adversely affect the stability of the *Tpv* sHSP 14.3 dimer/oligomer assembly and chaperone activity.

### **4.3 Effect of the *Tpv* sHSP14.3 Mutants with Gain of Charge (loss of hydrophobicity) on Structure and Function**

#### **Group 3 (A47D, G48E, G107D, and I108K Mutants)**

In this study, we have constructed four *Tpv* sHSP14.3 single mutants by replacing the hydrophobic amino acid within the ACD with the charged ones: A47D, G48E, G107D and I108K.

The residue A47 was found in the loop connecting the two anti-parallel  $\beta$ 3- $\beta$ 4 strands and a part of the P/A-G motif which is highly conserved in non-metazoan sHSPs. It was predicted that P/A-G doublet has direct role in the subunit-subunit interactions (Fu & Chang, 2006; Poulain *et al.*, 2010). The A47 residue was surrounded with the highly hydrophobic amino acids (*i.e.*, M-A-G-F) within the primary sequence of the *Tpv* sHSP14.3. The fine resolved crystal structure analysis of *Xanthomonas axonopodis* small heat shock protein XaHspA and human CRYAB implied that P62 in XaHspA and K82 in CRYAB that are equivalent to the A47 in *Tpv* sHSP14.3 could be key residues for the dimerization and intermolecular interactions (Hilario *et al.*, 2011; Ghosh & Clark 2005). Only one mutagenesis study revealed the importance of this motif in dimerization of Hsp16.3 of *M.tuberculosis* by substituting the glycine 59 residue (G59W and G59A). In this study, for the first time the conserved alanine residue of the P/A-G doublet was targeted for point mutation.

Structural analysis of A47D mutant showed that the intermolecular hydrophobic interactions between Ala47-Ala47 were lost upon mutation, which was thought to be responsible for the stability of interactions in the dimer interface. On the other hand, this mutation contributed to the formation of new intramolecular and intermolecular H-bond. These alterations in the side chain interactions within the dimer can be account for the structural differences noticed at the NTD, CTD and  $\beta$ 2/ $\beta$ 3 loop of the mutant sHSP. Not only side chain interactions, but also alpha carbon atom interactions were changed after the mutation. The resultant mutant sHSP lost long range hydrophobic contacts between the alpha carbon atom of Ala47 that could be a reason of the very low heat stability of the mutant protein at 60 °C as compared to the WT.

Reduced hydrophobicity together with the increased compactness of the protein structure as a result of decreasing distance between the residues can be additional factors responsible for the decline in the thermodynamic stability.

Above mentioned, changes in the structure can be associated with the decrease in chaperone activity of A47D mutant sHSP as compared to that of WT. Taking all these into consideration, mutation in the highly conserved P/A-G motif has negative

affect on structural stability and consequently, thermal stability and molecular chaperone function of the *Tpv* sHSP14.3.

The next residue G48, in the *Tpv* sHSP14.3 is located in the loop that connects two antiparallel  $\beta$ -strands, the  $\beta$ 3- $\beta$ 4 strands. Similar to A47, G48 is also located at the P/A-G motif. Equivalent Gly62 in *Methanococcus jannaschii* was suggested to be involved in intersubunit contacts as revealed by crystal structure analysis (Kim *et al.*, 1998). Site directed mutagenesis of *Mycobacterium tuberculosis* Hsp 16.3 at the equivalent position G59, which is also an equivalent residue for G48 of *Tpv* sHSP 14.3, also showed the importance of Pro-Gly doublet in subunit interaction (Fu *et al.*, 2006). Another study on human  $\alpha\beta$  crystallin was conducted at the corresponding conserved KH motif in metazoan sHSPs. The substitution of H83 by alanine in this motif confirmed importance of this residue in the anti-aggregation activity and substrate binding (P. Santhoshkumar & Sharma, 2006). H83 in human  $\alpha\beta$  crystallin is equivalent to G48 in *Tpv* sHSP14.3 as found by MSA analysis (Figure 3.13). Therefore, it was expected that our targeted mutation at position 48 could have impact on subunit interactions, dimerization and chaperone function of the *Tpv* sHSP14.3.

Our thermal aggregation measurements using CS and ADH as substrates showed that G48E mutant protein at lower chaperone concentrations displayed greatly reduced chaperone activity as compared to the WT (Figure 3.56 & Figure 3.90). However, the inactivation of the substrate enzymes, CS and ADH, was better protected by this mutant sHSP, as compared to the WT. This could be attributed to the fact of its tendency to coaggregate with ADH. A drastic increase in C-score of all the surrounding area of the mutation site as a result of the increase in interface can increase fibrillation propensity of this mutant sHSP. This might lead to its co-aggregation with aggregating substrate (the nucleus). However, the mutant protein itself did not aggregate when incubated alone at this temperature. Similarly, it was reported that human small heat shock protein Hsp20 enhanced ADH aggregation through formation of the insoluble complexes with it (Bukach *et al.*, 2004). On the other hand, G48E was very effective in preventing the thermal inactivation of the CS

and ADH, and proved to be even better than WT especially at higher weight ratios (CS: Chaperone w/w 1/500 and 1/250 ratios). Results from the ADH activity assay also indicated that extent of the protection from heat inactivation was equal to the that of positive control. The increased chaperone activity of G48E mutant could be attributed to the strengthening of the dimer structure by formation of new intermolecular electrostatic bonds between Arg81-Glu48. Previously, it has been shown that electrostatic interactions play an important role in chaperone functions of HSP18 from *Mycobacterium leprae* (Nandi *et al.*, 2015).

As reported in this research for G48E mutant, although sHSPs can be quite effective in protecting the client proteins from heat inactivation, they may not prevent their aggregation at high temperatures. Similar results were documented for the small heat shock proteins from *Mycobacterium tuberculosis* Hsp16.3, *Pisum sativum* HSP18.1 and HSP17.7 and murine small heat shock protein Hsp25 (Lee *et al.*, 1995; Chang *et al.*, 1996; Ehrnsperger *et al.*, 1997; Garrett J. Lee *et al.*, 1997).

Furthermore, BN-PAGE analysis showed that single amino acid substitution shifted the equilibrium of the oligomers towards formation of the smaller species ( $\leq 24$ -mer). The presence of smaller oligomeric specie, but disappearance of large oligomers (*e.g.*, 60-mer oligomers) may indicate the perturbation of oligomer assembly as a result of G48E point mutation. The proposed mechanism for the action of sHSPs depends on the binding of substrate protein to dimeric or other sub-oligomeric forms of the sHSPs (Basha, 2012). Therefore, it is possible that the enhanced chaperone activity of G48E mutant for the substrate proteins CS and ADH could be due to the increased population of the smaller species like dimer, octamer.

This mutation led to decrease in the RR distance of the monomer as well as total number of contacts from 831 to 820. Also, long-range contacts were less as compare to the WT that may be due to the lowest long-range contact value of Glu as reported before (Barlow and Thornton, 1983; Gromiha and Selvaraj, 2004). The loss of the long-range contacts in turn can decrease the thermodynamic stability of the G48E mutant protein.

Another point mutation, G107D, targeted the highly conserved G-x-L motif. Different from the G107A mutation we introduced that increased hydrophobicity, G107D mutation drastically changed the amino acid character from hydrophobic to the negatively charged one in the  $\beta 8/\beta 9$  loop. The mutations of equivalent residues G141 and G114 of human CRYAB and *Bradyrhizobium japonicum* HspH, respectively demonstrated the importance of the highly conserved motif for chaperone function and structure stability of the sHSP (Lentze *et al.*, 2003b; Muchowski *et al.*, 1999). In our study, as an effect of the G107D mutation, the mutant protein became completely inactivated at high temperatures. This residue was located in the putative dimer interface and it had topologically equivalent residues found in the dimer interfaces of the other sHSPs (*i.e.*, wheat Hsp16.9 and MjHSP16.5) (K. K. Kim *et al.*, 1998; R. L. van Montfort *et al.*, 2001). Therefore, such drastic change in that position can destabilize the structure, leading to its irreversible aggregation at high temperature. Therefore, the precipitated protein after heat treatment could be eliminated during centrifugation step so that SDS PAGE analysis showed the complete loss of the protein after heat treatment. For this reason, functional analyses could not be performed with this mutant protein. Secondly, the mutation caused formation of new set of electrostatic and hydrogen bond interactions between Asp107 and the residues found on  $\beta 3$ - $\beta 4$  loop, on  $\beta 6$  loop and on  $\beta 8$  strand. Also, such alterations may impede the dimer stability and cause structural disturbances. Thirdly, contact map analysis showed that the mutation decreased the total number within the monomeric structure of the G107D mutant sHSP. Elimination of polar-hydrophobic and hydrophobic-hydrophobic contacts at the targeted mutation point could be another reason of the inactivation of this mutant sHSP at 60°C. Shortening of the  $\beta 8/\beta 9$  loop after the mutation can be due to loss of long-range contacts after the mutation. Loop size can affect the thermodynamic and kinetic properties of the proteins (Gavrilov *et al.*, 2015; Bigman & Levy, 2020). Therefore, decreased thermodynamic stability of the mutant protein can be associated with the decrease in the loop size and overall structural alterations.



Although Gly107 in the WT was found a buried residue in the WT structure, the substituting aspartic acid residue was exposed to the surface as an another result of the structural changes. Overall, the G107D mutation can cause drastic structural changes including perturbation of the subunit-subunit interactions (*i.e.*, dimerization) that leads to loss of structural integrity. Therefore, this mutant could not be further characterized.

The other mutation regarding the substitution of the charged residue in the place of a hydrophobic amino acid was generated by replacing isoleucine with lysine at residue 108. This residue in the *Tpv* sHSP 14.3 was located in the highly conserved and functionally important G-x-L motif. I108 corresponded to amino acid residue V128 that was suggested as important for intersubunit contacts and dimerization of *M. jannaschii* and *X. axonopodis* (Hilario *et al.*, 2011; Kim *et al.*, 1998). Also, pin array technology conducted in human CRYAB demonstrated that equivalent V142 was capable of making intermolecular interactions and related core sequence <sub>141</sub>GVLTVNGP<sub>148</sub> could contribute to the formation of a dimer interface (Ghosh & Clark, 2005). Therefore, substitution mutation was done at this position, such that, not only the hydrophobicity was reduced, but also a charge was introduced at this position. It was hypothesized that introduction of the charged residue at a highly conserved hydrophobic position within the conserved alpha crystallin domain of *Tpv* sHSP 14.3 could induce important structural and/or functional alterations. Thermal aggregation measurements of CS in the presence of I108K mutant sHSP was reduced significantly and 1:10 CS: sHSP w/w ratio this protection effect exceeded the WT (Figure 3.57). This might suggest that five-fold more chaperone concentration was required to compensate the effect of the mutation and be effective as or better than the WT. On the other hand, this mutant was less potent than WT in preventing CS thermal inactivation. Decrease in the hydrophobicity by this mutation from 37.8 % to 36.59 % in the ACD could be reason of the reduced chaperone activity for the CS. On the other hand, structural perturbations caused by this mutation can be responsible for the alteration in its chaperone function for the model substrate, CS. Moreover, our results clearly indicated that I108K mutant differs in chaperone

capacities in protecting different model substrates. Although I108K mutant exhibited partial chaperone activity for CS, its chaperone efficiency for protecting the ADH from thermal inactivation was higher than that of positive control and the other *Tpv* sHSP14.3 variants including WT. The observed improvement in the chaperone activity of I108K single mutant can be due to stabilization of the dimer structure by formation of new electrostatic bonds (Lys108-Glu43) and hydrogen bonds (Lys108-Ala44). Also, it was reported that electrostatic interactions are known to have crucial roles in protein stability and function (Nakamura, 1996; Matthew, 1985; Nandi *et al.*, 2015).

However, failure of the I108K mutant protein to prevent ADH from thermal aggregation might be due to its increased aggregation propensity. Increase in area of interface and the C-score may shift the equilibrium towards higher propensity to form fibrils as revealed by Zipper DB analysis (Figure 3.144-b). It is known that high beta sheet content is stabilized by sets of hydrogen bonds, which decreases the exposed hydrophobic surfaces by polar or hydrophobic interactions (Do *et al.*, 2013). I108K mutation has resulted in formation of new intramolecular H-bonds between Lys108 and Ala44. Therefore, the flexibility of the I108K mutant dimer and also its exposed hydrophobic surfaces can be decreased. In addition to these, surface hydrophobicity analysis revealed a decrease in surface hydrophobicity of the mutant sHSP. Another reason that accounts for the decrease in chaperone activity of this mutant could be the altered oligomer equilibrium of this mutant. Although there are bands that correspond to dimer and octamer, for this mutant, 24-mer specie is absent. Similar to the Y34F mutant, 36 mer species instead of typical archaeal 24 mer was observed in I108K mutant.

In addition to this, the band densities indicate that most of the sHSP was present in the form of larger order oligomers and the protein, therefore, failed to dissociate efficiently, even upon heat stress. Both WT *Tpv* sHSP14.3 and I108K mutant variant show intermediate size oligomers of 12 and 8 subunits similar to the wheat Hsp 16.9 which had dodecameric structure as revealed by crystal structure analysis (R. L. van Montfort *et al.*, 2001). Although *Tpv* sHSP 14.3 and Hsp 16.9 have 23% identity in

their amino acid sequence, 12 mer oligomeric structure was common between two different sHSP. This suggests that I108K mutation which is at the highly conserved region of the *Tpv* sHSP14.3, altered the oligomeric distribution profile as compared to the WT *Tpv* sHSP14.3.

Increase in intramolecular interactions between the alpha carbon atom of the amino acids lead to stable protein structures, where protein folding and protein thermal stability depends on the distance among the alpha carbon atom of each residues in a polypeptide chain (Ponnuswamy & Michael Gromiha, 1994; Gromiha & Selvaraj, 2004). Contact map analysis of I108K mutant showed that this mutation introduced additional contacts particularly with hydrophobic residues which were not available in WT sHSP monomeric structure. In these respects, high heat stability of the I108K mutant at the temperatures above 60°C could be attributed to the increased number of such contacts as reported before (Gromiha, 2001; Gromiha & Selvaraj, 2004).

#### **4.4 Effect of the *Tpv* sHSP14.3 Double Mutants on Structure and Function**

##### **Group 4 (Y34FG48E and G48EI108K Double Mutants)**

There are studies that suggest that two mutations in a single protein at adjacent or far apart residues might result in additive effect on the function of the protein (Kasakov *et al.*, 2007; Shashidharamurty *et al.*, 2005). It is also possible that the negative effect of a single mutation is compensated by the other mutation (Phadte *et al.*, 2019). Based on this fact, in this study, two double mutants, Y34FG48E and G48EI108K, were designed and characterized regarding their chaperone functions and structural properties. In Y34FG48E, hydrophobic and charged residues were introduced to the positions 34 and 48, respectively. In the other mutant G48EI108K negative and positive charges were introduced in place of hydrophobic residues at positions 48 and 108, respectively.

The ability of Y34FG48E to prevent thermal aggregation of CS was found to be better than WT and respective single mutants in the presence of low (1:2 CS: sHSP w/w ratio) and high amount of the mutant protein (1:10 CS: sHSP w/w ratio) as

compared to respective single mutants. In agreement with our results, double mutation in human CRYAA,  $\alpha$ A-R157Q/G98R, increased the chaperone activity against ADH and insulin aggregation than the single mutant  $\alpha$ A-G98R, which had the characteristics of an increased oligomeric mass, reduced chaperone function, and loss of structural stability. It was stated that gain of the charged residue in the pathogenic cataract causing mutation was neutralized by the compensatory mutation R157Q (Phadte *et al.*, 2019).

Also, double mutation Y34FG48E did not induce drastic alterations in the structure of the mutant protein as compared to single point mutations. In case of surface hydrophobicity, there is a gain and loss of hydrophobicity at position 34 and 48, respectively. Only the difference from the single point mutation was the formation of the new intramolecular electrostatic interactions between Arg 69-Glu48, which might contribute to its chaperone function by maintaining the structural stability. Besides, double mutations reduced the fibrillation propensity by reducing the Rosetta Energy of all the hexapeptides that include the two mutation points. This feature may have assisted the mutant protein to keep the CS in a soluble state at 45°C as compared to the WT. In contrast to CS aggregation assay, Y34FG48E mutant protein was found to be less potent than WT in suppressing the thermal inactivation of CS. Although it showed appreciable protection at higher weight ratios for retaining the activity of enzyme, it did not further improve the chaperone function as compared to the WT and respective single variants. It might be due to the stable binding of the mutant protein to the substrate CS. Consistently, Hsp25 from murine small heat shock protein did not prevent the CS thermal inactivation at 43°C owing to the formation of the long-lived complex between Hsp25 and CS intermediates (Ehrnsperger *et al.*, 1997).

On the other hand, the Y34FG48E double mutant in ADH activity assay strongly repressed thermal inactivation of the ADH and its effect was better than the respective single variants. On the other hand, like the associated single mutant variants, Y34FG48E double mutant also has not been efficient in prevention of heat induced aggregation of ADH. As it was discussed before for other *Tpv* sHSP 14.3

mutants (*e.g.*, Y34F, G48E), the mutant protein can be prone to the coaggregate with the ADH at the assay temperatures (*i.e.*, 43°C).

The contact map analysis of the Y34FG48E mutant showed that 50 of the residue contacts were different than those found in the WT protein. Specifically, hydrophobic contact pairs were lost in the double mutant. Since hydrophobic amino acids in the core of thermophilic protein enables adaptation to higher temperatures (Sælensminde *et al.*, 2009), loss of hydrophobic contacts could be reason of low heat stability of the Y34FG48E mutant protein at 80°C. As reported before, hydrophobic residues in the molecular chaperones (*i.e.*, Tyr, Val, Ile and Phe) make more long range contacts than that of medium range contacts (Kumarevel *et al.*, 1998). On the other hand, our contact map analysis showed that long range contacts (*e.g.*, Phe34-Val121, Ile27-Ala47, and Tyr28-Ala67) which are found in the single mutant Y34F disappeared by double mutation, Y34FG48E. It was stated that loss of long range interactions may increase the entropy of unfolded state which in turn lower overall protein stability (Bigman & Levy, 2020).

In the second double mutant (G48EI108K), the conserved hydrophobic amino acids which are quite far apart from each other were replaced by charged ones. Suppression of heat induced aggregation of CS by this double mutant was as effective as the WT *Tpv* sHSP14.3 and I108K single mutant at low (1:2 CS/chaperone w/w ratio) and high (1/10 CS/chaperone w/w ratio) chaperone concentrations. In a study, it was shown that Hsp42 from *Saccharomyces cerevisiae* prevented the CS aggregation up to a molar ratio of 1:1, and the addition of excess chaperone did not provide stronger suppression of CS (Haslbeck, Braun, *et al.*, 2004). Protection efficiency of this mutant variant was about 20% and 10% higher than the single mutant variant G48E at 1:2 and 1:10 CS/sHSP w/w ratio, respectively. The increased protection against the heat induced aggregation of CS could be due to the effects of the sets of multi network interactions. Double mutation resulted in formation of intramolecular salt bridges, ion pairs and increased the number of intramolecular hydrogen bonding which were not present in the WT. These interactions may play a role in contributing to dimer stability, and consequently enhancing its chaperone activity. Similar

interactions in the SsHsp 14.1 from *Sulfolobus solfataricus* were implicated in the dynamic nature of the oligomers (dissociation/ association), which is critical for its chaperoning process (Wen *et al.*, 2010).

Consistent with the CS aggregation assay, G48EI108K mutant protected the activity of CS with almost same effectiveness as WT, but its chaperone performance was lower than G48E single mutant variant. ADH activity assay also revealed that G48EI108K double mutant prevented the heat induced inactivation of ADH almost at the same rate as WT sHSP, but effect was less potent than the G48E, and I108K single variants.

Aggregation of the model substrate ADH at high temperature was reduced up to 3-fold (at 1/3 ADH/chaperone w/w ratio, Figure 3.94) in the presence of the double mutant, G48EI108K as compared to WT sHSP. Related two single mutants, G48E, and I108K, however enhanced heat induced aggregation of ADH at 43°C, even far beyond the WT sHSP.

Besides, G48EI108K mutant proteins have favourable Rosetta energy levels with low tendency that in turn prevent its aggregation. Therefore, double mutation had considerable effect on the aggregation and activity properties of the model substrate proteins, CS and ADH. Enhanced chaperone function of sHSP14.3 could be arisen from the alterations in the bonding patterns after the mutation.

In addition to these, the double mutation affected the oligomeric state of the G48EI108K mutant protein. The increasing number of smaller species (*e.g.*, monomer, dimer and octamer) after the mutation were more obvious. The results of CS and ADH chaperone assays indicated that G48EI108K mutant variant enhanced chaperone activity even if the surface hydrophobicity was reduced which could be due to a shift in oligomeric equilibrium towards the smaller species. Similar to the our results, K137, 141E double mutant of human Hsp22 had higher chaperone activity than the single mutant K137E. Enhanced chaperone activity of the double mutant in terms of suppression of aggregation of insulin and rhodanese was explained by its lower ability to form high molecular weight oligomers (Kasakov *et al.*, 2007).

Moreover, the double mutation changed the inter-residue contact pairs occurred in the ACD of the protein, and reduced the RRDist of the protein significantly. Relatively short loop size (*i.e.*,  $\beta 2/\beta 3$  loop and  $\beta 8/\beta 9$  loop in the ACD) could be a factor increasing compactness of the mutant protein and its thermodynamic stability. Comparison of contact map of the double mutant with corresponding single variants showed small alterations in forming contacts. Additive effect of second mutation at position 108 resulted in disappearance of the polar-charged and hydrophobic contacts, which were present in the G48E protein. Similarly, polar charged interaction seen in the single mutant I108K were lost after the double mutation. In thermophilic proteins, the strength of the contacts between hydrophobic residues and also between polar-charged residues are thought to be one of the factors contributing to the thermal stability of the protein (Goldstein, 2007). Therefore, relatively low heat stability of the double mutant as compared to the respective single mutants at 70°C could be due to the loss of the crucial contacts which have roles in the thermal stability of the protein.

When the single mutants Y34F, G48E, and I108K are compared to their double mutants Y34FG48E and G48EI108K, the chaperone function of the double mutant is more effective, in terms of protecting the aggregation of CS and ADH along with CS activity protection at higher concentrations.

#### **4.5 Effect of the *Tpv* sHSP14.3 Mutant with Reversing of the Charge on Structure and Function**

##### **Group 5 (K87E Mutant)**

The ACD mutation K87E was produced by exchange between the opposite charged residues. The resulting K87E mutant effectively prevented the irreversible aggregation of CS at low and high w/w ratio of CS:sHSP) and at 45°C. Its aggregation suppression effect was more than of WT and other *Tpv* sHSP14.3 variants. The highly conserved 87<sup>th</sup> residue is found in the  $\beta 7$  strand of the *Tpv* sHSP 14.3 protein. In a non metazoan sHSP PsHsp18.1,  $\beta 7$  strand has been proposed to be

a potential substrate binding site (Jaya *et al.*, 2009). Strong interactions were suggested to occur in or near the  $\beta 7$  strand (*i.e.*, K112 and L114). Analysis of the crystal structure of wheat sHSP TaHsp16.9 revealed that the N-terminal part of the protein was close to the  $\beta 7$  strand of the ACD. Therefore, they suggested that strong interactions in this region of ACD could occur with the N terminal part (R. L. van Montfort *et al.*, 2001). In these respects, reversing the charge by K87E mutation may favor the sHSP and CS interaction, which can assist the keeping the substrate in a soluble state at higher temperature. On the other hand, the K87E mutant not only failed to prevent thermal aggregation of ADH, but also promotes its aggregation at high temperature. As observed in this study for other mutant variants and reported before for other sHSPs, they could differ in their preferences for substrate specificity. For example, Hsp27 was found to be more effective for preventing aggregation of rhodanese and CS than MDH and GAPDH (Mymrikov *et al.*, 2017). On the other hand, the sHSPs are generally less efficient in suppressing aggregation of larger protein substrates (Haslbeck & Vierling, 2015). Therefore, it is probable that K87E can be less potent for retarding the aggregation of tetrameric ADH than the dimeric CS.

In addition, *Tpv* K87E equivalent mutation is associated with certain genetic disorders in human such that R120G in  $\alpha$ B-crystallin, R116C in  $\alpha$ A-crystallin (HSPB4) and R127W in Hsp 27 have been implicated in desmin-related myopathy, central nuclear cataracts and Charcot–Marie–Tooth disease, respectively (Clark *et al.*, 2011). Our CS activity results showed that pronounced effect of K87E mutation was on the chaperone function of *Tpv* sHSP14.3. This mutant sHSP was found to be less effective in suppressing the thermal inactivation of CS as compared to WT and other *Tpv* sHSP 14.3 variants. Similar to our results, R116D mutation in rat  $\alpha$ A-crystallin significantly reduced chaperone activity against the heat induced aggregation of the client protein. It was indicated that positive charge at that position is crucial for structural and functional integrity of  $\alpha$ A-crystallin (Bera *et al.*, 2002). In human Hsp22, mutation of the homologous residue K141E was correlated to the development of distal hereditary motor neuropathy. Structural and functional



analyses results showed this mutation changed the secondary structure of the K141E mutant protein by increasing the unstructured region (M. V Kim *et al.*, 2006). Chaperone activity of the K141E mutant against aggregation of alcohol dehydrogenase and rhodanese was found to be reduced that was attributed to destabilization of its structure. Similarly, in another study, corresponding R96E mutation in *Bradyrhizobium japonicum* HspH showed a strong tendency to precipitate and form misfolded proteins (Lentze *et al.*, 2003a).

The mutation, K87E, were found to have an effect on the structure of the protein. Reversing the charge at the position 87 resulted in the elimination of the hydrophobic interactions within the same subunits. Equivalent disease related R120G mutation in the human CRYAB resulted in the loss of the two positive charge along the interface and gain of two new interface ion pairs between His83 and Asp80. This interaction closed the groove by forming new salt bridges and in turn disturb subunit exchange dynamics and impair the chaperone activity (Clark *et al.*, 2011). Moreover, same R120G mutation was found to be responsible for desmin filament aggregation (Perng *et al.*, 2004).

Oligomeric status of K87E mutant as revealed by BN-PAGE analysis showed a shift toward the formation of larger oligomeric species. Similarly, equivalent R107G mutation in the *Methanococcus jannaschii* and R120G mutation in the human  $\alpha$ B crystallin resulted in oligomers that were larger and more polydisperse than the wild type (Mani *et al.*, 2016; Quinlan *et al.*, 2013). Therefore, single point mutation in the 87th position of *Tpv* sHSP 14.3 altered the oligomeric distribution of the K87E mutant sHSP. On the other hand, K87E mutation decreased the number of the long range contacts as well as overall RRdist of the protein that may account for the reduced thermodynamic stability.

#### **4.6 Chaperone Activity of *Tpv* ACD Derived Mini Chaperones**

In this study, the ability of the synthetic peptides derived from the ACD of *Tpv* sHSP14.3 to prevent heat induced aggregation of ADH and CS has been tested.

Among four peptides we designed, 11 aa LVD peptide did not protect ADH from thermal aggregation, but it showed significant chaperone activity towards suppression of CS aggregation. Therefore, functional sequences in the LVD mini-peptide can be important for recognizing the partially unfolded CS and protecting it from irreversible aggregation. The LVD peptide sequence resides in the  $\beta$ 3- $\beta$ 4 region of the ACD of *Tpv* sHSP 14.3. This region corresponds to the F<sub>71</sub>-S<sub>81</sub> of the human alpha  $\alpha$ -crystallin (P. Santhoshkumar & Sharma, 2006). The related human mini- $\alpha$ A70-83 peptide (KFVIFLDVKHFSPE) also could not abolish thermal aggregation of ADH (Raju *et al.*, 2016; K. Krishna Sharma *et al.*, 2000). Two  $\alpha$ A-crystallin mini-peptides with overlapping sequences *i.e.*,  $\alpha$ A66-80 and  $\alpha$ A67-75 contained the core sequence of "FVIFLD" which was thought to contribute fibril formation under physiological conditions due to its sequence similarity to  $\beta$ -amyloid (Raju *et al.*, 2016; Puttur Santhoshkumar *et al.*, 2011). The sequence of the LVD peptide although corresponds to this region of human  $\alpha$ A-crystallin, the sequence similarity is less. Therefore, it is likely that LVD peptide can form fibrils and it aggregate together with ADH during heating.

It has been suggested that as the length of fibril-forming peptide increases their ability to form fibrils decrease (Raju *et al.*, 2014, 2016). We designed the second mini-chaperone SDS by addition of extra amino acids to the N-and C- termini of the LVD peptide. The sequence of this peptide is also located at the same region ( $\beta$ 3- $\beta$ 4 of the ACD), which was proposed to be functional sequence for chaperone function of mini  $\alpha$ A crystallin (Kannan *et al.*, 2012; K. Krishna Sharma *et al.*, 2000). The 20 aa sequences of the SDS peptide corresponds to the human mini-  $\alpha$ B 73-92 peptide sequence. The human mini peptide with 100 % efficiency prevented the heat induced aggregation of ADH and this region was identified as substrate binding site. (Bhattacharyya *et al.*, 2006; Raju *et al.*, 2016). Consistent with these findings, the SDS mini-peptide of *Tpv* sHSP14.3 protected the two client proteins, ADH and CS from thermal aggregation. This may imply that  $\beta$ 3- $\beta$ 4 region of ACD plays active role in chaperone activity. Furthermore, Raju *et al.*, (2016) showed that specific amino acid sequences in the peptide can be important for its proper functioning as a

chaperone. For example, deletion of the four or six amino acids from mini- $\alpha$ B 73-92 peptide (*i.e.*, mini- $\alpha$ B 73-88; mini- $\alpha$ B 73-86) impaired its chaperone activity. Similarly, inability of LVD peptide for protecting ADH against heat induced aggregation could be due to the absence of essential amino acid residues which are available in the SDS mini-peptide. Our results also indicated that chaperone efficiencies of the LVD and SDS mini-peptides vary in a substrate dependent manner.

Thermal aggregation assay results showed that anti-aggregation property of 18-aa YIE mini-peptide is concentration dependent manner. This peptide design is corresponding to  $\beta$ 6- $\beta$ 7 region within the ACD of *Tpv* sHSP which is known to be critical for dimerization /dimer stability. As an equivalent sequence found in YIE peptide, the interactive sequence  $_{113}\text{FISREFHR}_{120}$  for chaperone function of the human CRYAB was identified by pin array technology. This interactive sequence was suggested to contribute to the structural integrity of the dimer by connecting two sheets of the core domain(Ghosh *et al.*, 2005).

On the other hand, 8-aa mini peptide GIK displayed effective chaperone activity for ADH, but not for CS. Sequence of this mini peptide is positioned at  $\beta$ 9 region of the ACD. The homologous sequence  $_{141}\text{GVLTVNGP}_{148}$  of CRYAB was found to be important for substrate binding and subunit interaction (Ghosh *et al.*, 2005). Since ACD is highly conserved in different sHSPs, interactive sequences identified in human can be potential chaperone active sites in other sHSPs, including *Tpv* sHSP14.3.



## CHAPTER 5

### CONCLUSION

1. The mutations investigated in this project includes highly conserved residues in alpha crystallin domain of the small heat shock protein *Tpv* sHSP 14.3 that are specifically located at the putative dimer interface. Results from chaperone activity assays showed that *Tpv* sHSP 14.3 ACD mutant variants and the wild type prevented the thermal inactivation of CS and ADH at high temperature. The extent of the protection against the heat induced inactivation varies among mutants and is specific to the substrate proteins.
2. Suppression of thermal aggregation of CS in the presence of WT *Tpv* sHSP14.3 and its ACD mutants was more effective at CS/Chaperone 1:2 w/w ratio, except A47D mutant sHSP, which showed better protection at higher CS/Chaperone w/w ratios.
3. Hydrophobicity at the junction of ACD and N-terminal domain was found to be important for chaperone function of the *Tpv* sHSP14.3. Decreased hydrophobicity L33S mutation was accompanied by a decline in chaperone activity against thermal aggregation of CS. On the other hand, introduction of the highly hydrophobic amino acid phenylalanine at position 34 in place of tyrosine provided further improvement in the chaperone function of the sHSP protein against CS and ADH thermal inactivation.
4. The charged residues at  $\beta$ 3 region of the ACD were important for chaperone function, where loss of the negative charge at position of 45 (E45G) and of 43(E43V) resulted in slight change in the chaperone activity based on thermal aggregation and activity assay results, especially at higher sHSP/substrate ratios. This may imply an insignificant role for these glutamic acid residues in substrate interaction.

5. A47D mutation in the highly conserved P/A-G motif displayed the least effective chaperone function as shown by CS aggregation assay. Together with reduced thermostability, this result indicates the importance of the conserved motif in structural stability of *Tpv* sHSP14.3 (*i.e.*, in dimerization and higher order oligomerization) as previously reported for other sHSPs.
6. Our results showed that substituting the highly conserved glycine with different characteristic amino acid (G107D) or similar characteristic amino acid (G107A) on the conserved G-x-L motif had deleterious effect on the activity and stability of the *Tpv* sHSP14.3. G107D mutant was completely inactivated at higher temperatures; while G107A mutant showed increased propensity for fibrillation. Overall, structural analysis indicated that this residue is critical for structural integrity.
7. Introduction of charged residues by G48E, I108K single mutations and Y34FG48E, G48EI108K double mutations resulted in enhancement of the chaperone activity of *Tpv* sHSP against heat induced inactivation of ADH. Formation new intramolecular electrostatic interactions in mutant chaperones are suspected to be one of the factors for increased structural integrity as revealed by model structure analysis.
8. Comparison of the double mutant G48EI108K with its respective single mutants showed that the chaperone function of the double mutant is more effective, in terms of protecting the aggregation of CS and ADH at higher concentrations. This result indicates a synergistic effect of the single mutations when they are combined in the same protein. In our case, G48EI108K mutant sHSP possessed the highest heat-protection effect among the other mutant variants against aggregation of ADH.
9. Oligomer size distribution analysis of the WT *Tpv* sHSP 14.3 and its ACD mutants showed that archaeal characteristic 24-mer is common among mutants. However, in Y34F and I108K mutant sHSPs instead of it, 36-mer oligomers are available. Different from other mutants, K87E mutant sHSP mostly formed 24-mer species and larger oligomeric forms. This shift toward

formation of larger oligomeric species was also detected in the equivalent disease-causing mutation (cataract) in the human. Such pronounced alterations in the oligomer distribution are indicating the importance of above-mentioned residues in oligomer assemblance.

10. The introduced mutations affected the surface properties of *Tpv* sHSP 14.3 ACD mutants. Depending on the properties of the substituted amino acid, surface become hydrophobic to hydrophilic or vice versa. Gly107 in the WT is found as buried residue, but after exchange 107D residue was exposed to the surface of protein as a result of the structural changes.
11. Rosetta energy diagram of the WT *Tpv* sHSP14.3 showed that the ACD has higher tendency to form  $\beta$ -fibrils while the unstructured flanking regions NTD and CTD have less propensity for fibrillation. Our study showed that substitution with the negatively charged residues (*i.e.*, A47D, G48E, G107D, Y34FG48E, G48EI108K) was more effective than the exchange by the hydrophobic amino acids in reducing fibrillation propensity of the *Tpv* sHSP 14.3 protein.
12. Our results showed that the number of the long-range contacts between the alpha carbon atom of the residues decreased as a result of the *Tpv* sHSP14.3 ACD mutations. Its common effect in the structure of the mutant variants that loop size is decreased and compactness is increased, as revealed by RRdist map analysis. Among the group of *Tpv* sHSP14.3 ACD mutants, where negative charge is introduced, the most affected mutant variant was G107D which has lowest number of the contacts. This was followed by A47D, K87E and G48E mutants, indicating the importance of the negatively charged residues on reducing the long-range contacts. Those structural changes can be responsible for the change in the orientation of the NTD and CTD as can be seen in 3-D model structures.
13. Thermodynamic stability programmes showed the decreased thermodynamic stability of the mutant proteins, except E43V mutation which slightly increased thermodynamic stability of the mutant sHSP. We also found a

correlation between the long-range contacts and thermodynamic stability such that loss of the long-range contacts led to decrease in thermodynamic stability of the protein.

14. The novel mini chaperones we designed regarding their superior chaperone activities as compared to WT *Tpv* sHSP14.3 can be considered promising for chaperone therapy. Our peptide designs are corresponding to sequences in  $\beta$ 3- $\beta$ 4,  $\beta$ 6- $\beta$ 7, and  $\beta$ 9 regions within the ACD of *Tpv* sHSP, which are critical regions for dimerization, dimer stability, and consequently for the chaperone function.



## REFERENCES

- Aquilina, J. A., Benesch, J. L. P., Lin, L. D., Yaron, O., Horwitz, J., & Robinson, C. V. (2005). Subunit exchange of polydisperse proteins: Mass spectrometry reveals consequences of  $\alpha$ A-crystallin truncation. *Journal of Biological Chemistry*, *280*, 14485–14491. <https://doi.org/10.1074/jbc.M500135200>
- Arrigo, A. P. (2013). Human small heat shock proteins: Protein interactomes of homo- and hetero-oligomeric complexes: An update. *FEBS Letters*, *587*(13), 1959–1969. <https://doi.org/10.1016/j.febslet.2013.05.011>
- Barlow, D. J., & Thornton, J. M. (1983). Ion-pairs in proteins. *Journal of Molecular Biology*, *168*(4), 867–885. [https://doi.org/10.1016/S0022-2836\(83\)80079-5](https://doi.org/10.1016/S0022-2836(83)80079-5)
- Basha, E. (2012). Small heat shock proteins and alpha-crystallins: dynamic proteins with flexible function. *Trends Biochem. Sci.*, *37*(3), 106–117. <https://doi.org/10.1016/j.tibs.2011.11.005>
- Basha, E., Friedrich, K. L., & Vierling, E. (2006). The N-terminal arm of small heat shock proteins is important for both chaperone activity and substrate specificity. *Journal of Biological Chemistry*, *281*(52), 39943–39952. <https://doi.org/10.1074/jbc.M607677200>
- Bepperling, a., Alte, F., Kriehuber, T., Braun, N., Weinkauff, S., Groll, M., Haslbeck, M., & Buchner, J. (2012). Alternative bacterial two-component small heat shock protein systems. *Proceedings of the National Academy of Sciences*, *109*(50), 20407–20412. <https://doi.org/10.1073/pnas.1209565109>
- Bera, S., Thampi, P., Cho, W. J., & Abraham, E. C. (2002b). A positive charge preservation at position 116 of  $\alpha$ A-crystallin is critical for its structural and functional integrity. *Biochemistry*, *41*, 12421–12426. <https://doi.org/10.1021/bi0204140>
- Bertoldo, C., Dock, C., & Antranikian, G. (2004). Thermoacidophilic microorganisms and their novel biocatalysts. *Engineering in Life Sciences*, *4*(6), 521–532. <https://doi.org/10.1002/elsc.200402155>
- Bhandari, S., Biswas, S., Chaudhary, A., Dutta, S., & Suguna, K. (2019). Dodecameric structure of a small heat shock protein from *Mycobacterium marinum* M. *Proteins: Structure, Function and Bioinformatics*, *87*(5), 365–379. <https://doi.org/10.1002/prot.25657>

- Bhattacharyya, J., Padmanabha Udupa, E. G., Wang, J., & Sharma, K. K. (2006). Mini-alphaB-crystallin: a functional element of alphaB-crystallin with chaperone-like activity. *Biochemistry*, *45*, 3069–3076. <https://doi.org/10.1021/bi0518141>
- Bigman, L. S., & Levy, Y. (2020). Entropic Contributions to Protein Stability. *Israel Journal of Chemistry*, *60*(7), 705–712. <https://doi.org/10.1002/ijch.202000032>
- Boelens, W. C. (2020). Structural aspects of the human small heat shock proteins related to their functional activities. *Cell Stress and Chaperones*, *25*(4), 581–591. <https://doi.org/10.1007/s12192-020-01093-1>
- Bova, M. P., Huang, Q., Ding, L., & Horwitz, J. (2002). Subunit exchange, conformational stability, and chaperone-like function of the small heat shock protein 16.5 from *Methanococcus jannaschii*. *Journal of Biological Chemistry*, *277*(41), 38468–38475. <https://doi.org/10.1074/jbc.M205594200>
- Bukach, O. V., Seit-Nebi, A. S., Marston, S. B., & Gusev, N. B. (2004). Some properties of human small heat shock protein Hsp20 (HspB6). *European Journal of Biochemistry*, *271*, 291–302. <https://doi.org/10.1046/j.1432-1033.2003.03928.x>
- Carver, J. A., Grosas, A. B., Ecroyd, H., & Quinlan, R. A. (2017). The functional roles of the unstructured N- and C-terminal regions in  $\alpha$ B-crystallin and other mammalian small heat-shock proteins. *Cell Stress and Chaperones*, *22*(4), 627–638. <https://doi.org/10.1007/s12192-017-0789-6>
- Chang, Z., Primm, T. P., Jakana, J., Lee, I. H., Serysheva, I., Chiu, W., Gilbert, H. F., & Quiocho, F. a. (1996). *Mycobacterium tuberculosis* 16-kDa antigen (Hsp16.3) functions as an oligomeric structure in vitro to suppress thermal aggregation. *Journal of Biological Chemistry*, *271*(12), 7218–7223. <https://doi.org/10.1074/jbc.271.12.7218>
- Chen, J. E., Huang, C. C., & Ferrin, T. E. (2015). RRDistMaps: A UCSF Chimera tool for viewing and comparing protein distance maps. *Bioinformatics*, *31*(9), 1484–1486. <https://doi.org/10.1093/bioinformatics/btu841>
- Cheng, J., Randall, A. Z., Sweredoski, M. J., & Baldi, P. (2005). SCRATCH: A protein structure and structural feature prediction server. *Nucleic Acids Research*, *33*(SUPPL. 2), 72–76. <https://doi.org/10.1093/nar/gki396>
- Cheng, Jianlin, Randall, A., & Baldi, P. (2006). Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins: Structure, Function and Genetics*, *62*(4), 1125–1132. <https://doi.org/10.1002/prot.20810>

- Chung, C. T., Niemela, S. L., & 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*, 2172–2175. <https://doi.org/10.1073/pnas.86.7.2172>
- Clark, a. R., Naylor, C. E., Bagn eris, C., Keep, N. H., & Slingsby, C. (2011). Crystal structure of R120G disease mutant of human  $\alpha$ B-crystallin domain dimer shows closure of a groove. *Journal of Molecular Biology*, *408*(1), 118–134. <https://doi.org/10.1016/j.jmb.2011.02.020>
- Collier, M. P., & Benesch, J. L. P. (2020). Small heat-shock proteins and their role in mechanical stress. *Cell Stress and Chaperones*, *25*(4), 601–613. <https://doi.org/10.1007/s12192-020-01095-z>
- Dabbaghizadeh, A., Finet, S., Morrow, G., Moutaoufik, M. T., & Tanguay, R. M. (2017). *Cell Stress and Chaperones*, *22*, 577–588. <https://doi.org/10.1007/s12192-017-0784-y>
- Das, K. P., & Surewicz, W. K. (1995). Temperature-induced exposure of hydrophobic surfaces and its effect on the chaperone activity of  $\alpha$ -crystallin. *FEBS Letters*, *369*, 321–325. [https://doi.org/10.1016/0014-5793\(95\)00775-5](https://doi.org/10.1016/0014-5793(95)00775-5)
- Datskevich, P. N., Nefedova, V. V., Sudnitsyna, M. V., & Gusev, N. B. (2012). Mutations of small heat shock proteins and human congenital diseases. *Biochemistry (Moscow)*, *77*(13), 1500–1514. <https://doi.org/10.1134/S0006297912130081>
- Do, T. D., Economou, N. J., Lapointe, N. E., Kincannon, W. M., Bleiholder, C., Feinstein, S. C., Teplow, D. B., Buratto, S. K., & Bowers, M. T. (2013). Factors that drive peptide assembly and fibril formation: Experimental and theoretical analysis of Sup35 NNQNY mutants. *Journal of Physical Chemistry B*, *117*(28), 8436–8446. <https://doi.org/10.1021/jp4046287>
- Ehrensperger, M., Gr aber, S., Gaestel, M., & Buchner, J. (1997). Binding of non-native protein to Hsp25 during heat shock creates a reservoir of folding intermediates for reactivation. *EMBO Journal*, *16*(2), 221–229. <https://doi.org/10.1093/emboj/16.2.221>
- Fichna, J. P., Potulska-Chromik, A., Miszta, P., Redowicz, M. J., Kaminska, A. M., Zekanowski, C., & Filipek, S. (2017). A novel dominant D109A CRYAB mutation in a family with myofibrillar myopathy affects  $\alpha$ B-crystallin structure. *BBA Clinical*, *7*, 1–7. <https://doi.org/10.1016/j.bbacli.2016.11.004>

- Fu, X., & Chang, Z. (2006). Identification of a highly conserved pro-gly doublet in non-animal small heat shock proteins and characterization of its structural and functional roles in *Mycobacterium tuberculosis* Hsp 16.3. *Biochemistry (Moscow)*, *71*. <https://doi.org/10.1134/S0006297906130141>
- Fu, X., Shi, X., Yin, L., Liu, J., Joo, K., Lee, J., & Chang, Z. (2013). Small heat shock protein IbpB acts as a robust chaperone in living cells by hierarchically activating its multi-type substrate-binding residues. *Journal of Biological Chemistry*, *288*(17), 11897–11906. <https://doi.org/10.1074/jbc.M113.450437>
- Fu, X., Zhang, H., Zhang, X., Cao, Y., Jiao, W., Liu, C., Song, Y., Abulimiti, A., & Chang, Z. (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), 6337–6348. <https://doi.org/10.1074/jbc.M406319200>
- Fukuhara, S., Nishigaki, T., Miyata, K., Tsuchiya, N., Waku, T., & Tanaka, N. (2012). Mechanism of the chaperone-like and antichaperone activities of amyloid fibrils of Peptides from  $\alpha$ -Crystallin. *Biochemistry*, *51*(27), 5394–5401. <https://doi.org/10.1021/bi3004236>
- Garrido, C., Paul, C., Seigneuric, R., & Kampinga, H. H. (2012). The small heat shock proteins family: The long forgotten chaperones. *International Journal of Biochemistry and Cell Biology*, *44*(10), 1588–1592. <https://doi.org/10.1016/j.biocel.2012.02.022>
- Gavrilov, Y., Dagan, S., & Levy, Y. (2015). Shortening a loop can increase protein native state entropy. *Proteins: Structure, Function and Bioinformatics*, *83*(12), 2137–2146. <https://doi.org/10.1002/prot.24926>
- Ghosh, J. G., & Clark, J. I. (2005). Insights into the domains required for dimerization and assembly of human  $\alpha$ B crystallin. *Protein Science: A Publication of the Protein Society*, *14*, 684–695. <https://doi.org/10.1110/ps.041152805>
- Ghosh, J. G., Estrada, M. R., & Clark, J. I. (2005). Interactive domains for chaperone activity in the small heat shock protein, human  $\alpha$ B crystallin. *Biochemistry*, *44*, 14854–14869. <https://doi.org/10.1021/bi0503910>
- Ghosh, J. G., Houck, S. a, & Clark, J. I. (2009). Interactive sequences in the stress protein and molecular chaperone human  $\alpha$ B crystallin recognize and modulate the assembly of filaments. *International Journal of Biochemistry*, *39*(10), 1804–1815. <https://doi.org/10.1016/j.biocel.2007.04.027>

- Giese, K. C., Basha, E., Catague, B. Y., & Vierling, E. (2005). Evidence for an essential function of the N terminus of a small heat shock protein in vivo, independent of in vitro chaperone activity. *Proceedings of the National Academy of Sciences*, *102*, 18896–18901. <https://doi.org/10.1073/pnas.0506169103>
- Giese, K. C., & Vierling, E. (2002). Changes in oligomerization are essential for the chaperone activity of a small heat shock protein in vivo and in vitro. *Journal of Biological Chemistry*, *277*(48), 46310–46318. <https://doi.org/10.1074/jbc.M208926200>
- Girirajan, S., Campbell, C., & Eichler, E. (2011). Protein Residue Contacts and Prediction Methods. *Physiology & Behavior*, *176*(5), 139–148. <https://doi.org/10.1007/978-1-4939-3572-7>
- Goldstein, R. A. (2007). Amino-acid interactions in psychrophiles, mesophiles, thermophiles, and hyperthermophiles: Insights from the quasi-chemical approximation. *Protein Science*, *16*(9), 1887–1895. <https://doi.org/10.1110/ps.072947007>
- Gromiha, M. M. (2001). Important inter-residue contacts for enhancing the thermal stability of thermophilic proteins. *Biophysical Chemistry*, *91*(1), 71–77. [https://doi.org/10.1016/S0301-4622\(01\)00154-5](https://doi.org/10.1016/S0301-4622(01)00154-5)
- Gromiha, M. M., & Selvaraj, S. (1997). Influence of Medium and Long Range Interactions in Different Structural Classes of Globular Proteins. *Journal of Biological Physics*, *23*(3), 151–162. <https://doi.org/10.1023/A:1004981409616>
- Gromiha, M. M., & Selvaraj, S. (2004). Inter-residue interactions in protein folding and stability. *Progress in Biophysics and Molecular Biology*, *86*(2), 235–277. <https://doi.org/10.1016/j.pbiomolbio.2003.09.003>
- Haidar, M., Asselbergh, B., Adriaenssens, E., De Winter, V., Timmermans, J. P., Auer-Grumbach, M., Juneja, M., & Timmerman, V. (2019). Neuropathy-causing mutations in HSPB1 impair autophagy by disturbing the formation of SQSTM1/p62 bodies. *Autophagy*, *15*(6), 1051–1068. <https://doi.org/10.1080/15548627.2019.1569930>
- Hanazono, Y., Takeda, K., Oka, T., Abe, T., Tomonari, T., Akiyama, N., Aikawa, Y., Yohda, M., & Miki, K. (2013). Nonequivalence observed for the 16-meric structure of a small heat shock protein, SpHsp16.0, from *Schizosaccharomyces pombe*. *Structure*, *21*(2), 220–228. <https://doi.org/10.1016/j.str.2012.11.015>
- Hanazono, Y., Takeda, K., Yohda, M., & Miki, K. (2012). Structural Studies on the Oligomeric Transition of a Small Heat Shock Protein, StHsp14.0. *Journal of Molecular Biology*, *422*(1), 100–108. <https://doi.org/10.1016/j.jmb.2012.05.017>

- Hartl, F. U., Bracher, A., & Hayer-hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. *Nature*, *475*, 324–332. <https://doi.org/10.1038/nature10317>
- Haslbeck, M., Braun, N., Stromer, T., Richter, B., Model, N., Weinkauff, S., & Buchner, J. (2004). Hsp42 is the general small heat shock protein in the cytosol of *Saccharomyces cerevisiae*. *EMBO Journal*, *23*(3), 638–649. <https://doi.org/10.1038/sj.emboj.7600080>
- Haslbeck, M., Ignatiou, A., Saibil, H., Helmich, S., Frenzl, E., Stromer, T., & Buchner, J. (2004). A domain in the N-terminal part of Hsp26 is essential for chaperone function and oligomerization. *Journal of Molecular Biology*, *343*, 445–455. <https://doi.org/10.1016/j.jmb.2004.08.048>
- Haslbeck, M., & Vierling, E. (2015). A First Line of Stress Defense: Small Heat Shock Proteins and Their Function in Protein Homeostasis. *Journal of Molecular Biology*, *427*(7), 1537–1548. <https://doi.org/10.1016/j.jmb.2015.02.002>
- Haslbeck, M., Weinkauff, S., & Buchner, J. (2019). Small heat shock proteins: Simplicity meets complexity. *Journal of Biological Chemistry*, *294*, 2121–2132. <https://doi.org/10.1074/jbc.REV118.002809>
- Hilario, E., Martin, F. J. M., Bertolini, M. C., & Fan, L. (2011). Crystal structures of xanthomonas small heat shock protein provide a structural basis for an active molecular chaperone oligomer. *Journal of Molecular Biology*, *408*(1), 74–86. <https://doi.org/10.1016/j.jmb.2011.02.004>
- Hochberg, G. K. A., & Benesch, J. L. P. (2014). Dynamical structure of  $\alpha$ B-crystallin. *Progress in Biophysics and Molecular Biology*, *115*(1), 11–20. <https://doi.org/https://doi.org/10.1016/j.pbiomolbio.2014.03.003>
- Janowska, M. K., Baughman, H. E. R., Woods, C. N., & Klevit, R. E. (2019). Mechanisms of small heat shock proteins. *Cold Spring Harbor Perspectives in Biology*, *11*. <https://doi.org/10.1101/cshperspect.a034025>
- Jaya, N., Garcia, V., & 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*, 15604–15609. <https://doi.org/10.1073/pnas.0902177106>
- KAGI, J. H., & VALLEE, B. L. (1960). The role of zinc in alcohol dehydrogenase. V. The effect of metal-binding agents on the structure of the yeast alcohol dehydrogenase molecule. *The Journal of Biological Chemistry*, *235*, 3188–3192.

- Kampinga, H. H., de Boer, R., & Beerstra, N. (2015). The Multicolored World of the Human HSPB Family. In R. M. Tanguay & L. E. Hightower (Eds.), *The Big Book on Small Heat Shock Proteins* (pp. 3–26). Springer International Publishing. [https://doi.org/10.1007/978-3-319-16077-1\\_1](https://doi.org/10.1007/978-3-319-16077-1_1)
- Kannan, R., Sreekumar, P. G., & Hinton, D. R. (2012). Novel roles for  $\alpha$ -crystallins in retinal function and disease. *Prog Retin Eye Res.*, *31*(6), 576–604. <https://doi.org/10.1016/j.preteyeres.2012.06.001>.Novel
- Kasakov, A. S., Bukach, O. V., Seit-Nebi, A. S., Marston, S. B., & Gusev, N. B. (2007). Effect of mutations in the  $\beta$ 5- $\beta$ 7 loop on the structure and properties of human small heat shock protein HSP22 (HspB8, H11). *FEBS Journal*, *274*, 5628–5642. <https://doi.org/10.1111/j.1742-4658.2007.06086.x>
- Khan, S., & Vihinen, M. (2010). Performance of protein stability predictors. *Human Mutation*, *31*(6), 675–684. <https://doi.org/10.1002/humu.21242>
- Kim, D. R., Lee, I., Ha, S. C., & Kim, K. K. (2003). Activation mechanism of HSP16.5 from *Methanococcus jannaschii*. *Biochemical and Biophysical Research Communications*, *307*, 991–998. [https://doi.org/10.1016/S0006-291X\(03\)01302-0](https://doi.org/10.1016/S0006-291X(03)01302-0)
- Kim, K. K., Kim, R., & Kim, S. H. (1998). Crystal structure of a small heat-shock protein. *Nature*, *394*(August), 595–599. <https://doi.org/10.1038/29106>
- Kim, R., Kim, K. K., Yokota, H., & Kim, S. H. (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), 9129–9133. <https://doi.org/10.1073/pnas.95.16.9129>
- Kim, M. V, Kasakov, A. S., Seit-Nebi, A. S., Marston, S. B., & Gusev, N. B. (2006). Structure and properties of K141E mutant of small heat shock protein HSP22 (HspB8, H11) that is expressed in human neuromuscular disorders. *Archives of Biochemistry and Biophysics*, *454*(1), 32–41. <https://doi.org/10.1016/j.abb.2006.07.014>
- Kocabiyik, S., & Özel, H. (2007). An extracellular-Pepstatin insensitive acid protease produced by *Thermoplasma volcanium*. *Bioresource Technology*, *98*, 112–117. <https://doi.org/10.1016/j.biortech.2005.11.016>
- Kocabiyik, S., & 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*, *47*(11), 1676–1683. <https://doi.org/10.1016/j.procbio.2011.11.014>

- Kumarevel, T. S., Gromiha, M. M., & Ponnuswamy, M. N. (1998). Analysis of hydrophobic and charged patches and influence of medium- and long-range interactions in molecular chaperones. *Biophysical Chemistry*, 75(2), 105–113. [https://doi.org/10.1016/S0301-4622\(98\)00198-7](https://doi.org/10.1016/S0301-4622(98)00198-7)
- Lee, G. J., Pokala, N., & Vierling, E. (1995). Structure and in vitro molecular chaperone activity of cytosolic small heat shock proteins from pea. *Journal of Biological Chemistry*, 270(18), 10432–10438. <https://doi.org/10.1074/jbc.270.18.10432>
- Lee, Garrett J., Roseman, A. M., Saibil, H. R., & Vierling, E. (1997). A small heat shock protein stably binds heat-denatured model substrates and can maintain a substrate in a folding-competent state. *EMBO Journal*, 16(3), 659–671. <https://doi.org/10.1093/emboj/16.3.659>
- Lentze, N., Studer, S., & Narberhaus, F. (2003a). Structural and functional defects caused by point mutations in the  $\alpha$ -crystallin domain of a bacterial  $\alpha$ -heat shock protein. *Journal of Molecular Biology*, 328(4), 927–937. [https://doi.org/10.1016/S0022-2836\(03\)00356-5](https://doi.org/10.1016/S0022-2836(03)00356-5)
- Lentze, N., Studer, S., & Narberhaus, F. (2003b). Structural and functional defects caused by point mutations in the  $\alpha$ -crystallin domain of a bacterial  $\alpha$ -heat shock protein. *Journal of Molecular Biology*, 328, 927–937. [https://doi.org/10.1016/S0022-2836\(03\)00356-5](https://doi.org/10.1016/S0022-2836(03)00356-5)
- Liu, L., Chen, J. Y., Yang, B., Wang, F. H., Wang, Y. H., & Yun, C. H. (2015). Active-State Structures of a Small Heat-Shock Protein Revealed a Molecular Switch for Chaperone Function. *Structure*, 23(11), 2066–2075. <https://doi.org/10.1016/j.str.2015.08.015>
- Liu, L., Chen, J., Yang, B., & Wang, Y. (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*, 458(2), 429–434. <https://doi.org/10.1016/j.bbrc.2015.01.134>
- Liu, Y., Zhang, X., Luo, L., Wu, M., Zeng, R., Cheng, G., Hu, B., Liu, B., Liang, J. J., & Shang, F. (2006). A novel  $\alpha$ B-crystallin mutation associated with autosomal dominant congenital lamellar cataract. *Investigative Ophthalmology and Visual Science*, 47(3), 1069–1075. <https://doi.org/10.1167/iovs.05-1004>
- Luo, H., & Robb, F. T. (2011). Thermophilic protein folding systems. In *Extremophiles handbook* (pp. 583–599). Springer. <https://doi.org/10.1007/978-4-431-53898-1>
- Magami, K., Hachiya, N., Morikawa, K., Fujii, N., & Takata, T. (2021). Isomerization of Asp is essential for assembly of amyloid-like fibrils of  $\alpha$ A-crystallin-derived peptide. *PLoS ONE*, 16(4 April), 1–12. <https://doi.org/10.1371/journal.pone.0250277>



- Maiti, P., Manna, J., Veleri, S., & Frautschy, S. (2014). Molecular chaperone dysfunction in neurodegenerative diseases and effects of curcumin. *BioMed Research International*, 2014. <https://doi.org/10.1155/2014/495091>
- Mani, N., Bhandari, S., Moreno, R., Hu, L., Venkataram Prasad, B. V., & Suguna, K. (2016). Multiple oligomeric structures of a bacterial small heat shock protein. *Nature Publishing Group, April*, 1–12. <https://doi.org/10.1038/srep24019>
- Mao, Q., & Chang, Z. (2001). Site-Directed Mutation on the Only Universally Conserved Residue Leu122 of Small Heat Shock Protein Hsp16 . 3. *Biochemical and Biophysical Research Communications*, 1261, 1257–1261. <https://doi.org/10.1006/bbrc.2001.6062>
- Matthew, J. B. (1985). ELECTROSTATIC EFFECTS IN PROTEINS. *Ann. Rev. Biophys. Biophys. Chem*, 14, 387–417.
- McHaourab, H. S., Lin, Y. L., & Spiller, B. W. (2012). Crystal structure of an activated variant of small heat shock protein Hsp16.5. *Biochemistry*, 51, 5105–5112. <https://doi.org/10.1021/bi300525x>
- Mishra, S., Chandler, S. A., Williams, D., Claxton, D. P., Koteiche, H. A., Stewart, P. L., Benesch, J. L. P., & Mchaourab, H. S. (2018). Engineering of a Polydisperse Small Heat-Shock Protein Reveals Conserved Motifs of Oligomer Plasticity. *Structure*, 26, 1116-1126.e4. <https://doi.org/10.1016/j.str.2018.05.015>
- Mogk, A., Ruger-herreros, C., & Bukau, B. (2019). Cellular Functions and Mechanisms of Action of Small Heat Shock Proteins. *Annual Review of Microbiology*, 73, 89–110.
- Moutaoufik, M. T., Morrow, G., Maaroufi, H., Féraud, C., Finet, S., & Tanguay, R. M. (2017). Oligomerization and chaperone-like activity of *Drosophila melanogaster* small heat shock protein DmHsp27 and three arginine mutants in the alpha-crystallin domain. *Cell Stress and Chaperones*, 22, 455–466. <https://doi.org/10.1007/s12192-016-0748-7>
- Muchowski, P. J., Wu, G. J. S., Liang, J. J. N., Adman, E. T., & Clark, J. I. (1999). Site-directed mutations within the core “ $\alpha$ -crystallin” domain of the small heat-shock protein, human  $\alpha$ B-crystallin, decrease molecular chaperone functions. *Journal of Molecular Biology*, 289, 397–411. <https://doi.org/10.1006/jmbi.1999.2759>
- Muranova, L. K., Sudnitsyna, M. V., Strelkov, S. V., & Gusev, N. B. (2020). Mutations in HspB1 and hereditary neuropathies. *Cell Stress and Chaperones*, 25(4), 655–665. <https://doi.org/10.1007/s12192-020-01099-9>

- Mymrikov, E. V., Daake, M., Richter, B., Haslbeck, M., & Buchner, J. (2017). The Chaperone Activity and Substrate Spectrum of Human Small Heat Shock Proteins. *The Journal of Biological Chemistry*, 292(2), 672–684. <https://doi.org/10.1074/jbc.M116.760413>
- Nahomi, R. B., Wang, B., Raghavan, C. T., Voss, O., Doseff, A. I., Santhoshkumar, P., & Nagaraj, R. H. (2013). Chaperone peptides of  $\alpha$ -crystallin inhibit epithelial cell apoptosis, protein insolubilization, and opacification in experimental cataracts. *Journal of Biological Chemistry*, 288(18), 13022–13035. <https://doi.org/10.1074/jbc.M112.440214>
- Nakamura, H. (1996). Roles of electrostatic interaction in proteins. *Quarterly Reviews of Biophysics*, 1, 1–90.
- Nandi, Sandip K., Rehna, E. a a, Panda, A. K., Shiburaj, S., Dharmalingam, K., & Biswas, A. (2013). A S52P mutation in the “‘a-crystallin domain’ of *Mycobacterium leprae* HSP18 reduces its oligomeric size and chaperone function.” *FEBS Journal*, 280, 5994–6009. <https://doi.org/10.1111/febs.12519>
- Nandi, Sandip Kumar, Panda, A. K., Chakraborty, A., Ray, S. S., & Biswas, A. (2015). Role of subunit exchange and electrostatic interactions on the chaperone activity of *Mycobacterium leprae* HSP18. *PLoS ONE*, 10. <https://doi.org/10.1371/journal.pone.0129734>
- Narberhaus, F. (2002).  $\alpha$ -Crystallin-Type Heat Shock Proteins: Socializing Minichaperones in the Context of a Multichaperone Network. *Microbiology and Molecular Biology Reviews*, 66(1), 64–93. <https://doi.org/10.1128/mmbr.66.1.64-93.2002>
- Nefedova, V. V., Sudnitsyna, M. V., Strelkov, S. V., & Gusev, N. B. (2013). Structure and properties of G84R and L99M mutants of human small heat shock protein HspB1 correlating with motor neuropathy. *Archives of Biochemistry and Biophysics*, 538, 16–24. <https://doi.org/10.1016/j.abb.2013.07.028>
- Obuchowski, I., Karaś, P., & Liberek, K. (2021). The Small Ones Matter—sHsps in the Bacterial Chaperone Network. *Frontiers in Molecular Biosciences*, 8(May), 1–7. <https://doi.org/10.3389/fmolb.2021.666893>
- Pasta, S. Y., Raman, B., Ramakrishna, T., & Rao, C. M. (2003). Role of the conserved SRLFDQFFG region of  $\alpha$ -crystallin, a small heat shock protein: Effect on oligomeric size, subunit exchange, and chaperone-like activity. *Journal of Biological Chemistry*, 278, 51159–51166. <https://doi.org/10.1074/jbc.M307523200>

- Perng, M. Der, Wen, S. F., Ijssel, P. Van Den, Prescott, A. R., & Quinlan, R. A. (2004). *Desmin Aggregate Formation by R120G  $\alpha$ B-Crystallin Is Caused by Altered Filament Interactions and Is Dependent upon Network Status in Cells*. *15*(May), 2335–2346. <https://doi.org/10.1091/mbc.E03>
- Phadte, A. S., Mahalingam, S., & Santhoshkumar, P. (2019). Functional Rescue of Cataract-Causing  $\alpha$ A-G98R-Crystallin by Targeted Compensatory Suppressor Mutations in Human  $\alpha$ A-Crystallin. *Biochemistry*, *58*, 4148–4158. <https://doi.org/10.1021/acs.biochem.9b00374>
- Ponnuswamy, P. K., & Michael Gromiha, M. (1994). On the conformational stability of folded proteins. *Journal of Theoretical Biology*, *166*,63-74. <https://doi.org/10.1006/jtbi.1994.1005>
- Poulain, P., Gelly, J. C., & Flatters, D. (2010). Detection and architecture of small heat shock protein Monomers. *PLoS ONE*, *5*(4). <https://doi.org/10.1371/journal.pone.0009990>
- Quinlan, R. a, Zhang, Y., Lansbury, A., Williamson, I., Pohl, E., & Sun, F. (2013). Changes in the quaternary structure and function of MjHSP16.5 attributable to deletion of the IXI motif and introduction of the substitution, R107G, in the  $\alpha$ -crystallin domain. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*, 20120327. <https://doi.org/10.1098/rstb.2012.0327>
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Raju, M., Santhoshkumar, P., Henzl, T. M., & Sharma, K. K. (2011). Identification and characterization of a copper-binding site in  $\alpha$ a-crystallin. *Free Radical Biology and Medicine*, *50*(10), 1429–1436. <https://doi.org/10.1016/j.freeradbiomed.2011.01.036>
- Raju, M., Santhoshkumar, P., & Sharma, K. K. (2012).  $\alpha$ A-Crystallin-Derived Mini-Chaperone Modulates Stability and Function of Cataract Causing  $\alpha$ AG98R-Crystallin. *PLoS ONE*, *7*(9). <https://doi.org/10.1371/journal.pone.0044077>
- Raju, M., Santhoshkumar, P., & Sharma, K. K. (2016). Alpha-crystallin-derived peptides as therapeutic chaperones. *Biochimica et Biophysica Acta*, *1860*(1), 246–251. <https://doi.org/10.1016/j.bbagen.2015.06.010>
- Raju, M., Santhoshkumar, P., Xie, L., & Sharma, K. K. (2014). Addition of  $\alpha$ a-crystallin sequence 164-173 to a mini-chaperone DFVIFLDVKHFSPEDLT alters the conformation but not the chaperone-like activity. *Biochemistry*, *53*, 2615–2623. <https://doi.org/10.1021/bi4017268>

- Reinle, K., Mogk, A., & Bukau, B. (2021). The Diverse Functions of Small Heat Shock Proteins in the Proteostasis Network. *Journal of Molecular Biology*, *434*(1), 167–157. <https://doi.org/10.1016/j.jmb.2021.167157>
- Robertson, A. L., Headey, S. J., Saunders, H. M., Ecroyd, H., Scanlon, M. J., Carver, J. a, & Bottomley, S. P. (2010). Small heat-shock proteins interact with a flanking domain to suppress polyglutamine aggregation. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(23), 10424–10429. <https://doi.org/10.1073/pnas.0914773107>
- Rodrigues, C. H. M., Pires, D. E. V., & Ascher, D. B. (2021). DynaMut2: Assessing changes in stability and flexibility upon single and multiple point missense mutations. *Protein Science*, *30*(1), 60–69. <https://doi.org/10.1002/pro.3942>
- Rutsdottir, G., Härmark, J., Weide, Y., Hebert, H., Rasmussen, M. I., Wernersson, S., Respondek, M., Akke, M., Højrup, P., Koeck, P. J. B., Söderberg, C. A. G., & Emanuelsson, C. (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), 8103–8121. <https://doi.org/10.1074/jbc.M116.766816>
- Sælensminde, G., Halskau, Ø., & Jonassen, I. (2009). Amino acid contacts in proteins adapted to different temperatures: Hydrophobic interactions and surface charges play a key role. *Extremophiles*, *13*(1), 11–20. <https://doi.org/10.1007/s00792-008-0192-4>
- Santhoshkumar, P., & Sharma, K. K. (2006). Conserved F84 and P86 residues in  $\alpha$ B-crystallin are essential to effectively prevent the aggregation of substrate proteins. *Protein Science*, *15*, 2488–2498. <https://doi.org/10.1110/ps.062338206>
- Santhoshkumar, Puttur, Raju, M., & Sharma, K. K. (2011).  $\alpha$ A-crystallin peptide 66SDRDKFVIFLDVKHF80 accumulating in aging lens impairs the function of  $\alpha$ -crystallin and induces lens protein aggregation. *PLoS ONE*, *6*(4). <https://doi.org/10.1371/journal.pone.0019291>
- Schägger, H. (2004). Blue-native gels to isolate protein complexes from mitochondria. *Methods in Cell Biology*, *65*, 231–244. [https://doi.org/10.1016/s0091-679x\(01\)65014-3](https://doi.org/10.1016/s0091-679x(01)65014-3)
- Sharma, K. Krishna, Kumar, R. S., Kumar, G. S., & Quinn, P. T. (2000). Synthesis and characterization of a peptide identified as a functional element in  $\alpha$ A-crystallin. *Journal of Biological Chemistry*, *275*(6), 3767–3771. <https://doi.org/10.1074/jbc.275.6.3767>

- Sharma, K. K. (2001). Conformational specificity of mini- $\alpha$ A-crystallin as a molecular chaperone. *Journal of Peptide Research*, *57*, 428–434.
- Shashidharamurthy, R., Koteiche, H. a., Dong, J., & Mchaourab, H. S. (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*, *280*(7), 5281–5289. <https://doi.org/10.1074/jbc.M407236200>
- Shroff, N. P., Bera, S., Cherian-Shaw, M., & Abraham, E. C. (2001). Substituted hydrophobic and hydrophilic residues at methionine-68 influence the chaperone-like function of  $\alpha$ B-crystallin. *Molecular and Cellular Biochemistry*, *220*, 127–133. <https://doi.org/10.1023/A:1010834107809>
- Shroff, Nilufer P., Cherian-Shaw, M., Bera, S., & Abraham, E. C. (2000). Mutation of R116C results in highly oligomerized  $\alpha$ A-Crystallin with modified structure and defective chaperone-like function. *Biochemistry*, *39*, 1420–1426. <https://doi.org/10.1021/bi991656b>
- Singh, B. N., Rao, K. S., Ramakrishna, T., Rangaraj, N., & Rao, C. M. (2007). Association of  $\alpha$ B-Crystallin, a Small Heat Shock Protein, with Actin: Role in Modulating Actin Filament Dynamics in Vivo. *Journal of Molecular Biology*, *366*, 756–767. <https://doi.org/10.1016/j.jmb.2006.12.012>
- Son, H. J., Shin, E. J., Nam, S. W., Kim, D. E., & Jeon, S. J. (2007). Properties of the  $\alpha$  subunit of a Chaperonin from the hyperthermophilic Crenarchaeon *Aeropyrum pernix* K1. *FEMS Microbiology Letters*, *266*, 103–109. <https://doi.org/10.1111/j.1574-6968.2006.00513.x>
- Stamler, R., Kappe, G., Boelens, W., & Slingsby, C. (2005). Wrapping the  $\alpha$ -crystallin domain fold in a chaperone assembly. *Journal of Molecular Biology*, *353*, 68–79. <https://doi.org/10.1016/j.jmb.2005.08.025>
- Strauch, A., & Haslbeck, M. (2016). The function of small heat-shock proteins and their implication in proteostasis. *Essays in Biochemistry*, *60*(2), 163–172. <https://doi.org/10.1042/EBC20160010>
- Stromer, T., Fischer, E., Richter, K., Haslbeck, M., & Buchner, J. (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), 11222–11228. <https://doi.org/10.1074/jbc.M310149200>
- Studer, S., Obrist, M., Lentze, N., & Narberhaus, F. (2002). A critical motif for oligomerization and chaperone activity of bacterial  $\alpha$ -heat shock proteins. *European Journal of Biochemistry*, *269*, 3578–3586. <https://doi.org/10.1046/j.1432-1033.2002.03049>

- Sun, Y., & MacRae, T. H. (2005). Small heat shock proteins: Molecular structure and chaperone function. *Cellular and Molecular Life Sciences*, *62*(21), 2460–2476.  
<https://doi.org/10.1007/s00018-005-5190-4>
- Takemoto, L. (1994). Release of  $\alpha$ -A Sequence 158-173 Correlates with a Decrease in the Molecular Chaperone Properties of Native  $\alpha$ -Crystallin. *Experimental Eye Research*, *59*, 239–242. <https://doi.org/10.1006/exer.1994.1103>
- Tanaka, N., Tanaka, R., Tokuhara, M., Kunugi, S., Lee, Y. F., & Hamada, D. (2008). Amyloid fibril formation and chaperone-like activity of peptides from  $\alpha$ A-Crystallin. *Biochemistry*, *47*(9), 2961–2967. <https://doi.org/10.1021/bi701823g>
- Thériault, J. R., Lambert, H., Chávez-Zobel, A. T., Charest, G., Lavigne, P., & Landry, J. (2004). Essential role of the NH<sub>2</sub>-terminal WD/EPF motif in the phosphorylation-activated protective function of mammalian Hsp27. *Journal of Biological Chemistry*, *279*(22), 23463–23471.  
<https://doi.org/10.1074/jbc.M402325200>
- Tikhomirova, T. S., Selivanova, O. M., & Galzitskaya, O. V. (2017).  $\alpha$ -Crystallins are small heat shock proteins: Functional and structural properties. In *Biochemistry (Moscow)* (Vol. 82, Issue 2, pp. 106–121). <https://doi.org/10.1134/S0006297917020031>
- Ungelenk, S., Moayed, F., Ho, C., Grousl, T., Scharf, A., Mashaghi, A., Tans, S., Mayer, M. P., Mogk, A., & Bukau, B. (2016). Small heat shock proteins sequester misfolding proteins in near-native conformation for cellular protection and efficient refolding. *Nature Communications*, *7*, 1–14. <https://doi.org/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*, 134–144. <https://doi.org/10.1110/ps.03264204>
- Usui, K., Yoshida, T., Maruyama, T., & Yohda, M. (2001). Small heat shock protein of a hyperthermophilic archaeum, *Thermococcus* sp. strain KS-1, exists as a spherical 24 mer and its expression is highly induced under heat-stress conditions. *Journal of Bioscience and Bioengineering*, *92*(2), 161–166. <https://doi.org/10.1263/jbb.92.161>
- Usui, Keisuke, Hatipoglu, O. F., Ishii, N., & Yohda, M. (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*, 113–118.  
<https://doi.org/10.1016/j.bbrc.2004.01.031>

- Van Montfort, R. L. M., Basha, E., Friedrich, K. L., Slingsby, C., & Vierling, E. (2001). Crystal structure and assembly of a eukaryotic small heat shock protein. *Nature Structural Biology*, 8(12), 1025–1030. <https://doi.org/10.1038/nsb722>
- Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25, 1189–1191. <https://doi.org/10.1093/bioinformatics/btp033>
- Wen, Z. Z., Wang, Y. H., Xu, X., Yang, B., Li, W. C., & Xie, M. Q. (2010). Importance of a Potential Salt Bridge and Hydrophobic Core in the Function and Oligomerization of a Small Heat Shock Protein. *Protein and Peptide Letters*, 17, 751–758.
- Wittig, I., Braun, H. P., & Schägger, H. (2006). Blue native PAGE. *Nature Protocols*, 1, 418–428. <https://doi.org/10.1038/nprot.2006.62>
- Xi, D., Wei, P., Zhang, C., & Lai, L. (2014). The minimal  $\alpha$ -crystallin domain of Mj Hsp16.5 is functional at non-heat-shock conditions. *Proteins: Structure, Function and Bioinformatics*, 82(October), 1156–1167. <https://doi.org/10.1002/prot.24480>
- Yang, C., Salerno, J. C., & Koretz, J. F. (2005). NH2-terminal stabilization of small heat shock protein structure: a comparison of two NH2-terminal deletion mutants of alphaA-crystallin. *Molecular Vision*, 11, 641–647.
- Yasuda, M., Oyaizu, H., Yamagishi, a, & Oshima, T. (1995). Morphological variation of new *Thermoplasma acidophilum* isolates from Japanese hot springs. *Applied and Environmental Microbiology*, 61(9), 3482–3485.
- Zhang, J., Zhao, X., Cai, Y., Li, Y., Yu, X., & Lu, L. (2015). Protection of Retina by Mini-  $\alpha$  A in NaIO<sub>3</sub>-Induced Retinal Pigment Epithelium Degeneration Mice. *International Journal of Molecular Sciences*, 16, 1644–1656. <https://doi.org/10.3390/ijms16011644>





## APPENDICES

### A. Buffer Solutions and Compositions

#### **LB Agar**

10 g of NaCl

10 g of tryptone

5 g of yeast extract

Add double distilled H<sub>2</sub>O to a final volume of 1 liter

Adjust pH to 7.0 with 5 N NaOH , and add 2% of agar. Autoclave at 121°C for 20 min.

#### **NZY+ Broth (per 250mL)**

2.5 g of NZ amine

1.25 g of yeast extract

1.25 g of NaCl

Add double distilled H<sub>2</sub>O to a final volume of 250 mL

Adjust to pH 7.5 using 5M and 1 M NaOH

Autoclave

Add the following filter-sterilized supplements prior to use (for 250 mL):

3.125 ml of 1 M MgCl<sub>2</sub>

3.125 ml of 1 M MgSO<sub>4</sub>

2.5 ml of 2 M glucose)

#### **Running Buffer ( 500 mL)**

1.51 gram of Tris

7.2 gram of Glycine

0.5 gram of SDS

### **2X-SDS Loading Dye**

0.775 gram of Tris

10 mL of Glycerol dissolve with 17.5 ml of ultra pure water, and pH is adjust with HCL and add

2 gram of SDS

5 mL of 2mercaptoethanol

1 mg Bromophenol Blue

Volume is completed with 50 mL of water.

### **50x TAE Buffer (500 mL), pH 8.4**

121.0 gram of Tris base

28.55 mL of glacial acetic acid

50 mL of 0.5 M EDTA

### **1X Dulbecco Phopshate Buffer Saline Preparation (200 mL)**

0.0402 gram of KCl

0.049 gram of  $\text{KH}_2\text{PO}_4$

1.6 gram of NaCl

0.536 gram of  $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$

Volume Completed to the 200 ml using sterilized ultra pure water, and pH is adjusted to the 7.2 with 5N HCl

Solution is sterilized by autoclave for 20 min.

## B. Molecular Size Markers

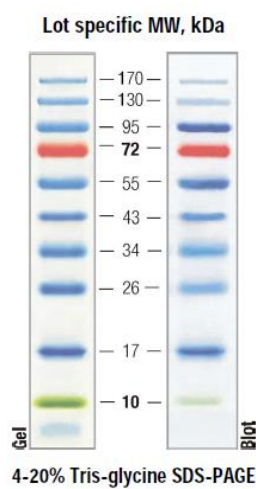


Figure B. 1. PageRuler Prestained Protein Ladder

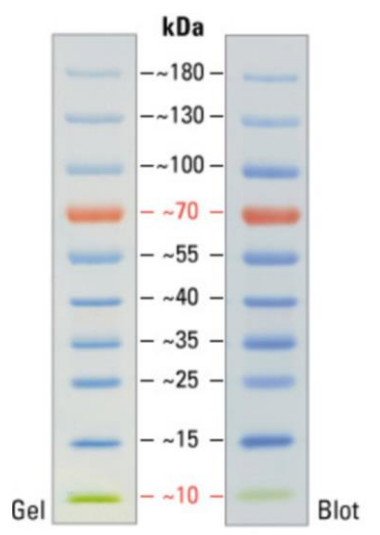


Figure B. 2. PageRuler Prestained Protein Ladder

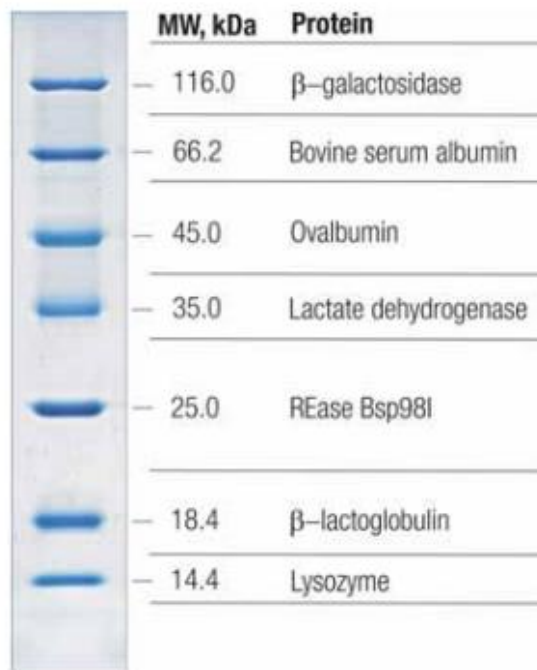


Figure B. 3. Unstained Protein Molecular Weight Marker (SM0431).

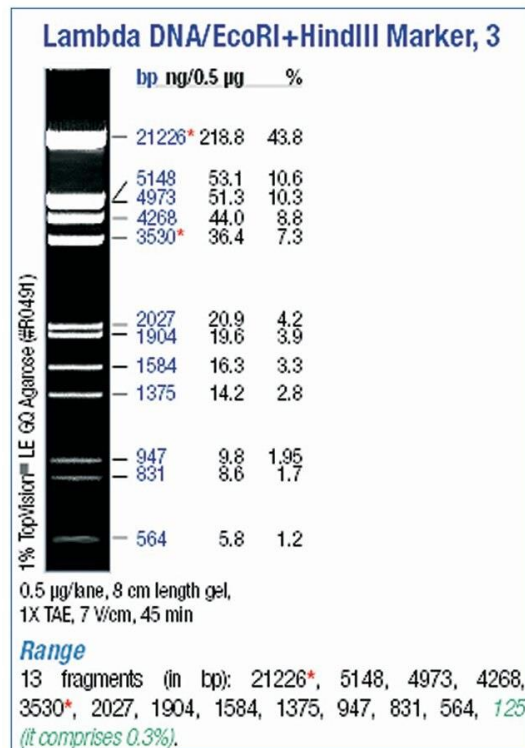


Figure B. 4. Lambda DNA/EcoRI plus HindIII Marker

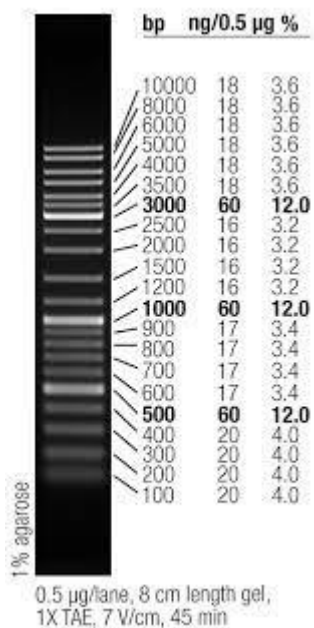


Figure B. 5. O'Gene ruler DNA Ladder mix

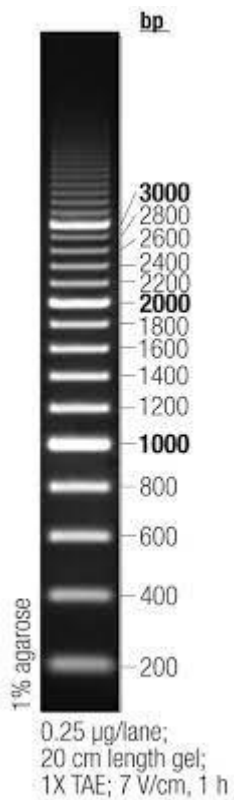


Figure B. 6. O' range ruler 200 bp DNA ladder

O'RangeRuler 500 bp DNA Ladder, ready-to-use

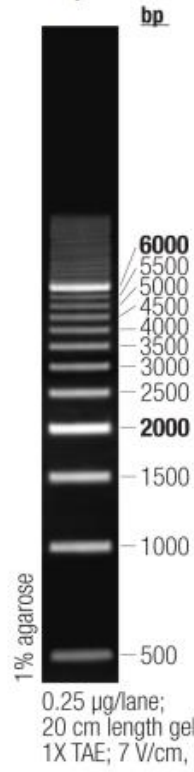


Figure B. 7. O' range ruler 500 bp DNA ladder

### C. Expression Vector Map

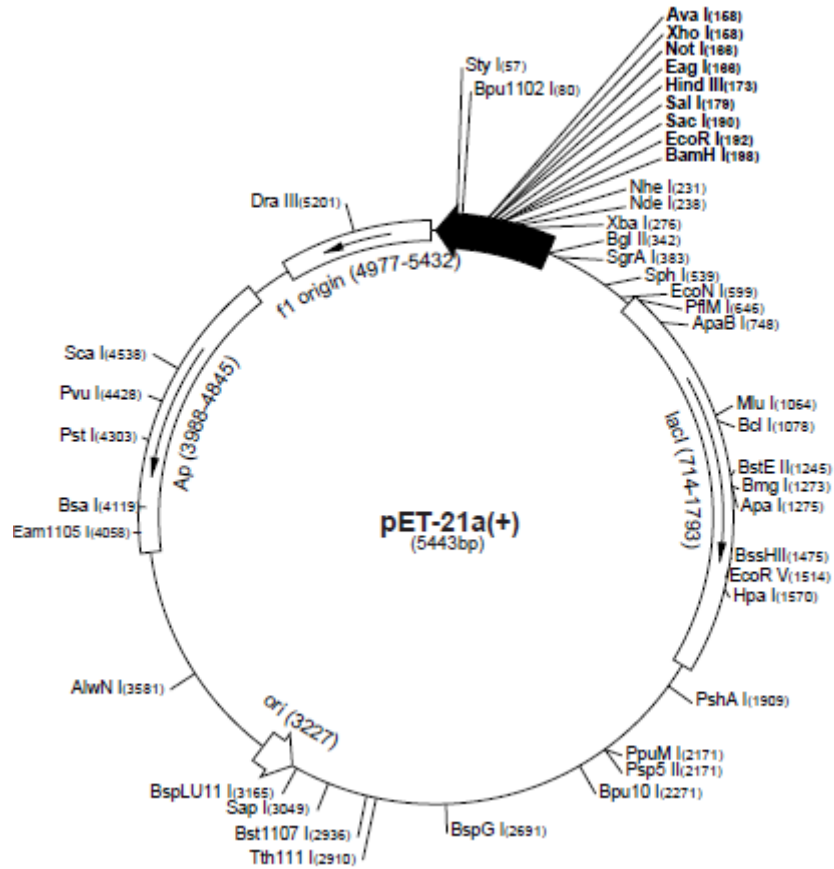


Figure C. 1. pET-21a (+) vector map

## D. pET 21-9 Sequencing Results

CLUSTAL O(1.2.4) multiple sequence alignment

```

pET21-9_T7ter      ACAGTATCCACATTCACCACCTGGAATTGATCTTTTCCCGGCGCTATTCATGCATACGTG
TVN0775            -----
pET21-9_T7        -----

pET21-9_T7ter      AAAGGTTTGCGCCATTCGATGTATCCGGGATCTAGACGCTCTACCTTATGCGACTCCTGC
TVN0775            -----
pET21-9_T7        -----

pET21-9_T7ter      ATTTAGGAAGCAGCCCAGTAGTAGTTTGAGGCCGTTGAGCACCGCCGCCAAGGAATGG
TVN0775            -----
pET21-9_T7        -----

pET21-9_T7ter      TGCATGCAAGGAGATGGCGCCCAACAGTCCCCGGCCACGGGGCCTGCCACCATACCCAC
TVN0775            -----
pET21-9_T7        -----

pET21-9_T7ter      GCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCATCGGTGATGTC
TVN0775            -----
pET21-9_T7        -----

pET21-9_T7ter      GGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCC
TVN0775            -----
pET21-9_T7        -----

pET21-9_T7ter      GGCGTAGAGGATCGAGATCTCGATCCCGCAAATTAATACGACTCACTATAGGGGAATTG
TVN0775            -----
pET21-9_T7        -----

pET21-9_T7ter      TGAGCGGATA-ACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATA
TVN0775            -----
pET21-9_T7        ---GCTGAAAGGAACGTTCCCTCTAGAA-TAATTTTGTTTAACTTTAAGAAGGAGATATA

pET21-9_T7ter      CATATGTATACACCCATAAAGTTCTTTACGAATGAGATGATAAAAAACGTATCGAATACT
TVN0775            ---ATGTATACACCCATAAAGTTCTTTACGAATGAGATGATAAAAAACGTATCGAATACT
pET21-9_T7        CATATGTATACACCCATAAAGTTCTTTACGAATGAGATGATAAAAAACGTATCGAATACT
                    *****

pET21-9_T7ter      GTGAAAGAGGTCTCATCCTTTATATATCCACCAGTCACGTTATATCAAGATAGCTCTGAT
TVN0775            GTGAAAGAGGTCTCATCCTTTATATATCCACCAGTCACGTTATATCAAGATAGCTCTGAT
pET21-9_T7        GTGAAAGAGGTCTCATCCTTTATATATCCACCAGTCACGTTATATCAAGATAGCTCTGAT
                    *****

pET21-9_T7ter      CTGGTATTGGAAGCAGAAATGGCCGGGTTTGACAAGAAAAACATAAAGGTCTCGGTAAT
TVN0775            CTGGTATTGGAAGCAGAAATGGCCGGGTTTGACAAGAAAAACATAAAGGTCTCGGTAAT
pET21-9_T7        CTGGTATTGGAAGCAGAAATGGCCGGGTTTGACAAGAAAAACATAAAGGTCTCGGTAAT
                    *****

```



**“Figure D.1 (Cont’d)”**

```

pET21-9_T7ter      AAGAATGTA...
TVN0775            AAGAATGTA...
pET21-9_T7        AAGAATGTA...
*****

pET21-9_T7ter      CAGCGCGTTGACAAAGTGTATGAGGTAGT...
TVN0775            CAGCGCGTTGACAAAGTGTATGAGGTAGT...
pET21-9_T7        CAGCGCGTTGACAAAGTGTATGAGGTAGT...
*****

pET21-9_T7ter      ATATCTGCTAAGTATAGTGAAGGCATACTTACAGTTAGAATGAAAACCAAGAACATAAAG
TVN0775            ATATCTGCTAAGTATAGTGAAGGCATACTTACAGTTAGAATGAAAACCAAGAACATAAAG
pET21-9_T7        ATATCTGCTAAGTATAGTGAAGGCATACTTACAGTTAGAATGAAAACCAAGAACATAAAG
*****

pET21-9_T7ter      AACGTAGAAATAGAAATAAATCATT...
TVN0775            AACGTAGAAATAGAAATAAATCATT...
pET21-9_T7        AACGTAGAAATAGAAATAAATCATT...
*****

pET21-9_T7ter      ATTGCTATAAGTATGCTTGATGTGAT...
TVN0775            ATTGCTATAAGTATGCTTGATGTGAT...
pET21-9_T7        ATTGCTATAAGTATGCTTGATGTGAT...

pET21-9_T7ter      CTGCGGCCGCACTCGAGCACCACCACCACCACCCTGAGATCCGGCTGCTAACAAAGCC
TVN0775            CTGCGGCCGCACTCGAGCACCACCACCACCACCCTGAGATCCGGCTGCTAACAAAGCC
pET21-9_T7        CTGCGGCCGCACTCGAGCACCACCACCACCACCCTGAGATCCGGCTGCTAACAAAGCC

pET21-9_T7ter      CGAAAGAA-GCTTATTTTAC-----
TVN0775            CGAAAGAA-GCTTATTTTAC-----
pET21-9_T7        CGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGG

pET21-9_T7ter      -----
TVN0775            -----
pET21-9_T7        GCCTCTAAACGGGTCTTGAGGGGTTTTTGGCTGAAAGGAGGAACTATATCCGGATTGGCG

pET21-9_T7ter      -----
TVN0775            -----
pET21-9_T7        AATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGGGTGTGGTGGTTACGCGCAGCG

pET21-9_T7ter      -----
TVN0775            -----
pET21-9_T7        TGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTCTTCCCTTCCCTTTC

pET21-9_T7ter      -----
TVN0775            -----
pET21-9_T7        TCGCCACGTTGCGCCGCTTTCGCCGTCAGCTCTAAATCGGGGGCTCCCTTAGGGTTCC

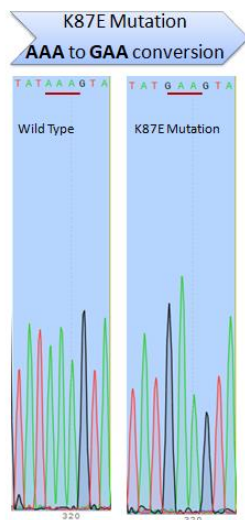
pET21-9_T7ter      -----
TVN0775            -----
pET21-9_T7        GATTTAGTGCTTTACGGCACCTCGACCCCAAAAACCTGATTAGGGTGATGGGTTACCGT

```

**“Figure D.1 (Cont’d)”**

```
pET21-9_T7ter -----  
TVN0775 -----  
pET21-9_T7      AGTGGCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACCGTTCTTTAA
```

```
pET21-9_T7ter -----  
TVN0775 -----  
pET21-9_T7      TAGTGGACTCTTGTTCCAAC
```



**Figure D. 1. Sequence result analysis of the pET21\_tvshSP9 recombinant plasmid with nucleotide msa analysis and chromatogram analysis. Start and stop codon of the *Tpv* sHSP 14.3 gene is colored by yellow and green, respectively. Point mutation at the position 87 by A to G transition is colored by blue. The codon change is underlined by red color in the chromatogram.**

## E. Wild Type Sequencing Results

CLUSTAL O(1.2.4) multiple sequence alignment

### pET21-2

```
pEt21_2_T7ter      TCACCCGTCCTTTAGCCACGAAAGCCGGACTTGGACCTTGGATTCAACCACGTGCGGTAA
TVN0775            -----
pET21-2_T7        -----
```

```
pEt21_2_T7ter      CGGTCTGATAAGAGACCACCGGGCTAGATCTGCGAGCTTTGTAATAGCCGTAACCTGTTGC
TVN0775            -----
pET21-2_T7        -----
```

```
pEt21_2_T7ter      ACATTCATCCACCCTGAATGGCGTCTTTTTCCGGACCTTTCATGGCATAACCGCGAAAGGT
TVN0775            -----
pET21-2_T7        -----
```

```
pEt21_2_T7ter      ATGCGCCATCGATGGTGTCCGGGATGTCGACGCTCTCCCATATGCGACTCCTACATTAG
TVN0775            -----
pET21-2_T7        -----
```

```
pEt21_2_T7ter      GAAGCAGCCCAGTAGTAGGTGAGGGCCGTTGAGCACCGCCGCCGCAAGAAATGGTGCAT
TVN0775            -----
pET21-2_T7        -----
```

```
pEt21_2_T7ter      GCAAGGAGATGGCGCCCAACAGTCCCCTCGGCCACGGGGCCTGCCACCATACCCCACGCC
TVN0775            -----
pET21-2_T7        -----
```

```
pEt21_2_T7ter      GAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCATCGGTGATGTCGGC
TVN0775            -----
pET21-2_T7        -----
```

```
pEt21_2_T7ter      GATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGGC
TVN0775            -----
pET21-2_T7        -----
```

```
pEt21_2_T7ter      GTAGAGGATCGAGATCTCGATCCCGGAAATTAATACGACTCACTATAGGGGAATTGTGA
TVN0775            -----
pET21-2_T7        -----
```

```
pEt21_2_T7ter      GCGGATAACAATTCCCCTCTAGAAATAATTTGTTTAACTTTAAGAAGGAGATATACATA
TVN0775            -----
pET21-2_T7        GGTTTAAGGTACATTCCCCTCTAGAATAATTTGTTTAACTTTAAGAAGGAGATATACATA
*

```

```
pEt21_2_T7ter      TGTATACACCATAAAGTTCTTTACGAATGAGATGATAAAAAACGTATCGAATACTGTGA
TVN0775            TGTATACACCATAAAGTTCTTTACGAATGAGATGATAAAAAACGTATCGAATACTGTGA
pET21-2_T7        TGTATACACCATAAAGTTCTTTACGAATGAGATGATAAAAAACGTATCGAATACTGTGA
*****
```

```
pEt21_2_T7ter      AAGAGGTCTCATCCTTTATATATCCACCAGTCACGTTATATCAAGATAGCTCTGATCTGG
TVN0775            AAGAGGTCTCATCCTTTATATATCCACCAGTCACGTTATATCAAGATAGCTCTGATCTGG
pET21-2_T7        AAGAGGTCTCATCCTTTATATATCCACCAGTCACGTTATATCAAGATAGCTCTGATCTGG
*****
```

**“Figure E.1 (Cont’d)”**

```

pEt21_2_T7ter      TATTGGAAGCAGAAATGGCCGGGTTTGACAAGAAAAACATAAAGGTCFCGGTAAATAAGA
TVN0775            TATTGGAAGCAGAAATGGCCGGGTTTGACAAGAAAAACATAAAGGTCFCGGTAAATAAGA
pET21-2_T7        TATTGGAAGCAGAAATGGCCGGGTTTGACAAGAAAAACATAAAGGTCFCGGTAAATAAGA
*****

pEt21_2_T7ter      ATGTACTCACTATAAGTGCCGGAGAGAAAGAGAGAATACTCTACCGTATATATCGATCAGC
TVN0775            ATGTACTCACTATAAGTGCCGGAGAGAAAGAGAGAATACTCTACCGTATATATCGATCAGC
pET21-2_T7        ATGTACTCACTATAAGTGCCGGAGAGAAAGAGAGAATACTCTACCGTATATATCGATCAGC
*****

pEt21_2_T7ter      GCGTTGACAAAGTGTATAAAGTAGTTAAGCTGCCCGTAGAGATTGAGCAGCAGGACATAT
TVN0775            GCGTTGACAAAGTGTATAAAGTAGTTAAGCTGCCCGTAGAGATTGAGCAGCAGGACATAT
pET21-2_T7        GCGTTGACAAAGTGTATAAAGTAGTTAAGCTGCCCGTAGAGATTGAGCAGCAGGACATAT
*****

pEt21_2_T7ter      CTGCTAAGTATAGTGAAGGCATACTTACAGTTAGAATGAAAACCAAGAACATAAAGAACG
TVN0775            CTGCTAAGTATAGTGAAGGCATACTTACAGTTAGAATGAAAACCAAGAACATAAAGAACG
pET21-2_T7        CTGCTAAGTATAGTGAAGGCATACTTACAGTTAGAATGAAAACCAAGAACATAAAGAACG
*****

pEt21_2_T7ter      TAGAAATAGAAATAAATCATTTTTTAATAATAATATATGAAAAGTATTGCATCTATTG
TVN0775            TAGAAATAGAAATAAATCATTTTTTAATAATAATATATGAAAAGTATTGCATCTATTG
pET21-2_T7        TAGAAATAGAAATAAATCATTTTTTAATAATAATATATGAAAAGTATTGCATCTATTG
*****

pEt21_2_T7ter      CTATAAGTATGCTTGATGTGATTGGGTGGGATCCGAATTCGAGCTCCGTCGACAAGCTTG
TVN0775            CTATAAGTATGCTTGATGTGATTGGGTGGGATCCGAATTCGAGCTCCGTCGACAAGCTTG
pET21-2_T7        CTATAAGTATGCTTGATGTGATTGGGTGGGATCCGAATTCGAGCTCCGTCGACAAGCTTG

pEt21_2_T7ter      CGGCCGCACTCGAGCACCACCACCACCACCCTGAGATCCGGCTGCTAACAAAGCCCGAA
TVN0775            CGGCCGCACTCGAGCACCACCACCACCACCCTGAGATCCGGCTGCTAACAAAGCCCGAA
pET21-2_T7        CGGCCGCACTCGAGCACCACCACCACCACCCTGAGATCCGGCTGCTAACAAAGCCCGAA

pEt21_2_T7ter      AGGAGGCGAGGGGGTAAGTTAGA-----
TVN0775            AGGAGGCGAGGGGGTAAGTTAGA-----
pET21-2_T7        AGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCT

pEt21_2_T7ter      -----
TVN0775            -----
pET21-2_T7        CTAAACGGGTCTTGAGGGGTTTTTGTCTGAAAGGAGGAACATATCCGGATTGGCGAATG

pEt21_2_T7ter      -----
TVN0775            -----
pET21-2_T7        GGACGCGCCCTGTAGCGCGCATTAAGCGCGCGGGTGTGGTGGTTACGCGCAGCGTGAC

pEt21_2_T7ter      -----
TVN0775            -----
pET21-2_T7        CGCTACACTTGCCAGCGCCCTAGCGCCGCTCCTTTCGCTTCTTCCCTTCCCTTCTCGC

pEt21_2_T7ter      -----
TVN0775            -----
pET21-2_T7        CACGTCGCGCGCTTTCGCCGTCAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCGGATT

pEt21_2_T7ter      -----
TVN0775            -----
pET21-2_T7        TAGTGCTTTACGGCACCTCGACCCCAAAAACCTTGATTAGGGTGATGGTTCACGTAGTGG

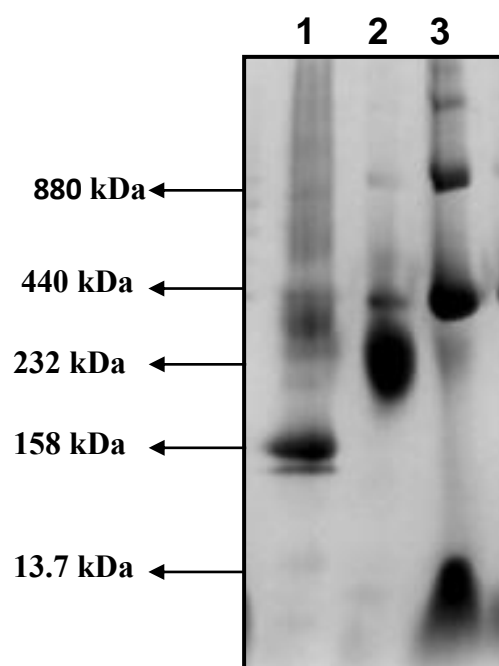
```

**“Figure E.1 (Cont’d)”**

```
pEt21_2_T7ter -----  
TVN0775 -----  
pET21-2_T7 GCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTATAGT  
  
pEt21_2_T7ter -----  
TVN0775 -----  
pET21-2_T7 GGACTCTTGTTCCTCAACTGGAACACACTCAACCCTATCTCGTCTATTCTTTTGATTTATAG  
  
pEt21_2_T7ter -----  
TVN0775 -----  
pET21-2_T7 ATTTGCCGATTTTCGCTATTGATAAAAAAATGAGCTGATTACAAAACCTACGCGATTTAAC  
  
pEt21_2_T7ter -----  
TVN0775 -----  
pET21-2_T7 AAATAATTACGCTTACATAAGTGCCACGTTTCGGCAATGTGCGCCGAACCCCTAT
```

**Figure E. 1. Sequence analysis of the cloned gene.** Sequence of the recombinant plasmids are verified by sequencing using both forward (T7) and reverse (T7 ter) primers. Start and stop codon of the *Tpv* sHSP 14.3 gene is colored by green and red, respectively. The gene is colored by yellow.

**F. Standart Proteins used in the BN-PAGE Analysis**



**Figure F. 1. Protein Standards. 1: Aldolase (158 kDa), 2: Catalase (232 kDa), 3: Ferritin (880 and 440 kDa) +Ribonuclease A (13.7 kDa) were used as standards.**

## G. Structural Bond Analysis of the *Tpv* sHSP 14.3 WT and its Mutants

**Table G. 1. Structural analysis of the *Tpv* sHSP 14.3 L33S mutant variant**

| Hydrogen Bond in WT       | Distance | Hydrophobic interactions in WT | Distance              | Hydrogen Bond in L33S     | Distance |
|---------------------------|----------|--------------------------------|-----------------------|---------------------------|----------|
| <b>Intramolecular</b>     |          |                                | <b>Intramolecular</b> |                           |          |
| B:LEU33:CA -<br>B:VAL41:O | 3,3591   | A:LEU33 - A:LEU42              | 4,66586               | B:SER33:CA -<br>B:VAL41:O | 3,16388  |
| A:LEU33:CA -<br>A:VAL41:O | 3,35931  | B:LEU33 - B:LEU42              | 4,66627               | A:SER33:CA -<br>A:VAL41:O | 3,16395  |
| <b>Intermolecular</b>     |          |                                | <b>Intermolecular</b> |                           |          |
| B:LEU33:N -<br>A:ILE78:O  | 3,07118  | A:LEU33 - B:ILE78              | 5,12996               | B:SER33:N -<br>A:ILE78:O  | 2,83934  |
| B:ILE78:N -<br>A:LEU33:O  | 3,02307  | A:ILE78 - B:LEU33              | 4,69937               | B:ILE78:N -<br>A:SER33:O  | 2,67419  |
|                           |          | A:TYR77 - B:LEU33              | 5,45735               | A:ILE78:N -<br>B:SER33:O  | 2,03009  |

**Table G. 2. Structural analysis of the *Tpv* sHSP 14.3 Y34F mutant variant**

| Hydrogen Bond in WT      | Distance | Hydrophobic interactions in WT | Distance | Hydrogen Bond in Y34F    | Distance | Hydrophobic interactions in Y34F | Distance |
|--------------------------|----------|--------------------------------|----------|--------------------------|----------|----------------------------------|----------|
| <b>Intramolecular</b>    |          |                                |          | <b>Intramolecular</b>    |          |                                  |          |
| B:TYR34:N -<br>B:VAL41:O | 3,12704  | A:TYR34 -<br>A:VAL41           | 4,03417  | B:PHE34:N -<br>B:VAL41:O | 3,00232  | B:VAL41:CB -<br>B:PHE34          | 3,86048  |
| B:VAL41:N -<br>B:TYR34:O | 2,98035  | B:TYR34 -<br>B:VAL41           | 4,03407  | B:VAL41:N -<br>B:PHE34:O | 3,00422  | A:VAL41:CB -<br>A:PHE34          | 3,86101  |
| A:TYR34:N -<br>A:VAL41:O | 3,12661  |                                |          | A:PHE34:N -<br>A:VAL41:O | 3,00165  |                                  |          |
| A:VAL41:N -<br>A:TYR34:O | 2,97915  |                                |          | A:VAL41:N -<br>A:PHE34:O | 3,00486  |                                  |          |

**Table G. 3. Structural analysis of the *Tpv* sHSP 14.3 G107A mutant variant**

| Hydrogen Bond in WT                | Distance | Hydrogen Bond in G107A | Distance                           | Hydrophobic Interactions in G107A | Distance |
|------------------------------------|----------|------------------------|------------------------------------|-----------------------------------|----------|
| <b>Intramolecular Interactions</b> |          |                        | <b>Intramolecular Interactions</b> |                                   |          |
| A:GLY107:CA -<br>A:TYR104:OH       | 3,50579  | -                      | -                                  | B:ALA47 - B:ALA107                | 2,93535  |
| B:GLY107:CA -<br>B:TYR104:OH       | 3,50606  | -                      | -                                  | A:ALA47 - A:ALA107                | 2,935    |

**Table G. 4. Structural analysis of the *Tpv* sHSP 14.3 E43V mutant variant**

| Hydrogen Bond in WT       | Distance | Electrostatic Interactions in WT        | Distance              | Hydrogen Bond in E43V | Distance |
|---------------------------|----------|---|-----------------------|-----------------------|----------|
| <b>Intramolecular</b>     |          |   | <b>Intramolecular</b> |                       |          |
| A:GLU43:N - A:THR32:O     | 2,82401  |   |                       | A:THR32:N - A:VAL43:O | 2,74732  |
| A:THR32:N - A:GLU43:O     | 2,7499   |   |                       | A:VAL43:N - A:THR32:O | 2,75528  |
| B:THR32:N - B:GLU43:O     | 2,74994  |   |                       | B:THR32:N - B:VAL43:O | 2,74676  |
| B:GLU43:N - B:THR32:O     | 2,8242   |   |                       | B:VAL43:N - B:THR32:O | 2,75591  |
| <b>Intermolecular</b>     |          |   | <b>Intermolecular</b> |                       |          |
| A:ARG81:CD - B:GLU43:O    | 3,36078  | A:ARG81:NH1 - B:GLU43:OE1 (salt bridge) | 2,9427                | None                  |          |
| A:ARG81:NH1 - B:GLU43:OE1 | 2,9427   | B:ARG81:NH1 - A:GLU43:OE1 (salt bridge) | 3,14283               |                       |          |
| B:ARG81:NH1 - A:GLU43:OE1 | 3,14283  | B:ARG81:NH2 - A:GLU43:OE2               | 5,54493               |                       |          |
|                           |          | A:ARG81:NH2 - B:GLU43:OE2               | 5,06808               |                       |          |

**Table G. 5. Structural analysis of the *Tpv* sHSP 14.3 E45G mutant variant**

| Hydrogen Bond in WT       | Distance | Electrostatic Interactions in WT        | Distance              | Hydrogen Bond in E45G | Distance |
|---------------------------|----------|---|-----------------------|-----------------------|----------|
| <b>Intramolecular</b>     |          |   | <b>Intramolecular</b> |                       |          |
| A:GLU45:N - A:PRO30:O     | 3,29     |   |                       | A:GLY45:N - A:PRO30:O | 3,2304   |
| B:GLU45:N - B:PRO30:O     | 3,29     |   |                       | B:GLY45:N - B:PRO30:O | 3,23079  |
| <b>Intermolecular</b>     |          |   | <b>Intermolecular</b> |                       |          |
| B:GLY48:CA - A:GLU45:OE2  | 2,94     | A:ARG81:NH2 - B:GLU45:OE2               | 4,67                  | None                  |          |
| A:GLY48:CA - B:GLU45:OE2  | 2,89     | B:ARG81:NH2 - A:GLU45:OE2               | 4,54                  |                       |          |
| A:ARG69:NH2 - B:GLU45:OE2 | 1,92     | A:ARG69:NH2 - B:GLU45:OE2 (salt bridge) | 1,92                  |                       |          |
| B:ARG69:NH2 - A:GLU45:OE2 | 2,03     | B:ARG69:NH2 - A:GLU45:OE2 (salt bridge) | 2,03                  |                       |          |
|                           |          | A:ARG69:NH1 - B:GLU45:OE1               | 4,92                  |                       |          |
|                           |          | B:ARG69:NH1 - A:GLU45:OE1               | 5,04                  |                       |          |
|                           |          | B:GLU45:OE2 - A:PHE49 (pi -anion)       | 4,73                  |                       |          |



**Table G. 6. Structural analysis of the *Tpv* sHSP 14.3 A47D mutant variant**

| Hydrogen Bond in WT          | Distance | Hydrophobic interactions in WT | Distance              | Hydrogen Bond in A47D       | Distance |
|------------------------------|----------|--------------------------------|-----------------------|-----------------------------|----------|
| <b>Intramolecular</b>        |          |                                | <b>Intramolecular</b> |                             |          |
| A: TYR28: OH<br>- A: ALA47:O | 3,22644  |                                |                       | A: TYR28: OH - A:ASP47:O    | 3,22425  |
| B: TYR28: OH<br>- B: ALA47:O | 3,2261   |                                |                       | B: TYR28: OH - B:ASP47:O    | 3,22411  |
|                              |          |                                |                       | A: ARG69:NH1 - A:ASP47:O    | 3,34672  |
|                              |          |                                |                       | B: ARG69:NH1 - B:ASP47:O    | 3,34701  |
|                              |          |                                |                       | B: GLY107:CA - B: ASP47:OD1 | 2,74495  |
|                              |          |                                |                       | A: GLY107:CA - A:ASP47:OD1  | 2,74594  |
| <b>Intermolecular</b>        |          |                                | <b>Intermolecular</b> |                             |          |
|                              |          | A:ALA47 B:ALA47                | 2,49937               | A:GLY107:CA -B:ASP47:OD1    | 1,77459  |
|                              |          |                                |                       | B:GLY107:CA -A:ASP47:OD1    | 1,59856  |

**Table G. 7. Structural analysis of the *Tpv* sHSP 14.3 G48E mutant variant**

| Hydrogen Bond in WT         | Distance | Hydrogen Bond in G48E      | Distance | Electrostatic Interaction in G48E                   | Distance |
|-----------------------------|----------|----------------------------|----------|---|----------|
| <b>Intramolecular</b>       |          | <b>Intramolecular</b>      |          |   |          |
| A:ARG69:NH2<br>- A: GLY48:O | 2,6697   | A:ARG69:CD -<br>A:GLU48:O  | 3,44711  | B:ARG81:NH2 -<br>B:GLU48:OE2<br>(Attractive Charge) | 5,50358  |
|                             |          | A:ARG69:NH1 -<br>A:GLU48:O | 3,10521  | A:ARG81:NH2 -<br>A:GLU48:OE2<br>(Attractive Charge) | 5,50316  |
| B:ARG69:NH2<br>- B:GLY48:O  | 2,66905  | B:ARG69:CD -<br>B:GLU48:O  | 3,44671  |   |          |
|                             |          | B:ARG69:NH1 -<br>B:GLU48:O | 3,10438  |   |          |
| <b>Intermolecular</b>       |          | <b>Intermolecular</b>      |          |   |          |
| B:GLY48:CA -<br>A:GLU45:OE2 | 2,9427   | A:GLU48:N -<br>B:GLU106:O  | 2,51986  | None  |          |
| A:GLY48:CA -<br>B:GLU45:OE2 | 2,89017  | B:GLU48:N -<br>A:GLU106:O  | 2,43155  |   |          |

**Table G. 8. Structural analysis of the *Tpv* sHSP 14.3 G107D mutant variant**

| Hydrogen Bond in WT                | Distance | Hydrogen Bond in G107D     | Distance                           | Electrostatic interaction in G107D                 | Distance |
|------------------------------------|----------|----------------------------|------------------------------------|--|----------|
| <b>Intramolecular Interactions</b> |          |                            | <b>Intramolecular Interactions</b> |  |          |
| A:GLY107:CA - A:TYR104:OH          | 3,50579  | A:ASP107:CA - A:TYR104:OH  | 3,71738                            |  |          |
| -                                  | -        | A:MET46:N - A:ASP107:O     | 2,95758                            |  |          |
|                                    |          |                            |                                    |  |          |
| B:GLY107:CA - B:TYR104:OH          | 3,50606  | B:ASP107:CA - B:TYR104:OH  | 3,71691                            |  |          |
| -                                  | -        | B:MET46:N - B:ASP107:O     | 2,95846                            |  |          |
| <b>Intermolecular Interactions</b> |          |                            | <b>Intermolecular Interactions</b> |  |          |
|                                    |          | A:TYR104:OH - B:ASP107:OD2 | 2,808                              | A:ARG69:NH1<br>B:ASP107:OD1<br>(Attractive charge) | 5,16406  |
|                                    |          | B:TYR104:OH - A:ASP107:OD2 | 2,79406                            | B:ARG69:NH1<br>A:ASP107:OD1<br>(Attractive charge) | 5,25879  |

**Table G. 9. Structural analysis of the *Tpv* sHSP 14.3 I108K mutant variant**

| Hydrogen Bond in WT     | Distance | Hydrogen Bond in I108K  | Distance              | Electrostatic Interaction in 108K | Distance |
|-------------------------|----------|-------------------------|-----------------------|-----------------------------------|----------|
| <b>Intramolecular</b>   |          |                         | <b>Intramolecular</b> |                                   |          |
| A:ILE108:N - A:SER105:O | 3,01432  | A:LYS108:N - A:SER105:O | 3,07421               | A:LYS108:NZ - A:GLU43:OE2         | 3,63498  |
| A:SER105:N - A:ILE108:O | 2,92962  | A:SER105:N - A:LYS108:O | 3,0008                | B:LYS108:NZ - B:GLU43:OE2         | 3,63513  |
| -                       | -        | A:LYS108:CA - A:ALA44:O | 3,29847               |                                   |          |
| B:ILE108:N - B:SER105:O | 3,01372  | B:LYS108:N - B:SER105:O | 3,07462               |                                   |          |
| B:SER105:N - B:ILE108:O | 2,92959  | B:SER105:N - B:LYS108:O | 3,00081               |                                   |          |
| -                       | -        | B:LYS108:CA - B:ALA44:O | 3,29739               |                                   |          |

**Table G. 10. Structural analysis of the *Tpv* sHSP 14.3 Y34FG48E double mutant variant**

| H-Bond in WT               | D    | Hydrophobic interactions in WT | D    | Hydrogen Bond in Y34F/G48E | D    | Hydrophobic Interactions in Y34FG48E | D    | Electrostatic Interaction in Y34FG48E         | D    |
|----------------------------|------|--------------------------------|------|----------------------------|------|--------------------------------------|------|---|------|
| <b>Intramolecular</b>      |      |                                |      | <b>Intramolecular</b>      |      |                                      |      |   |      |
| A:TYR34:N - A:VAL41:O      | 3,13 | A:TYR34 - A:VAL41              | 4,03 | A:PHE34:N - A:VAL41:O      | 3,11 | B:PHE34 - B:VAL41                    | 4,59 | A:ARG69:NH2 - A:GLU48:OE1 (Attractive Charge) | 5,36 |
| A:VAL41:N - A:TYR34:O      | 2,98 | B:TYR34 - B:VAL41              | 4,03 | A:VAL41:N - A:PHE34:O      | 2,96 | A:PHE34 - A:VAL41                    | 4,59 | B:ARG69:NH2 - B:GLU48:OE1 (Attractive Charge) | 5,36 |
| A:ARG69:NH2 - A:GLY48:O    | 2,67 |                                |      | A:ARG69:NH2 - A:GLU48:O    | 2,85 |                                      |      |   |      |
| --                         | --   |                                |      | A:ARG69:NE - A:GLU48:O     | 2,96 |                                      |      |   |      |
| B:TYR34:N - B:VAL41:O      | 3,12 |                                |      | B:PHE34:N - B:VAL41:O      | 3,11 |                                      |      |   |      |
| B:VAL41:N - B:TYR34:O      | 2,98 |                                |      | B:VAL41:N - B:PHE34:O      | 2,96 |                                      |      |   |      |
| B:ARG69:NH2 - B:GLY48:O    | 2,67 |                                |      | B:ARG69:NH2 - B:GLU48:O    | 2,85 |                                      |      |   |      |
| -                          | -    |                                |      | B:ARG69:NE - B:GLU48:O     | 2,97 |                                      |      |   |      |
| <b>Intermolecular</b>      |      |                                |      | <b>Intermolecular</b>      |      |                                      |      |   |      |
| B:GLY48:C A - A:GLU45:O E2 | 2,94 | None                           |      | A:GLU48:N - B:GLU106:O     | 2,60 | None                                 |      | None  |      |
| A:GLY48:C A - B:GLU45:O E2 | 2,89 |                                |      | B:GLU48:N - A:GLU106:O     | 2,71 |                                      |      |   |      |

**Table G. 11. Structural analysis of the *Tpv* sHSP 14.3 G48E1108K double mutant variant**

| Hydrogen Bond in WT      | Distance | Hydrogen Bond in G48E1108K | Distance | Electrostatic for G48E1108K                      | Distance |
|--------------------------|----------|----------------------------|----------|--|----------|
| <b>Intramolecular</b>    |          | <b>Intramolecular</b>      |          |  |          |
| A:ARG69:NH2 - A:GLY48:O  | 2,6697   | A:ARG69:CD - A:GLU48:O     | 3,3746   | A:ARG69:NH1 - A:GLU48:OE1                        | 5,07262  |
| A:ILE108:N - A:SER105:O  | 3,01432  | A:LYS108:N - A:SER105:O    | 3,15091  | A:ARG81:NH1 - A:GLU48:OE1<br>Salt bridge         | 3,73297  |
| A:SER105:N - A:ILE108:O  | 2,92962  | A:SER105:N - A:LYS108:O    | 2,84333  | A:ARG81:NH2 - A:GLU48:OE2                        | 4,67967  |
|                          |          | A:ARG69:NH1 - A:GLU48:O    | 3,17     | B:ARG69:NH1 - B:GLU48:OE1                        | 5,07148  |
|                          |          | A:ARG81:NH1 - A:GLU48:OE1  | 3,73297  | B:ARG81:NH1 - B:GLU48:OE1<br>Salt bridge         | 3,73292  |
|                          |          | A:ARG81:NH2 - A:GLU48:OE1  | 3,98175  | B:ARG81:NH2 - B:GLU48:OE2                        | 4,67994  |
| B:ILE108:N - B:SER105:O  | 3,01372  | B:LYS108:N - B:SER105:O    | 3,15139  |  |          |
| B:SER105:N - B:ILE108:O  | 2,92959  | B:SER105:N - B:LYS108:O    | 2,84324  |  |          |
| B:ARG69:NH2 - B:GLY48:O  | 2,66905  | B:ARG69:CD - B:GLU48:O     | 3,37458  |  |          |
|                          |          | B:ARG81:NH2 - B:GLU48:OE1  | 3,98092  |  |          |
|                          |          | B:ARG69:NH1 - B:GLU48:O    | 3,1696   |  |          |
|                          |          | B:ARG81:NH1 - B:GLU48:OE1  | 3,73292  |  |          |
| <b>Intermolecular</b>    |          | <b>Intermolecular</b>      |          |  |          |
| B:GLY48:CA - A:GLU45:OE2 | 2,9427   | -                          | -        | A:LYS108:NZ - B:GLU48:OE1<br>(Attractive Charge) | 4,70258  |
| A:GLY48:CA - B:GLU45:OE2 | 2,89017  | -                          | -        | B:LYS108:NZ - A:GLU48:OE1<br>(Attractive Charge) | 4,63377  |

**Table G. 12. Structural analysis of the *Tpv* sHSP 14.3 K87E mutant variant**

| <b>Hydrogen Bond in WT</b> | <b>Distance</b> | <b>Hydrophobic Interaction WT</b> | <b>Distance</b>       | <b>Hydrogen Bond in K87E</b> | <b>Distance</b> |
|----------------------------|-----------------|-----------------------------------|-----------------------|------------------------------|-----------------|
| <b>Intramolecular</b>      |                 |                                   | <b>Intramolecular</b> |                              |                 |
| A:ILE65:N - A:LYS87:O      | 3,01675         | A:LYS87 - A:ILE65                 | 4,21535               | A:ILE65:N - A:GLU87:O        | 2,94861         |
| A:LYS87:N - A:ILE65:O      | 2,82245         | A:LYS87 - A:VAL89                 | 5,20757               | A:GLU87:N - A:ILE65:O        | 2,75443         |
|                            |                 | A:VAL31 - A:LYS87                 | 5,26139               |                              |                 |
| A:THR64:CA - A:LYS87:O     | 3,32424         | B:LYS87 - B:ILE65                 | 4,21512               | A:THR64:CA - A:GLU87:O       | 3,24831         |
| B:ILE65:N - B:LYS87:O      | 3,01784         | B:LYS87 - B:VAL89                 | 5,20753               | B:ILE65:N - B:GLU87:O        | 2,94707         |
| B:LYS87:N - B:ILE65:O      | 2,82231         | B:VAL31 - B:LYS87                 | 5,26119               | B:GLU87:N - B:ILE65:O        | 2,75463         |
| B:THR64:CA - B:LYS87:O     | 3,32428         |                                   |                       | B:THR64:CA - B:GLU87:O       | 3,24668         |

## H. Rosetta Energy Level Analysis

Table H. 1. Comparative analysis of the rosetta energy levels of *Tpv* sHSP WT and mutants (L33S, Y34F, E43V, E45G, A47D, G48E)

| Position | WILD TYPE |                |              |                   |         | MUTANTS  |          |                |              |                   |         |
|----------|-----------|----------------|--------------|-------------------|---------|----------|----------|----------------|--------------|-------------------|---------|
|          | Sequence  | Rosetta Energy | Shape compl. | Area of Interface | C score | Position | Sequence | Rosetta Energy | Shape compl. | Area of Interface | C score |
| L33      | VTSYOD    | -17.700        | 0.8605       | 0                 | -30.608 | S33      | VTSYOD   | -17.900        | 0.6703       | 78                | -35.755 |
|          | TYODSS    | -21.200        | 0.7483       | 0                 | -32.425 |          | TYODSS   | -21.100        | 0.7483       | 43                | -36.624 |
|          | LYODSS    | -20.300        | 0.8456       | 0                 | -32.984 |          | SYODSS   | -22.600        | 0.7481       | 55                | -39.272 |
|          | VTLYOD    | -17.700        | 0.8605       | 0                 | -30.608 |          | VTLYOD   | -15.500        | 0.8605       | 60                | -34.358 |
|          | TLYODS    | -21.200        | 0.7483       | 0                 | -32.425 |          | TLYODS   | -21.200        | 0.8817       | 68                | -41.225 |
| Y34      | LYODSS    | -20.300        | 0.8456       | 0                 | -32.984 | F34      | LYODSS   | -19.500        | 0.8582       | 33                | -35.673 |
|          | YODSSD    | -18.000        | 0.8061       | 0                 | -30.091 |          | FODSSD   | -17.100        | 0.7947       | 62                | -35.221 |
|          | SDLVLE    | -20.900        | 0.8840       | 74                | -41.560 |          | SDLVLY   | -22.800        | 0.8840       | 74                | -43.460 |
|          | DI VLEA   | -20.400        | 0.0000       | 0                 | 0.000   |          | DI VLEA  | -24.400        | 0.0000       | 0                 | 0.000   |
|          | LVI VAE   | -20.500        | 0.9220       | 52                | -39.480 |          | LVI VAE  | -23.600        | 0.9126       | 51                | -42.389 |
| E43      | VLEAEM    | -18.100        | 0.7390       | 0                 | 0.000   | V43      | VLEAEM   | -21.700        | 0.8873       | 62                | -41.160 |
|          | EAE MAG   | -17.800        | 0.8715       | 0                 | -30.872 |          | LV AEMA  | -22.400        | 0.7994       | 90                | -43.341 |
|          | EAE MAG   | -17.800        | 0.8715       | 0                 | -30.872 |          | VAEMAG   | -20.000        | 0.8349       | 69                | -39.374 |
|          | LVI VAE   | -20.500        | 0.9220       | 52                | -39.480 |          | LVI VAE  | -20.300        | 0.9299       | 50                | -39.199 |
|          | LVAEMA    | -19.600        | 0.7850       | 0                 | 0.000   |          | LVAEMA   | -19.700        | 0.8170       | 94                | -41.355 |
| E45      | EAE MAG   | -17.800        | 0.8715       | 0                 | -30.872 | G45      | EAE MAG  | -19.000        | 0.8156       | 51                | -36.334 |
|          | EAE MAG   | -22.800        | 0.8961       | 0                 | -36.242 |          | AGMAGF   | -22.900        | 0.9061       | 55                | -41.992 |
|          | EMAGFD    | -18.300        | 0.8730       | 78                | -39.145 |          | GMAGFD   | -21.700        | 0.8136       | 65                | -40.404 |
|          | LEAEMA    | -19.600        | 0.7850       | 0                 | 0.000   |          | LEAEMD   | -16.700        | 0.8169       | 92                | -38.103 |
|          | EAE MAG   | -17.800        | 0.8715       | 0                 | -30.872 |          | EAE MDC  | -15.700        | 0.8117       | 96                | -37.425 |
| A47      | EAE MAG   | -22.800        | 0.8961       | 0                 | -36.242 | D47      | EAE MDC  | -19.700        | 0.9061       | 55                | -38.742 |
|          | EMAGFD    | -18.300        | 0.8730       | 78                | -39.145 |          | EMDGF    | -16.100        | 0.8250       | 92                | -37.625 |
|          | MA GFDK   | -18.800        | 0.8732       | 0                 | -31.898 |          | MDGFDK   | -16.200        | 0.8860       | 54                | -34.890 |
|          | AGFDKK    | -18.300        | 0.8331       | 0                 | -30.797 |          | DGFDKK   | -18.600        | 0.7607       | 58                | -35.810 |
|          | EAE MAG   | -17.800        | 0.8715       | 0                 | -30.872 |          | EAE MAF  | -17.100        | 0.8715       | 81                | -38.273 |
| G48      | EAE MAG   | -22.800        | 0.8961       | 0                 | -36.242 | E48      | EAE MAF  | -21.200        | 0.8649       | 83                | -42.474 |
|          | EMAGFD    | -18.300        | 0.8730       | 78                | -39.145 |          | EMAFD    | -19.000        | 0.8272       | 71                | -38.458 |
|          | MAGFDK    | -18.800        | 0.8732       | 0                 | -31.898 |          | MAEFDK   | -18.900        | 0.8718       | 90                | -40.977 |
|          | AGFDKK    | -18.300        | 0.8331       | 0                 | -30.797 |          | AEFDKK   | -18.200        | 0.8332       | 67                | -37.398 |
|          | GFDKKN    | -18.800        | 0.8127       | 0                 | -30.991 |          | EFDKKN   | -16.900        | 0.6701       | 64                | -33.301 |

Table H. 2. Comparative analysis of the rosetta energy levels of *Tpn* sHSP WT and mutants (G107A, G107D, I108K, K87E)

| Position | WILD TYPE |                |              |                   |         |          | MUTANTS  |                |              |                   |         |  |
|----------|-----------|----------------|--------------|-------------------|---------|----------|----------|----------------|--------------|-------------------|---------|--|
|          | Sequence  | Rosetta Energy | Shape compl. | Area of Interface | C score | Position | Sequence | Rosetta Energy | Shape compl. | Area of Interface | C score |  |
| G107     | AKYSEG    | -22.300        | 0.8246       | 0                 | -34.669 | A107     | AKYSEA   | -22.800        | 0.8203       | 29                | -37.955 |  |
|          | KYSEGI    | -21.800        | 0.8090       | 77                | -41.635 |          | KYSEAI   | -23.400        | 0.7246       | 72                | -41.469 |  |
|          | YSEGIL    | -20.400        | 0.7200       | 75                | -38.650 |          | YSEAIL   | -21.500        | 0.7200       | 72                | -39.450 |  |
|          | SEGLIT    | -21.700        | 0.7545       | 0                 | -33.017 |          | SEAILT   | -23.600        | 0.8411       | 77                | -43.916 |  |
|          | EGILTV    | -23.000        | 0.8820       | 80                | -44.230 |          | EAILTV   | -23.900        | 0.8820       | 80                | -45.130 |  |
|          | GILTVR    | -23.000        | 0.7120       | 59                | -39.580 |          | AILTVR   | -23.300        | 0.7786       | 59                | -40.879 |  |
|          |           |                |              |                   |         |          |          |                |              |                   |         |  |
| G107     | AKYSEG    | -22.300        | 0.8246       | 0                 | -34.669 | D107     | AKYSED   | -21.100        | 0.8203       | 29                | -36.255 |  |
|          | KYSEGI    | -21.800        | 0.8090       | 77                | -41.635 |          | KYSEDI   | -23.100        | 0.7098       | 99                | -43.597 |  |
|          | YSEGIL    | -20.400        | 0.7200       | 75                | -38.650 |          | YSEDIL   | -18.800        | 0.7200       | 75                | -37.050 |  |
|          | SEGLIT    | -21.700        | 0.7545       | 0                 | -33.017 |          | SEDILT   | -21.600        | 0.8460       | 76                | -41.890 |  |
|          | EGILTV    | -23.000        | 0.8820       | 80                | -44.230 |          | EDILTV   | -21.400        | 0.8820       | 80                | -42.630 |  |
|          | GILTVR    | -23.000        | 0.7120       | 59                | -39.580 |          | DILTVR   | -19.500        | 0.8066       | 70                | -38.599 |  |
|          |           |                |              |                   |         |          |          |                |              |                   |         |  |
| I108     | KYSEGI    | -21.800        | 0.8090       | 77                | -41.635 | K108     | KYSEGK   | -20.300        | 0.8482       | 75                | -40.523 |  |
|          | YSEGLI    | -20.400        | 0.7200       | 75                | -38.650 |          | YSEGKL   | -19.600        | 0.8943       | 67                | -39.715 |  |
|          | SEGLIT    | -21.700        | 0.7545       | 0                 | -33.017 |          | SEGLIT   | -21.400        | 0.7560       | 78                | -40.490 |  |
|          | EGILTV    | -23.000        | 0.8820       | 80                | -44.230 |          | EGKLTV   | -22.200        | 0.7950       | 72                | -41.275 |  |
|          | GILTVR    | -23.000        | 0.7120       | 59                | -39.580 |          | GKLTVR   | -20.800        | 0.7228       | 57                | -37.292 |  |
|          | ILTVRM    | -18.500        | 0.9104       | 0                 | -32.156 |          | KLTVRM   | -17.600        | 0.8396       | 104               | -40.544 |  |
|          | VDKVKY    | -19.400        | 0.8397       | 0                 | -31.995 | E87      | VDKVYE   | -18.500        | 0.8370       | 0                 | 0.000   |  |
| K87      | DKVYKV    | -20.600        | 0.7780       | 62                | -38.470 |          | DKVYEV   | -19.300        | 0.8125       | 69                | -38.388 |  |
|          | KVYKVV    | -23.800        | 0.8717       | 0                 | -36.876 |          | KVYEVV   | -23.200        | 0.8717       | 49                | -41.175 |  |
|          | VYKVVK    | -23.600        | 0.9509       | 0                 | -37.864 |          | VYEVVK   | -24.200        | 0.8020       | 71                | -43.330 |  |
|          | YKVVKL    | -23.900        | 0.8105       | 0                 | -36.057 |          | YEVVKL   | -23.400        | 0.7810       | 53                | -40.415 |  |
|          |           |                |              |                   |         |          |          |                |              |                   |         |  |
|          |           |                |              |                   |         |          |          |                |              |                   |         |  |
|          |           |                |              |                   |         |          |          |                |              |                   |         |  |

Table H. 3. Comparative analysis of the rosetta energy levels of *Trp* sHSP WT and mutants (Y34FG48E, and G48E1108K)

| Position                         | WILD TYPE               |                         |              |                   |         | Position | MUTANTS                          |                         |              |                   |         |         |
|----------------------------------|-------------------------|-------------------------|--------------|-------------------|---------|----------|----------------------------------|-------------------------|--------------|-------------------|---------|---------|
|                                  | Sequence                | Rosetta Energy          | Shape compl. | Area of Interface | C score |          | Sequence                         | Rosetta Energy          | Shape compl. | Area of Interface | C score |         |
| Y34G48                           | VTL <b>Y</b> QD         | -17.700                 | 0.8605       | 0                 | -30.608 | F34E48   | VTL <b>F</b> QD                  | -15.500                 | 0.8605       | 60                | -34.358 |         |
|                                  | TL <b>Y</b> QDS         | -21.200                 | 0.7483       | 0                 | -32.425 |          | TL <b>F</b> QDS                  | -21.200                 | 0.8817       | 68                | -41.225 |         |
|                                  | LYQDSS                  | -20.300                 | 0.8456       | 0                 | -32.984 |          | L <b>F</b> QDSS                  | -19.500                 | 0.8582       | 33                | -35.673 |         |
|                                  | <b>Y</b> QDSSD          | -18.000                 | 0.8061       | 0                 | -30.091 |          | <b>F</b> QDSSD                   | -17.100                 | 0.7947       | 62                | -35.221 |         |
|                                  | EA <b>M</b> AG          | -17.800                 | 0.8715       | 0                 | -30.872 |          | EA <b>M</b> AE                   | -17.100                 | 0.8715       | 81                | -38.273 |         |
|                                  | AE <b>M</b> AG <b>F</b> | -22.800                 | 0.8961       | 0                 | -36.242 |          | AE <b>M</b> AE <b>F</b>          | -21.200                 | 0.8649       | 83                | -42.474 |         |
|                                  | EMAG <b>F</b> D         | -18.300                 | 0.8730       | 78                | -39.145 |          | EMAE <b>F</b> D                  | -19.000                 | 0.8272       | 71                | -38.458 |         |
|                                  | MA <b>G</b> FDK         | -18.800                 | 0.8732       | 0                 | -31.898 |          | MA <b>F</b> FDK                  | -18.900                 | 0.8718       | 90                | -40.977 |         |
|                                  | A <b>G</b> FDKK         | -18.300                 | 0.8331       | 0                 | -30.797 |          | A <b>F</b> FDKK                  | -18.200                 | 0.8332       | 67                | -37.398 |         |
|                                  | <b>G</b> FDKKN          | -18.800                 | 0.8127       | 0                 | -30.991 |          | <b>F</b> FDKKN                   | -16.900                 | 0.6701       | 64                | -33.301 |         |
|                                  | G481108                 | EA <b>M</b> AG          | -17.800      | 0.8715            | 0       | -30.872  | E48K108                          | EA <b>M</b> AE          | -17.100      | 0.8715            | 81      | -38.273 |
|                                  |                         | AE <b>M</b> AG <b>F</b> | -22.800      | 0.8961            | 0       | -36.242  |                                  | AE <b>M</b> AE <b>F</b> | -21.200      | 0.8649            | 83      | -42.474 |
|                                  |                         | EMAG <b>F</b> D         | -18.300      | 0.8730            | 78      | -39.145  |                                  | EMAE <b>F</b> D         | -19.000      | 0.8272            | 71      | -38.458 |
|                                  |                         | MA <b>G</b> FDK         | -18.800      | 0.8732            | 0       | -31.898  |                                  | MA <b>F</b> FDK         | -18.900      | 0.8718            | 90      | -40.977 |
| A <b>G</b> FDKK                  |                         | -18.300                 | 0.8331       | 0                 | -30.797 |          | A <b>F</b> FDKK                  | -18.200                 | 0.8332       | 67                | -37.398 |         |
| <b>G</b> FDKKN                   |                         | -18.800                 | 0.8127       | 0                 | -30.991 |          | <b>F</b> FDKKN                   | -16.900                 | 0.6701       | 64                | -33.301 |         |
| K <b>Y</b> SE <b>G</b> I         |                         | -21.800                 | 0.8090       | 77                | -41.635 |          | K <b>Y</b> SE <b>G</b> K         | -20.300                 | 0.8482       | 75                | -40.523 |         |
| <b>Y</b> SE <b>G</b> I <b>L</b>  |                         | -20.400                 | 0.7200       | 75                | -38.650 |          | <b>Y</b> SE <b>G</b> K <b>L</b>  | -19.600                 | 0.8943       | 67                | -39.715 |         |
| SE <b>G</b> I <b>L</b> T         |                         | -21.700                 | 0.7545       | 0                 | -33.017 |          | SE <b>G</b> K <b>L</b> T         | -21.400                 | 0.7560       | 78                | -40.490 |         |
| E <b>G</b> I <b>L</b> T <b>V</b> |                         | -23.000                 | 0.8820       | 80                | -44.230 |          | E <b>G</b> K <b>L</b> T <b>V</b> | -22.200                 | 0.7950       | 72                | -41.275 |         |
| <b>G</b> I <b>L</b> T <b>V</b> R |                         | -23.000                 | 0.7120       | 59                | -39.580 |          | <b>G</b> K <b>L</b> T <b>V</b> R | -20.800                 | 0.7228       | 57                | -37.292 |         |
| <b>I</b> L <b>T</b> V <b>R</b> M |                         | -18.500                 | 0.9104       | 0                 | -32.156 |          | <b>K</b> L <b>T</b> V <b>R</b> M | -17.600                 | 0.8396       | 104               | -40.544 |         |



## I. Contact Map Analysis

**Table I. 1. List of the specific contacts which are present in the WT *Tpv* sHSP14.3, but not for the Y34F *Tpv* sHSP14.3 mutant and vice versa. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| WT <i>Tpv</i> sHSP14.3 Specific Contacts |           |         |                       |
|--|-----------|---------|-----------------------|
| Residue 1                                | Residue 2 | P-score | Contacts evaluation   |
| Ile27                                    | Ala67     | 0.5001  | long-range contacts   |
| Pro29                                    | Ile123    | 0.50904 | long-range contacts   |
| Pro30                                    | Ser66     | 0.5037  | long-range contacts   |
| Val31                                    | Ile123    | 0.50164 | long-range contacts   |
| Tyr34                                    | Val89     | 0.5068  | long-range contacts   |
| Gln35                                    | Gly107    | 0.5103  | long-range contacts   |
| Val41                                    | Glu124    | 0.52839 | long-range contacts   |
| Leu42                                    | Ile123    | 0.51384 | long-range contacts   |
| Phe49                                    | Leu109    | 0.508   | long-range contacts   |
| Ile54                                    | Ser105    | 0.50174 | long-range contacts   |
| Ile54                                    | Tyr104    | 0.51169 | long-range contacts   |
| Ile54                                    | Ile108    | 0.51304 | long-range contacts   |
| Ile54                                    | Leu109    | 0.51623 | long-range contacts   |
| Val56                                    | Gly107    | 0.51962 | long-range contacts   |
| Ser57                                    | Ser101    | 0.50227 | long-range contacts   |
| Ser57                                    | Lys103    | 0.50486 | long-range contacts   |
| Val58                                    | Glu106    | 0.50057 | long-range contacts   |
| Val58                                    | Ile100    | 0.53535 | long-range contacts   |
| Lys60                                    | Ser105    | 0.50049 | long-range contacts   |
| Lys60                                    | Ala102    | 0.50136 | long-range contacts   |
| Val62                                    | Ile100    | 0.53919 | long-range contacts   |
| Leu63                                    | Ile123    | 0.5108  | long-range contacts   |
| Ile65                                    | Ile123    | 0.50447 | long-range contacts   |
| Ile65                                    | Lys90     | 0.51142 | long-range contacts   |
| Ile65                                    | Asp99     | 0.51814 | long-range contacts   |
| Ala67                                    | Gln98     | 0.50843 | long-range contacts   |
| Leu42                                    | Lys61     | 0.51032 | Medium-range contacts |
| Asn61                                    | Gln80     | 0.51041 | Medium-range contacts |
| Ile65                                    | Tyr77     | 0.5043  | Medium-range contacts |
| Ser66                                    | Val88     | 0.50573 | Medium-range contacts |
| Ser66                                    | Val89     | 0.50728 | Medium-range contacts |
| Thr75                                    | Lys87     | 0.51489 | Medium-range contacts |
| Ala102                                   | Lys114    | 0.50025 | Medium-range contacts |
| Arg112                                   | Glu124    | 0.50503 | Medium-range contacts |
| Met12                                    | Val20     | 0.50916 | Short-range contacts  |
| Ile13                                    | Val20     | 0.62073 | Short-range contacts  |
| Tyr77                                    | Tyr86     | 0.50164 | Short-range contacts  |
| Ile78                                    | Val85     | 0.53982 | Short-range contacts  |

“Table I.1 (cont’d)”

| <b>Y34F <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|--|------------------|----------------|----------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| <b>Tyr28</b>   | Glu45            | 0.50476        | long-range contacts        |
| <b>Pro29</b>   | Leu91            | 0.51323        | long-range contacts        |
| <b>Pro29</b>   | Pro92            | 0.52828        | long-range contacts        |
| <b>Thr32</b>   | Asn61            | 0.5089         | long-range contacts        |
| <b>Phe34</b>   | Lys60            | 0.50224        | long-range contacts        |
| <b>Phe34</b>   | Lys114           | 0.5088         | long-range contacts        |
| <b>Phe34</b>   | Ser66            | 0.51614        | long-range contacts        |
| <b>Ser38</b>   | Ser66            | 0.50423        | long-range contacts        |
| <b>Ser38</b>   | Lys114           | 0.50794        | long-range contacts        |
| <b>Leu40</b>   | Thr115           | 0.50585        | long-range contacts        |
| <b>Leu40</b>   | Gln80            | 0.51102        | long-range contacts        |
| <b>Leu40</b>   | Val82            | 0.52113        | long-range contacts        |
| <b>Val41</b>   | Val82            | 0.5062         | long-range contacts        |
| <b>Leu42</b>   | Arg81            | 0.56162        | long-range contacts        |
| <b>Glu43</b>   | Gln80            | 0.50195        | long-range contacts        |
| <b>Lys52</b>   | Thr115           | 0.5009         | long-range contacts        |
| <b>Lys55</b>   | Pro92            | 0.50278        | long-range contacts        |
| <b>Val56</b>   | Tyr86            | 0.50921        | long-range contacts        |
| <b>Lys60</b>   | Val88            | 0.50213        | long-range contacts        |
| <b>Asn61</b>   | Val85            | 0.50029        | long-range contacts        |
| <b>Thr64</b>   | Thr110           | 0.5038         | long-range contacts        |
| <b>Phe26</b>   | Ala47            | 0.51891        | Medium-range contacts      |
| <b>Ile27</b>   | Ala47            | 0.52713        | Medium-range contacts      |
| <b>Tyr28</b>   | Ala44            | 0.55085        | Medium-range contacts      |
| <b>Tyr28</b>   | Met46            | 0.57889        | Medium-range contacts      |
| <b>Pro29</b>   | Gly48            | 0.52693        | Medium-range contacts      |
| <b>Gln35</b>   | Ser57            | 0.50806        | Medium-range contacts      |
| <b>Glu43</b>   | Asn59            | 0.50871        | Medium-range contacts      |
| <b>Asn61</b>   | Val82            | 0.52963        | Medium-range contacts      |
| <b>Asn61</b>   | Lys84            | 0.53198        | Medium-range contacts      |
| <b>Val62</b>   | Asp79            | 0.51149        | Medium-range contacts      |
| <b>Thr64</b>   | Val82            | 0.55751        | Medium-range contacts      |
| <b>Ile65</b>   | Val82            | 0.56827        | Medium-range contacts      |
| <b>Ala67</b>   | Val82            | 0.50637        | Medium-range contacts      |
| <b>Val76</b>   | Leu91            | 0.50745        | Medium-range contacts      |
| <b>Tyr86</b>   | Met              | 0.50585        | Medium-range contacts      |
| <b>Tyr86</b>   | Ile108           | 0.50607        | Medium-range contacts      |
| <b>Ile108</b>  | Val121           | 0.50349        | Medium-range contacts      |
| <b>Met1</b>  | Phe7             | 0.51255        | Short-range contacts       |
| <b>Ile5</b>  | Met12            | 0.56454        | Short-range contacts       |
| <b>Lys6</b>  | Ile13            | 0.55133        | Short-range contacts       |
| <b>Lys6</b>  | Met12            | 0.5719         | Short-range contacts       |
| <b>Phe49</b>   | Lys55            | 0.50121        | Short-range contacts       |
| <b>Ala67</b>   | Thr75            | 0.50273        | Short-range contacts       |
| <b>Lys70</b>   | Asp79            | 0.53037        | Short-range contacts       |
| <b>Glu72</b>   | Ile78            | 0.52369        | Short-range contacts       |
| <b>Gln80</b>   | Val88            | 0.54254        | Short-range contacts       |
| <b>Lys90</b>   | Asp99            | 0.50481        | Short-range contacts       |

**Table I. 2. List of residues forming contacts with the residue Y34 in WT *Tpv* sHSP14.3 and F34 in Y34F *Tpv* sHSP14.3 mutant.** Specific contacts between the WT and the mutant were highlighted by yellow color. Probability score is the probability of the two residues in contact within the 8 Å distance.

| WT <i>Tpv</i> sHSP14.3 Contacts at Position 34 |           |         |                      |
|--|-----------|---------|----------------------|
| Residue 1                                      | Residue 2 | P-Score | Contact evaluation   |
| Tyr34  | Val89     | 0.5068  | long-range contacts  |
| Tyr34  | Val58     | 0.51045 | long-range contacts  |
| Tyr34  | Val121    | 0.5154  | long-range contacts  |
| Tyr34  | Leu109    | 0.5287  | long-range contacts  |
| Tyr34  | Leu63     | 0.53396 | long-range contacts  |
| Tyr34  | Val88     | 0.53549 | long-range contacts  |
| Tyr34  | Lys87     | 0.54221 | long-range contacts  |
| Tyr34  | Ile65     | 0.56108 | long-range contacts  |
| Tyr34  | Ala44     | 0.50735 | Short-range contacts |
| Tyr34  | Glu43     | 0.62804 | Short-range contacts |
| Tyr34  | Val41     | 0.74313 | Short-range contacts |
| Tyr34  | Leu42     | 0.74691 | Short-range contacts |
| Tyr34  | Leu40     | 0.75437 | Short-range contacts |

| Y34F <i>Tpv</i> sHSP14.3 Mutant Contacts at Position 34 |           |         |                      |
|---|-----------|---------|----------------------|
| Residue 1   | Residue 2 | P-Score | Contact evaluation   |
| Phe34   | Val121    | 0.50087 | long-range contacts  |
| Phe34   | Leu109    | 0.50851 | long-range contacts  |
| Phe34   | Val58     | 0.52511 | long-range contacts  |
| Phe34   | Leu63     | 0.54192 | long-range contacts  |
| Phe34   | Ile65     | 0.55263 | long-range contacts  |
| Phe34   | Val88     | 0.55897 | long-range contacts  |
| Phe34   | Lys87     | 0.56625 | long-range contacts  |
| Phe34   | Lys60     | 0.50224 | long-range contacts  |
| Phe34   | Lys114    | 0.5088  | long-range contacts  |
| Phe34   | Ser66     | 0.51614 | long-range contacts  |
| Phe34   | Ala44     | 0.50891 | Short-range contacts |
| Phe34   | Glu43     | 0.6405  | Short-range contacts |
| Phe34   | Leu42     | 0.76226 | Short-range contacts |
| Phe34   | Val41     | 0.76607 | Short-range contacts |
| Phe34   | Leu40     | 0.77637 | Short-range contacts |

**Table I. 3. List of the specific contacts which are present in the WT *Tpv* sHSP14.3, but not for the G107A *Tpv* sHSP14.3 mutant and vice versa.** Probability score is the probability of the two residues in contact within the 8 Å distance.

| <b>WT <i>Tpv</i> sHSP14.3 Specific Contacts</b> |                  |                |                           |
|---|------------------|----------------|---------------------------|
| <b>Residue 1</b>                                | <b>Residue 2</b> | <b>P-score</b> | <b>Contact evaluation</b> |
| Ile27   | Ala67            | 0.5001         | long-range contacts       |
| Tyr28   | Ile123           | 0.51044        | long-range contacts       |
| Pro29   | Ile123           | 0.50904        | long-range contacts       |
| Pro30   | Ile123           | 0.53963        | long-range contacts       |
| Val31   | Ile123           | 0.50164        | long-range contacts       |
| Thr32   | Glu122           | 0.5108         | long-range contacts       |
| Thr32   | Ile123           | 0.51263        | long-range contacts       |
| Leu33   | Glu122           | 0.51057        | long-range contacts       |
| Leu33   | Val121           | 0.52608        | long-range contacts       |
| Tyr34   | Val121           | 0.5154         | long-range contacts       |
| Ser38   | Thr115           | 0.50279        | long-range contacts       |
| Ser38   | Met113           | 0.53874        | long-range contacts       |
| Asp39   | Glu124           | 0.55505        | long-range contacts       |
| Asp39   | Ile123           | 0.55527        | long-range contacts       |
| Leu40   | Glu122           | 0.53661        | long-range contacts       |
| Val41   | Glu124           | 0.52839        | long-range contacts       |
| Leu42   | Asn61            | 0.51032        | long-range contacts       |
| Leu42   | Ile123           | 0.51384        | long-range contacts       |
| Leu42   | Asp79            | 0.5186         | long-range contacts       |
| Ala47   | Ala102           | 0.50309        | long-range contacts       |
| Ala47   | Glu106           | 0.52557        | long-range contacts       |
| Phe49   | Leu109           | 0.508          | long-range contacts       |
| Ile54   | Ser105           | 0.50174        | long-range contacts       |
| Ile54   | Tyr104           | 0.51169        | long-range contacts       |
| Ile54   | Ile108           | 0.51304        | long-range contacts       |
| Ile54   | Leu109           | 0.51623        | long-range contacts       |
| Val56   | Gly107           | 0.51962        | long-range contacts       |
| Ser57   | Lys103           | 0.50486        | long-range contacts       |
| Lys60   | Ser105           | 0.50049        | long-range contacts       |
| Lys60   | Ala102           | 0.50136        | long-range contacts       |
| Lys60   | Val89            | 0.54637        | long-range contacts       |
| Asn61   | Gln80            | 0.51041        | long-range contacts       |

“Table I.3 (cont’d)”

| WT <i>Tpv</i> sHSP14.3 Specific Contacts |           |         |                      |
|--|-----------|---------|----------------------|
| Residue 1                                | Residue 2 | P-score | Contacts evaluation  |
| Asn61                                    | Pro92     | 0.54997 | long-range contacts  |
| Val62                                    | Ile123    | 0.53086 | long-range contacts  |
| Leu63                                    | Ile123    | 0.5108  | long-range contacts  |
| Ile65                                    | Ile123    | 0.50447 | long-range contacts  |
| Ile65                                    | Lys90     | 0.51142 | long-range contacts  |
| Ser66                                    | Val88     | 0.50573 | long-range contacts  |
| Ser66                                    | Val82     | 0.54109 | long-range contacts  |
| Ala67                                    | Gln98     | 0.50843 | long-range contacts  |
| Ala67                                    | Tyr86     | 0.54075 | long-range contacts  |
| Ala67                                    | Val85     | 0.55527 | long-range contacts  |
| Glu68                                    | Gln80     | 0.54192 | long-range contacts  |
| Glu68                                    | Val85     | 0.56551 | long-range contacts  |
| Thr75                                    | Leu91     | 0.50275 | long-range contacts  |
| Thr75                                    | Pro92     | 0.51105 | long-range contacts  |
| Thr75                                    | Lys87     | 0.51489 | long-range contacts  |
| Ala102                                   | Lys114    | 0.50025 | long-range contacts  |
| Met12                                    | Val20     | 0.50916 | Short-range contacts |
| Ile13                                    | Val20     | 0.62073 | Short-range contacts |
| Arg69                                    | Gln80     | 0.50341 | Short-range contacts |
| Arg69                                    | Asp79     | 0.58868 | Short-range contacts |
| Arg71                                    | Ile78     | 0.5772  | Short-range contacts |
| Val76                                    | Lys87     | 0.56952 | Short-range contacts |
| Tyr77                                    | Tyr86     | 0.50164 | Short-range contacts |

“Table I.3 (cont’d)”.

| <b>G107A <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>  | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Pro29   | Leu91            | 0.50184        | long-range contacts        |
| Pro29   | Pro92            | 0.51503        | long-range contacts        |
| Thr32   | Asn61            | 0.50879        | long-range contacts        |
| Tyr34   | Val111           | 0.50146        | long-range contacts        |
| Gln35   | Thr64            | 0.50293        | long-range contacts        |
| Ser38   | Ser66            | 0.50236        | long-range contacts        |
| Leu40   | Leu109           | 0.53487        | long-range contacts        |
| Leu42   | Arg81            | 0.55778        | long-range contacts        |
| Glu43   | Gln80            | 0.51832        | long-range contacts        |
| Glu43   | Tyr104           | 0.51308        | long-range contacts        |
| Glu43   | Val121           | 0.50108        | long-range contacts        |
| Lys52   | Thr115           | 0.50278        | long-range contacts        |
| Asn53   | Val93            | 0.50463        | long-range contacts        |
| Lys55   | Pro92            | 0.50771        | long-range contacts        |
| Val56   | Tyr86            | 0.50568        | long-range contacts        |
| Ser57   | Val85            | 0.51307        | long-range contacts        |
| Val58   | Ser101           | 0.50042        | long-range contacts        |
| Asn61   | Val85            | 0.50313        | long-range contacts        |
| Asp83   | Arg112           | 0.50205        | long-range contacts        |
| Phe26   | Ala47            | 0.50652        | Medium-range contacts      |
| Tyr28   | Ala44            | 0.50292        | Medium-range contacts      |
| Tyr28   | Met46            | 0.5122         | Medium-range contacts      |
| Gln35   | Ser57            | 0.51115        | Medium-range contacts      |
| Glu43   | Asn59            | 0.50558        | Medium-range contacts      |
| Lys60   | Arg81            | 0.50095        | Medium-range contacts      |
| Asn61   | Val82            | 0.53047        | Medium-range contacts      |
| Asn61   | Lys84            | 0.51356        | Medium-range contacts      |
| Val62   | Asp79            | 0.53978        | Medium-range contacts      |
| Thr64   | Tyr77            | 0.50928        | Medium-range contacts      |
| Thr64   | Val82            | 0.51002        | Medium-range contacts      |
| Ile65   | Val82            | 0.52125        | Medium-range contacts      |
| Ser74   | Lys90            | 0.55044        | Medium-range contacts      |

“Table I.3 (cont’d)”

| <b>G107A <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                      |                |                            |
|---|----------------------|----------------|----------------------------|
| <b>Residue<br/>1</b>                                      | <b>Residue<br/>2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Ser74   | Leu91                | 0.54745        | Medium-range contacts      |
| Asp83   | Tyr104               | 0.50294        | Medium-range contacts      |
| Tyr86   | Asp99                | 0.50746        | Medium-range contacts      |
| Tyr86   | Ile108               | 0.50591        | Medium-range contacts      |
| Ile108  | Val121               | 0.50373        | Medium-range contacts      |
| Met1  | Phe7                 | 0.5032         | Short-range contacts       |
| Ala67   | Thr75                | 0.53247        | Short-range contacts       |
| Glu68   | Ser74                | 0.53224        | Short-range contacts       |
| Lys90   | Asp99                | 0.5066         | Short-range contacts       |

**Table I. 4. List of residues forming contacts with the residue G107 in WT *Tpv* sHSP14.3 and A107 in G107A *Tpv* sHSP14.3 mutant.** Probability score is the probability of the two residues in contact within the 8 Å distance.

| <b>WT <i>Tpv</i> sHSP14.3 Contacts at Position 107</b> |                      |                     |                           |
|--|----------------------|---------------------|---------------------------|
| <b>Residue<br/>1</b>                                   | <b>Residue<br/>2</b> | <b>P-<br/>Score</b> | <b>Contact evaluation</b> |
| Gly107   | Ile123               | 0.64729             | Medium-range contacts     |
| Gly107   | Glu124               | 0.64104             | Medium-range contacts     |

| <b>G107A <i>Tpv</i> sHSP14.3 Mutant Contacts at Position 107</b> |                      |                |                           |
|--|----------------------|----------------|---------------------------|
| <b>Residue<br/>1</b>   | <b>Residue<br/>2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| Ala107   | Ile123               | 0.63422        | Medium-range contacts     |
| Ala107   | Glu124               | 0.62475        | Medium-range contacts     |

**Table I. 5. List of the specific contacts which are present in the WT *Tpv* sHSP14.3, but not for the L33S *Tpv* sHSP14.3 mutant and vice versa. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| WT <i>Tpv</i> sHSP14.3 Specific Contacts |              |         |                       |
|--|--------------|---------|-----------------------|
| Residue<br>1                             | Residue<br>2 | P-score | Contacts evaluation   |
| Ile27                                    | Ala67        | 0.5001  | long-range contacts   |
| Val31                                    | Ile123       | 0.50164 | long-range contacts   |
| Gln35                                    | Gly107       | 0.5103  | long-range contacts   |
| Asp39                                    | Ile123       | 0.55527 | long-range contacts   |
| Val41                                    | Glu124       | 0.52839 | long-range contacts   |
| Leu42                                    | Ile123       | 0.51384 | long-range contacts   |
| Ala47                                    | Ala102       | 0.50309 | long-range contacts   |
| Phe49                                    | Leu109       | 0.508   | long-range contacts   |
| Ile54                                    | Ser105       | 0.50174 | long-range contacts   |
| Ile54                                    | Tyr104       | 0.51169 | long-range contacts   |
| Ile54                                    | Ile108       | 0.51304 | long-range contacts   |
| Ile54                                    | Leu109       | 0.51623 | long-range contacts   |
| Val56                                    | Gly107       | 0.51962 | long-range contacts   |
| Ser57                                    | Ser101       | 0.50227 | long-range contacts   |
| Ser57                                    | Lys103       | 0.50486 | long-range contacts   |
| Val58                                    | Glu106       | 0.50057 | long-range contacts   |
| Val58                                    | Ile100       | 0.53535 | long-range contacts   |
| Lys60                                    | Ser105       | 0.50049 | long-range contacts   |
| Lys60                                    | Ala102       | 0.50136 | long-range contacts   |
| Val62                                    | Ile100       | 0.53919 | long-range contacts   |
| Leu63                                    | Ile123       | 0.5108  | long-range contacts   |
| Ile65                                    | Ile123       | 0.50447 | long-range contacts   |
| Ile65                                    | Asp99        | 0.51814 | long-range contacts   |
| Ala67                                    | Gln98        | 0.50843 | long-range contacts   |
| Leu42                                    | Asn61        | 0.51032 | Medium-range contacts |
| Asn61                                    | Gln80        | 0.51041 | Medium-range contacts |
| Ile65                                    | Tyr77        | 0.5043  | Medium-range contacts |
| Ser66                                    | Val88        | 0.50573 | Medium-range contacts |
| Ser66                                    | Val89        | 0.50728 | Medium-range contacts |
| Thr75                                    | Lys87        | 0.51489 | Medium-range contacts |
| Ala102                                   | Lys114       | 0.50025 | Medium-range contacts |
| Arg112                                   | Glu124       | 0.50503 | Medium-range contacts |



**“Table I. 5 (cont’d)”**

| <b>WT <i>Tpv</i> sHSP14.3 Specific Contacts</b> |                      |                |                            |
|---|----------------------|----------------|----------------------------|
| <b>Residue<br/>1</b>                            | <b>Residue<br/>2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Met12   | Val20                | 0.50916        | Short-range contacts       |
| Ile13   | Val20                | 0.62073        | Short-range contacts       |
| Tyr77   | Tyr86                | 0.50164        | Short-range contacts       |
| Ile78   | Val85                | 0.53982        | Short-range contacts       |
| Pro92   | Ser101               | 0.50251        | Short-range contacts       |

| <b>L33S <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                      |                |                            |
|--|----------------------|----------------|----------------------------|
| <b>Residue<br/>1</b>                                     | <b>Residue<br/>2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Pro29  | Leu91                | 0.50523        | long-range contacts        |
| Pro29  | Pro92                | 0.5193         | long-range contacts        |
| Thr32  | Asn61                | 0.50935        | long-range contacts        |
| Tyr34  | Val111               | 0.50224        | long-range contacts        |
| Ser38  | Ser66                | 0.50164        | long-range contacts        |
| Ser38  | Lys114               | 0.50846        | long-range contacts        |
| Leu40  | Gln80                | 0.51389        | long-range contacts        |
| Leu40  | Val82                | 0.52149        | long-range contacts        |
| Leu40  | Leu109               | 0.50124        | long-range contacts        |
| Leu40  | Thr115               | 0.5089         | long-range contacts        |
| Val41  | Arg69                | 0.50061        | long-range contacts        |
| Val41  | Val82                | 0.50521        | long-range contacts        |
| Val41  | Ser101               | 0.50041        | long-range contacts        |
| Leu42  | Arg81                | 0.562          | long-range contacts        |
| Glu43  | Gln80                | 0.5004         | long-range contacts        |
| Glu43  | Asn120               | 0.50061        | long-range contacts        |
| Lys52  | Thr115               | 0.50195        | long-range contacts        |
| Lys55  | Pro92                | 0.50318        | long-range contacts        |
| Val56  | Tyr86                | 0.51178        | long-range contacts        |
| Asn61  | Val85                | 0.50146        | long-range contacts        |
| Thr64  | Thr110               | 0.50258        | long-range contacts        |
| Phe26  | Ala47                | 0.50838        | Medium-range contacts      |
| Tyr28  | Ala44                | 0.50764        | Medium-range contacts      |

“Table I. 5 (cont’d)”

| <b>L33S <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                      |                |                            |
|--|----------------------|----------------|----------------------------|
| <b>Residue<br/>1</b>                                     | <b>Residue<br/>2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Tyr28  | Met46                | 0.51574        | Medium-range contacts      |
| Gln35  | Ser57                | 0.50799        | Medium-range contacts      |
| Glu43  | Asn59                | 0.50144        | Medium-range contacts      |
| Asn61  | Val82                | 0.52732        | Medium-range contacts      |
| Asn61  | Lys84                | 0.53278        | Medium-range contacts      |
| Val62  | Asp79                | 0.50624        | Medium-range contacts      |
| Thr64  | Val82                | 0.55511        | Medium-range contacts      |
| Ile65  | Val82                | 0.56396        | Medium-range contacts      |
| Ala67  | Val82                | 0.50451        | Medium-range contacts      |
| Arg69  | Arg81                | 0.50044        | Medium-range contacts      |
| Val76  | Leu91                | 0.50479        | Medium-range contacts      |
| Tyr86  | Asp99                | 0.50782        | Medium-range contacts      |
| Tyr86  | Ile108               | 0.50538        | Medium-range contacts      |
| Ile108   | Val121               | 0.50329        | Medium-range contacts      |
| Met1   | Phe7                 | 0.51212        | Short-range contacts       |
| Ile5   | Met12                | 0.5648         | Short-range contacts       |
| Lys6   | Met12                | 0.57287        | Short-range contacts       |
| Lys6   | Ile13                | 0.55274        | Short-range contacts       |
| Lys70  | Asp79                | 0.52687        | Short-range contacts       |
| Glu72  | Ile78                | 0.52285        | Short-range contacts       |
| Gln80  | Val88                | 0.54176        | Short-range contacts       |
| Lys90  | Asp99                | 0.50663        | Short-range contacts       |

**Table I. 6. List of residues forming contacts with the residue L33 in WT *Tpv* sHSP14.3 and Ser33 in L33S *Tpv* sHSP14.3 mutant. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| <b>WT <i>Tpv</i> sHSP14.3 Contacts at Position 33</b> |                  |                |                           |
|---|------------------|----------------|---------------------------|
| <b>Residue 1</b>                                      | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| Leu33   | Lys90            | 0.50829        | long-range contacts       |
| Leu33   | Glu122           | 0.51057        | long-range contacts       |
| Leu33   | Ala67            | 0.51851        | long-range contacts       |
| Leu33   | Val121           | 0.52608        | long-range contacts       |
| Leu33   | Ser66            | 0.53156        | long-range contacts       |
| Leu33   | Glu68            | 0.53764        | long-range contacts       |
| Leu33   | Arg112           | 0.54309        | long-range contacts       |
| Leu33   | Val89            | 0.55257        | long-range contacts       |
| Leu33   | Met113           | 0.57316        | long-range contacts       |
| Leu33   | Lys114           | 0.57329        | long-range contacts       |
| Leu33   | Lys87            | 0.58567        | long-range contacts       |
| Leu33   | Leu109           | 0.58911        | long-range contacts       |
| Leu33   | Val62            | 0.59187        | long-range contacts       |
| Leu33   | Val111           | 0.59394        | long-range contacts       |
| Leu33   | Val88            | 0.60011        | long-range contacts       |
| Leu33   | Thr64            | 0.60119        | long-range contacts       |
| Leu33   | Leu63            | 0.65279        | long-range contacts       |
| Leu33   | Ile65            | 0.66173        | long-range contacts       |
| Leu33   | Glu45            | 0.51847        | Medium-range contacts     |
| Leu33   | Asp39            | 0.54576        | Short-range contacts      |
| Leu33   | Ala44            | 0.70129        | Short-range contacts      |
| Leu33   | Glu43            | 0.76174        | Short-range contacts      |
| Leu33   | Leu40            | 0.81852        | Short-range contacts      |
| Leu33   | Val41            | 0.83408        | Short-range contacts      |
| Leu33   | Leu42            | 0.88898        | Short-range contacts      |

| <b>L33S <i>Tpv</i> sHSP14.3 Mutant Contacts at Position 33</b> |                  |                |                           |
|--|------------------|----------------|---------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| Ser33  | Val62            | 0.61783        | long-range contacts       |
| Ser33  | Leu63            | 0.60243        | long-range contacts       |
| Ser33  | Thr64            | 0.62654        | long-range contacts       |
| Ser33  | Ile65            | 0.63392        | long-range contacts       |
| Ser33  | Ser66            | 0.5352         | long-range contacts       |
| Ser33  | Ala67            | 0.52142        | long-range contacts       |
| Ser33  | Glu68            | 0.54088        | long-range contacts       |
| Ser33  | Lys87            | 0.57881        | long-range contacts       |
| Ser33  | Val88            | 0.60079        | long-range contacts       |
| Ser33  | Val89            | 0.57216        | long-range contacts       |
| Ser33  | Lys90            | 0.50906        | long-range contacts       |
| Ser33  | Leu109           | 0.61923        | long-range contacts       |
| Ser33  | Val111           | 0.59582        | long-range contacts       |
| Ser33  | Arg112           | 0.54337        | long-range contacts       |
| Ser33  | Met113           | 0.58804        | long-range contacts       |
| Ser33  | Lys114           | 0.57208        | long-range contacts       |
| Ser33  | Val121           | 0.5536         | long-range contacts       |
| Ser33  | Glu122           | 0.51009        | long-range contacts       |
| Ser33  | Glu45            | 0.53642        | Medium-range contacts     |
| Ser33  | Asp39            | 0.58504        | Short-range contacts      |

“Table I. 6 (cont’d)”

| <b>L33S <i>Tpv</i> sHSP14.3 Mutant Contacts at Position 33</b> |                  |                |                           |
|--|------------------|----------------|---------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| Ser33  | Leu40            | 0.79949        | Short-range contacts      |
| Ser33  | Val41            | 0.84885        | Short-range contacts      |
| Ser33  | Leu42            | 0.89139        | Short-range contacts      |
| Ser33  | Glu43            | 0.77433        | Short-range contacts      |
| Ser33  | Ala44            | 0.69635        | Short-range contacts      |

**Table I. 7. List of the specific contacts which are present in the WT *Tpv* sHSP14.3, but not for the G48E *Tpv* sHSP14.3 mutant and vice versa. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| <b>WT <i>Tpv</i> sHSP14.3 Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>                                | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Ile27   | Ala67            | 0.5001         | long-range contacts        |
| Tyr28   | Ile123           | 0.51044        | long-range contacts        |
| Pro29   | Ile123           | 0.50904        | long-range contacts        |
| Val31   | Ile123           | 0.50164        | long-range contacts        |
| Thr32   | Glu122           | 0.5108         | long-range contacts        |
| Thr32   | Ile123           | 0.51263        | long-range contacts        |
| Leu33   | Glu122           | 0.51057        | long-range contacts        |
| Leu33   | Val121           | 0.52608        | long-range contacts        |
| Tyr34   | Val121           | 0.5154         | long-range contacts        |
| Ser38   | Thr115           | 0.50279        | long-range contacts        |
| Ser38   | Ile123           | 0.53874        | long-range contacts        |
| Val41   | Glu124           | 0.52839        | long-range contacts        |
| Leu42   | Ile123           | 0.51384        | long-range contacts        |
| Ala47   | Ala102           | 0.50309        | long-range contacts        |
| Ala47   | Glu106           | 0.52557        | long-range contacts        |
| Phe49   | Leu109           | 0.508          | long-range contacts        |
| Lys51   | Met113           | 0.50842        | long-range contacts        |
| Ile54   | Ser105           | 0.50174        | long-range contacts        |
| Ile54   | Tyr104           | 0.51169        | long-range contacts        |
| Ile54   | Ile108           | 0.51304        | long-range contacts        |
| Ile54   | Leu109           | 0.51623        | long-range contacts        |
| Val56   | Gly107           | 0.51962        | long-range contacts        |
| Ser57   | Ser101           | 0.50227        | long-range contacts        |

**“Table I.7 (cont’d)”**

| <b>WT <i>Tpy</i> sHSP14.3 Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>                                | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Ser57   | Lys103           | 0.50486        | long-range contacts        |
| Val58   | Glu106           | 0.50057        | long-range contacts        |
| Val58   | Ile100           | 0.53535        | long-range contacts        |
| Lys60   | Ser105           | 0.50049        | long-range contacts        |
| Lys60   | Ala102           | 0.50136        | long-range contacts        |
| Lys60   | Val89            | 0.54637        | long-range contacts        |
| Asn61   | Pro92            | 0.54997        | long-range contacts        |
| Leu63   | Ile123           | 0.5108         | long-range contacts        |
| Ile65   | Ile123           | 0.50447        | long-range contacts        |
| Ile65   | Asp99            | 0.51814        | long-range contacts        |
| Ala67   | Gln98            | 0.50843        | long-range contacts        |
| Leu42   | Asn61            | 0.51032        | Medium-range contacts      |
| Asn61   | Gln80            | 0.51041        | Medium-range contacts      |
| Ala67   | Tyr86            | 0.54075        | Medium-range contacts      |
| Ala67   | Val85            | 0.55527        | Medium-range contacts      |
| Glu68   | Val85            | 0.56551        | Medium-range contacts      |
| Thr75   | Leu91            | 0.50275        | Medium-range contacts      |
| Thr75   | Lys87            | 0.51489        | Medium-range contacts      |
| Ala102  | Lys114           | 0.50025        | Medium-range contacts      |
| Met12   | Val20            | 0.50916        | Short-range contacts       |
| Ile13   | Val20            | 0.62073        | Short-range contacts       |
| Arg69   | Gln80            | 0.50341        | Short-range contacts       |
| Arg69   | Asp79            | 0.58868        | Short-range contacts       |
| Arg71   | Ile78            | 0.5772         | Short-range contacts       |
| Tyr77   | Tyr86            | 0.50164        | Short-range contacts       |

“Table I.7 (cont’d)”

| <b>G48E <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|--|------------------|----------------|----------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Pro29  | Leu91            | 0.50504        | long-range contacts        |
| Pro29  | Pro92            | 0.51721        | long-range contacts        |
| Thr32  | Asn61            | 0.50717        | long-range contacts        |
| Tyr34  | Val111           | 0.50297        | long-range contacts        |
| Ser38  | Ser66            | 0.5027         | long-range contacts        |
| Leu40  | Gln80            | 0.51156        | long-range contacts        |
| Leu40  | Val82            | 0.51797        | long-range contacts        |
| Leu40  | Leu109           | 0.53802        | long-range contacts        |
| Val41  | Val82            | 0.50258        | long-range contacts        |
| Leu42  | Arg81            | 0.56179        | long-range contacts        |
| Glu43  | Tyr104           | 0.5121         | long-range contacts        |
| Lys52  | Thr115           | 0.50053        | long-range contacts        |
| Asn53  | Val93            | 0.50696        | long-range contacts        |
| Lys55  | Pro92            | 0.50831        | long-range contacts        |
| Val56  | Tyr86            | 0.51169        | long-range contacts        |
| Asn61  | Val85            | 0.5072         | long-range contacts        |
| Thr64  | Thr110           | 0.50296        | long-range contacts        |
| Phe26  | Ala47            | 0.50905        | Medium-range contacts      |
| Tyr28  | Ala44            | 0.50509        | Medium-range contacts      |
| Tyr28  | Met46            | 0.51566        | Medium-range contacts      |
| Gln35  | Ser57            | 0.50754        | Medium-range contacts      |
| Glu43  | Asn59            | 0.50308        | Medium-range contacts      |
| Asn61  | Val82            | 0.52764        | Medium-range contacts      |
| Asn61  | Lys84            | 0.532          | Medium-range contacts      |
| Val62  | Asp79            | 0.52762        | Medium-range contacts      |
| Thr64  | Val82            | 0.54773        | Medium-range contacts      |
| Ile65  | Val82            | 0.5499         | Medium-range contacts      |
| Ser74  | Lys90            | 0.52134        | Medium-range contacts      |
| Ser74  | Pro92            | 0.50304        | Medium-range contacts      |
| Tyr86  | Asp99            | 0.50725        | Medium-range contacts      |
| Tyr86  | Ile108           | 0.50397        | Medium-range contacts      |
| Ile108   | Val121           | 0.50248        | Medium-range contacts      |
| Met1   | Phe7             | 0.50375        | Short-range contacts       |
| Ala67  | Thr75            | 0.53228        | Short-range contacts       |
| Glu68  | Ser74            | 0.53389        | Short-range contacts       |
| Gln80  | Val88            | 0.54062        | Short-range contacts       |
| Lys90  | Asp99            | 0.5069         | Short-range contacts       |

**Table I. 8. List of residues forming contacts with the residue Gly48 in WT *Tpv* sHSP14.3 and Glu48 in L33S *Tpv* sHSP14.3 mutant. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| <b>WT <i>Tpv</i> sHSP14.3 Contacts at Position 48</b> |                  |                |                           |
|---|------------------|----------------|---------------------------|
| <b>Residue 1</b>                                      | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact Evaluation</b> |
| Gly48   | Gly107           | 0.67628        | long-range contacts       |
| Gly48   | Ile108           | 0.53602        | long-range contacts       |
| Gly48   | Leu109           | 0.5333         | long-range contacts       |
| Gly48   | Ile54            | 0.66209        | Short-range contacts      |

| <b>G48E <i>Tpv</i> sHSP14.3 Mutant Contacts at Position 48</b> |                  |                |                           |
|--|------------------|----------------|---------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| Glu48  | Gly107           | 0.6517         | long-range contacts       |
| Glu48  | Ile108           | 0.51746        | long-range contacts       |
| Glu48  | Leu109           | 0.51841        | long-range contacts       |
| Glu48  | Ile54            | 0.67025        | Short-range contacts      |

**Table I. 9. List of the specific contacts which are present in the WT *Tpv* sHSP14.3, but not for the I108K *Tpv* sHSP14.3 mutant and vice versa. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| <b>WT <i>Tpv</i> sHSP14.3 Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>                                | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Ile27   | Ala67            | 0.5001         | long-range contacts        |
| Val31   | Ile123           | 0.50164        | long-range contacts        |
| Thr32   | Ile123           | 0.51263        | long-range contacts        |
| Asp39   | Glu124           | 0.55505        | long-range contacts        |
| Asp39   | Ile123           | 0.55527        | long-range contacts        |
| Val41   | Glu124           | 0.52839        | long-range contacts        |
| Leu42   | Ile123           | 0.51384        | long-range contacts        |
| Ala47   | Ala102           | 0.50309        | long-range contacts        |
| Phe49   | Leu109           | 0.508          | long-range contacts        |
| Lys51   | Met113           | 0.50842        | long-range contacts        |
| Ile54   | Leu109           | 0.51623        | long-range contacts        |
| Ile54   | Ser105           | 0.50174        | long-range contacts        |
| Ile54   | Tyr104           | 0.51169        | long-range contacts        |
| Lys55   | Ile108           | 0.5146         | long-range contacts        |
| Val56   | Gly107           | 0.51962        | long-range contacts        |
| Ser57   | Lys103           | 0.50486        | long-range contacts        |
| Ser57   | Ser101           | 0.50227        | long-range contacts        |
| Val58   | Glu106           | 0.50057        | long-range contacts        |
| Val58   | Ile100           | 0.53535        | long-range contacts        |
| Lys60   | Ala102           | 0.50136        | long-range contacts        |
| Lys60   | Ser105           | 0.50049        | long-range contacts        |

“Table I.9 (cont’d)”

| <b>WT <i>Tpv</i> sHSP14.3 Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>                                | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Leu63   | Ile123           | 0.5108         | long-range contacts        |
| Ile65   | Asp99            | 0.51814        | long-range contacts        |
| Ile65   | Ile123           | 0.50447        | long-range contacts        |
| Ala67   | Gln98            | 0.50843        | long-range contacts        |
| Leu42   | Asn61            | 0.51032        | Medium-range contacts      |
| Asn61   | Gln80            | 0.51041        | Medium-range contacts      |
| Ile65   | Tyr77            | 0.5043         | Medium-range contacts      |
| Ser66   | Val88            | 0.50573        | Medium-range contacts      |
| Ser66   | Val89            | 0.50728        | Medium-range contacts      |
| Thr75   | Lys87            | 0.51489        | Medium-range contacts      |
| Ala102  | Lys114           | 0.50025        | Medium-range contacts      |
| Arg112  | Glu124           | 0.50503        | Medium-range contacts      |
| Met12   | Val20            | 0.50916        | Short-range contacts       |
| Ile13   | Val20            | 0.62073        | Short-range contacts       |
| Tyr77   | Tyr86            | 0.50164        | Short-range contacts       |
| Ile78   | Val85            | 0.53982        | Short-range contacts       |

| <b>I108K <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>  | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Pro29   | Leu91            | 0.50444        | long-range contacts        |
| Pro29   | Pro92            | 0.51712        | long-range contacts        |
| Thr32   | Asn61            | 0.50958        | long-range contacts        |
| Leu33   | Lys108           | 0.50729        | long-range contacts        |
| Tyr34   | Val111           | 0.50366        | long-range contacts        |
| Gln35   | Thr64            | 0.50294        | long-range contacts        |
| Ser38   | Ser66            | 0.50328        | long-range contacts        |
| Ser38   | Lys114           | 0.50692        | long-range contacts        |
| Leu40   | Gln80            | 0.51337        | long-range contacts        |
| Leu40   | Val82            | 0.51926        | long-range contacts        |
| Leu40   | Leu109           | 0.50119        | long-range contacts        |
| Leu40   | Thr115           | 0.50601        | long-range contacts        |
| Val41   | Val82            | 0.50428        | long-range contacts        |
| Val41   | Ser101           | 0.50053        | long-range contacts        |
| Leu42   | Arg81            | 0.56329        | long-range contacts        |
| Glu43   | Gln80            | 0.50189        | long-range contacts        |
| Glu43   | Asn120           | 0.50107        | long-range contacts        |
| Lys52   | Thr115           | 0.50153        | long-range contacts        |
| Lys55   | Pro92            | 0.50475        | long-range contacts        |
| Val56   | Tyr86            | 0.51119        | long-range contacts        |
| Lys60   | Val88            | 0.50206        | long-range contacts        |
| Asn61   | Val85            | 0.50153        | long-range contacts        |
| Thr64   | Thr110           | 0.50305        | long-range contacts        |
| Phe26   | Ala47            | 0.50679        | Medium-range contacts      |
| Tyr28   | Ala44            | 0.50222        | Medium-range contacts      |
| Tyr28   | Met46            | 0.51296        | Medium-range contacts      |
| Gln35   | Ser57            | 0.51068        | Medium-range contacts      |
| Glu43   | Asn59            | 0.50594        | Medium-range contacts      |
| Lys52   | Arg71            | 0.50038        | Medium-range contacts      |
| Asn61   | Val82            | 0.52731        | Medium-range contacts      |



“Table I.9 (cont’d)”

| <b>I108K <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>  | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Asn61   | Lys84            | 0.5307         | Medium-range contacts      |
| Val62   | Asp79            | 0.50851        | Medium-range contacts      |
| Thr64   | Val82            | 0.55535        | Medium-range contacts      |
| Ile65   | Val82            | 0.56623        | Medium-range contacts      |
| Ala67   | Val82            | 0.50595        | Medium-range contacts      |
| Arg69   | Arg81            | 0.50261        | Medium-range contacts      |
| Val76   | Leu91            | 0.50551        | Medium-range contacts      |
| Tyr86   | Asp99            | 0.50783        | Medium-range contacts      |
| Tyr86   | Lys108           | 0.5027         | Medium-range contacts      |
| Lys108  | Val121           | 0.53955        | Medium-range contacts      |
| Met1  | Phe7             | 0.5039         | Short-range contacts       |
| Lys70   | Asp79            | 0.52941        | Short-range contacts       |
| Glu72   | Ile78            | 0.52398        | Short-range contacts       |
| Gln80   | Val88            | 0.5412         | Short-range contacts       |
| Lys90   | Asp99            | 0.50748        | Short-range contacts       |

**Table I. 10. List of residues forming contacts with the residue I108 in WT *Tpv* sHSP14.3 and K108 in I108K *Tpv* sHSP14.3 mutant. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| <b>WT <i>Tpv</i> sHSP14.3 Contacts at Position 108</b> |                  |                |                           |
|--|------------------|----------------|---------------------------|
| <b>Residue 1</b>                                       | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| Lys55  | Ile108           | 0.5146         | long-range contacts       |
| Ile108   | Glu122           | 0.53133        | Medium-range contacts     |
| Ile108   | Ile123           | 0.68024        | Medium-range contacts     |
| Ile108   | Glu124           | 0.7122         | Medium-range contacts     |

| <b>I108K <i>Tpv</i> sHSP14.3 Specific Contacts at Position 108</b> |                  |                |                           |
|--|------------------|----------------|---------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| Lys108   | Val121           | 0.53955        | Medium-range contacts     |
| Lys108   | Glu122           | 0.59284        | Medium-range contacts     |
| Lys108   | Ile123           | 0.70546        | Medium-range contacts     |
| Lys108   | Glu124           | 0.74008        | Medium-range contacts     |

**Table I. 11. List of the specific contacts which are present in the WT *Tpv* sHSP14.3, but not for the G107D *Tpv* sHSP14.3 mutant and vice versa. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| <b>WT <i>Tpv</i> sHSP14.3 Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>                                | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Ile27   | Ala67            | 0.5001         | long-range contacts        |
| Tyr28   | Ile123           | 0.51044        | long-range contacts        |
| Pro29   | Ile123           | 0.50904        | long-range contacts        |
| Pro30   | Ile123           | 0.53963        | long-range contacts        |
| Val31   | Ile123           | 0.50164        | long-range contacts        |
| Thr32   | Glu122           | 0.5108         | long-range contacts        |
| Thr32   | Ile123           | 0.51263        | long-range contacts        |
| Leu33   | Glu122           | 0.51057        | long-range contacts        |
| Leu33   | Val121           | 0.52608        | long-range contacts        |
| Tyr34   | Val121           | 0.5154         | long-range contacts        |
| Gln35   | Gly107           | 0.5103         | long-range contacts        |
| Ser38   | Thr115           | 0.50279        | long-range contacts        |
| Ser38   | Met113           | 0.53874        | long-range contacts        |
| Asp39   | Glu124           | 0.55505        | long-range contacts        |
| Asp39   | Ile123           | 0.55527        | long-range contacts        |
| Leu40   | Glu122           | 0.53661        | long-range contacts        |
| Val41   | Gly107           | 0.50254        | long-range contacts        |
| Val41   | Glu124           | 0.52839        | long-range contacts        |
| Leu42   | Ile123           | 0.51384        | long-range contacts        |
| Leu42   | Asp79            | 0.5186         | long-range contacts        |
| Ala47   | Ala102           | 0.50309        | long-range contacts        |
| Ala47   | Glu106           | 0.52557        | long-range contacts        |
| Phe49   | Leu109           | 0.508          | long-range contacts        |
| Ile54   | Ser105           | 0.50174        | long-range contacts        |
| Ile54   | Tyr104           | 0.51169        | long-range contacts        |
| Ile54   | Ile108           | 0.51304        | long-range contacts        |
| Ile54   | Leu109           | 0.51623        | long-range contacts        |
| Val56   | Gly107           | 0.51962        | long-range contacts        |
| Ser57   | Lys103           | 0.50486        | long-range contacts        |
| Lys60   | Ser105           | 0.50049        | long-range contacts        |
| Lys60   | Ala102           | 0.50136        | long-range contacts        |
| Lys60   | Val89            | 0.54637        | long-range contacts        |
| Asn61   | Pro92            | 0.54997        | long-range contacts        |
| Leu63   | Ile123           | 0.5108         | long-range contacts        |
| Ile65   | Ile123           | 0.50447        | long-range contacts        |
| Ile65   | Lys90            | 0.51142        | long-range contacts        |
| Ala67   | Gln98            | 0.50843        | long-range contacts        |
| Leu42   | Asn61            | 0.51032        | Medium-range contacts      |
| Asn61   | Gln80            | 0.51041        | Medium-range contacts      |
| Ser66   | Val88            | 0.50573        | Medium-range contacts      |
| Ser66   | Val82            | 0.54109        | Medium-range contacts      |
| Ala67   | Tyr86            | 0.54075        | Medium-range contacts      |
| Ala67   | Val85            | 0.55527        | Medium-range contacts      |
| Glu68   | Gln80            | 0.54192        | Medium-range contacts      |
| Glu68   | Val85            | 0.56551        | Medium-range contacts      |
| Thr75   | Leu91            | 0.50275        | Medium-range contacts      |
| Thr75   | Pro92            | 0.51105        | Medium-range contacts      |
| Thr75   | Lys87            | 0.51489        | Medium-range contacts      |

“Table I.11 (cont’d)”

| <b>WT <i>Tpv</i> sHSP14.3 Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>                                | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Leu91   | Gly107           | 0.5258         | Medium-range contacts      |
| Ala102  | Lys114           | 0.50025        | Medium-range contacts      |
| Met12   | Val20            | 0.50916        | Short-range contacts       |
| Ile13   | Val20            | 0.62073        | Short-range contacts       |
| Arg69   | Gln80            | 0.50341        | Short-range contacts       |
| Arg69   | Asp79            | 0.58868        | Short-range contacts       |
| Arg71   | Ile78            | 0.5772         | Short-range contacts       |
| Val76   | Lys87            | 0.56952        | Short-range contacts       |
| Tyr77   | Tyr86            | 0.50164        | Short-range contacts       |

| <b>G107D <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>  | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Pro29   | Leu91            | 0.50265        | long-range contacts        |
| Pro29   | Pro92            | 0.51697        | long-range contacts        |
| Thr32   | Asn61            | 0.50988        | long-range contacts        |
| Tyr34   | Val111           | 0.50043        | long-range contacts        |
| Gln35   | Thr64            | 0.50242        | long-range contacts        |
| Ser38   | Ser66            | 0.5041         | long-range contacts        |
| Leu40   | Leu109           | 0.53591        | long-range contacts        |
| Leu42   | Arg81            | 0.55725        | long-range contacts        |
| Glu43   | Gln80            | 0.51717        | long-range contacts        |
| Glu43   | Tyr104           | 0.51102        | long-range contacts        |
| Glu43   | Val121           | 0.50231        | long-range contacts        |
| Lys52   | Thr115           | 0.50424        | long-range contacts        |
| Asn53   | Val93            | 0.50347        | long-range contacts        |
| Lys55   | Pro92            | 0.50655        | long-range contacts        |
| Val56   | Tyr86            | 0.50488        | long-range contacts        |
| Ser57   | Val85            | 0.51245        | long-range contacts        |
| Asn61   | Val85            | 0.50157        | long-range contacts        |
| Asp83   | Arg112           | 0.50012        | long-range contacts        |
| Phe26   | Ala47            | 0.50898        | Medium-range contacts      |
| Tyr28   | Ala44            | 0.50546        | Medium-range contacts      |
| Tyr28   | Met46            | 0.51578        | Medium-range contacts      |
| Gln35   | Ser57            | 0.51027        | Medium-range contacts      |
| Glu43   | Asn59            | 0.50505        | Medium-range contacts      |
| Lys60   | Arg81            | 0.50189        | Medium-range contacts      |
| Asn61   | Val82            | 0.52981        | Medium-range contacts      |
| Asn61   | Lys84            | 0.51269        | Medium-range contacts      |
| Val62   | Asp79            | 0.53887        | Medium-range contacts      |
| Thr64   | Tyr77            | 0.5088         | Medium-range contacts      |
| Thr64   | Val82            | 0.50913        | Medium-range contacts      |
| Ile65   | Val82            | 0.52166        | Medium-range contacts      |
| Ser74   | Lys90            | 0.55132        | Medium-range contacts      |
| Ser74   | Leu91            | 0.54693        | Medium-range contacts      |
| Tyr86   | Asp99            | 0.50814        | Medium-range contacts      |
| Tyr86   | Ile108           | 0.50466        | Medium-range contacts      |
| Ile108  | Val121           | 0.50248        | Medium-range contacts      |
| Met1  | Phe7             | 0.50347        | Short-range contacts       |
| Ala67   | Thr75            | 0.5329         | Short-range contacts       |
| Glu68   | Ser74            | 0.53404        | Short-range contacts       |
| Lys90   | Asp99            | 0.50648        | Short-range contacts       |

**Table I. 12. List of residues forming contacts with the residue G107 in WT *Tpv* sHSP14.3 and D107 in G107D *Tpv* sHSP14.3 mutant. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| WT <i>Tpv</i> sHSP14.3 Contacts at position 107 |           |         |                     |
|---|-----------|---------|---------------------|
| Residue 1                                       | Residue 2 | P-score | Contacts evaluation |
| Gly107  | Glu124    | 0.64104 | long-range contact  |
| Gly107  | Ile123    | 0.64729 | long-range contact  |

| G107D <i>Tpv</i> sHSP14.3 Mutant Contacts at position 107 |           |         |                     |
|---|-----------|---------|---------------------|
| Residue 1   | Residue 2 | P-score | Contacts evaluation |
| Asp107  | Glu124    | 0.60446 | long-range contact  |
| Asp107  | Ile123    | 0.61399 | long-range contact  |

**Table I. 13. List of the specific contacts which are present in the WT *Tpv* sHSP14.3, but not for the A47D *Tpv* sHSP14.3 mutant and vice versa. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| WT <i>Tpv</i> sHSP14.3 Specific Contacts |           |         |                     |
|--|-----------|---------|---------------------|
| Residue 1                                | Residue 2 | P-score | Contacts evaluation |
| Ile27                                    | Ala67     | 0.5001  | long-range contacts |
| Tyr28                                    | Ile123    | 0.51044 | long-range contacts |
| Pro29                                    | Ile123    | 0.50904 | long-range contacts |
| Pro30                                    | Ile123    | 0.53963 | long-range contacts |
| Val31                                    | Ile123    | 0.50164 | long-range contacts |
| Thr32                                    | Glu122    | 0.5108  | long-range contacts |
| Thr32                                    | Ile123    | 0.51263 | long-range contacts |
| Leu33                                    | Glu122    | 0.51057 | long-range contacts |
| Leu33                                    | Val121    | 0.52608 | long-range contacts |
| Tyr34                                    | Val121    | 0.5154  | long-range contacts |
| Ser38                                    | Thr115    | 0.50279 | long-range contacts |
| Ser38                                    | Met113    | 0.53874 | long-range contacts |
| Asp39                                    | Glu124    | 0.55505 | long-range contacts |
| Asp39                                    | Ile123    | 0.55527 | long-range contacts |
| Val41                                    | Glu124    | 0.52839 | long-range contacts |
| Leu42                                    | Ile123    | 0.51384 | long-range contacts |
| Ala47                                    | Ala102    | 0.50309 | long-range contacts |
| Ala47                                    | Pro92     | 0.50486 | long-range contacts |
| Ala47                                    | Leu91     | 0.5216  | long-range contacts |
| Ala47                                    | Glu106    | 0.52557 | long-range contacts |
| Phe49                                    | Leu109    | 0.508   | long-range contacts |
| Lys51                                    | Met113    | 0.50842 | long-range contacts |
| Ile54                                    | Ser105    | 0.50174 | long-range contacts |
| Ile54                                    | Tyr104    | 0.51169 | long-range contacts |
| Ile54                                    | Ile108    | 0.51304 | long-range contacts |
| Ile54                                    | Leu109    | 0.51623 | long-range contacts |
| Val56                                    | Gly107    | 0.51962 | long-range contacts |
| Ser57                                    | Ser101    | 0.50227 | long-range contacts |
| Ser57                                    | Lys103    | 0.50486 | long-range contacts |
| Val58                                    | Glu106    | 0.50057 | long-range contacts |
| Val58                                    | Ile100    | 0.53535 | long-range contacts |

“Table I.13 (cont’d)”

| <b>WT <i>Tpv</i> sHSP14.3 Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>                                | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Lys60   | Ser105           | 0.50049        | long-range contacts        |
| Lys60   | Ala102           | 0.50136        | long-range contacts        |
| Lys60   | Val89            | 0.54637        | long-range contacts        |
| Asn61   | Pro92            | 0.54997        | long-range contacts        |
| Val62   | Ile100           | 0.53919        | long-range contacts        |
| Leu63   | Ile123           | 0.5108         | long-range contacts        |
| Ile65   | Ile123           | 0.50447        | long-range contacts        |
| Ile65   | Asp99            | 0.51814        | long-range contacts        |
| Ala67   | Gln98            | 0.50843        | long-range contacts        |
| Thr32   | Ala47            | 0.51           | Medium-range contacts      |
| Leu42   | Asn61            | 0.51032        | Medium-range contacts      |
| Asn61   | Gln80            | 0.51041        | Medium-range contacts      |
| Ala67   | Tyr86            | 0.54075        | Medium-range contacts      |
| Ala67   | Val85            | 0.55527        | Medium-range contacts      |
| Glu68   | Val85            | 0.56551        | Medium-range contacts      |
| Thr75   | Leu91            | 0.50275        | Medium-range contacts      |
| Thr75   | Lys87            | 0.51489        | Medium-range contacts      |
| Ala102  | Lys114           | 0.50025        | Medium-range contacts      |
| Arg112  | Glu124           | 0.50503        | Medium-range contacts      |
| Met12   | Val20            | 0.50916        | Short-range contacts       |
| Ile13   | Val20            | 0.62073        | Short-range contacts       |
| Arg69   | Gln80            | 0.50341        | Short-range contacts       |
| Arg69   | Asp79            | 0.58868        | Short-range contacts       |
| Arg71   | Ile78            | 0.5772         | Short-range contacts       |
| Tyr77   | Tyr86            | 0.50164        | Short-range contacts       |
| Pro92   | Ser101           | 0.50251        | Short-range contacts       |

| <b>A47D <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|--|------------------|----------------|----------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Pro29  | Leu91            | 0.50436        | long-range contacts        |
| Pro29  | Pro92            | 0.51729        | long-range contacts        |
| Thr32  | Asn61            | 0.50858        | long-range contacts        |
| Tyr34  | Val111           | 0.50343        | long-range contacts        |
| Gln35  | Thr64            | 0.50319        | long-range contacts        |
| Ser38  | Ser66            | 0.50141        | long-range contacts        |
| Leu40  | Gln80            | 0.51356        | long-range contacts        |
| Leu40  | Val82            | 0.51971        | long-range contacts        |
| Leu40  | Leu109           | 0.53889        | long-range contacts        |
| Val41  | Val82            | 0.5067         | long-range contacts        |
| Val41  | Ser101           | 0.50014        | long-range contacts        |
| Leu42  | Arg81            | 0.56367        | long-range contacts        |
| Glu43  | Gln80            | 0.50365        | long-range contacts        |
| Glu43  | Tyr104           | 0.51439        | long-range contacts        |
| Lys52  | Thr115           | 0.50015        | long-range contacts        |
| Asn53  | Val93            | 0.50755        | long-range contacts        |
| Lys55  | Pro92            | 0.50956        | long-range contacts        |
| Val56  | Tyr86            | 0.51033        | long-range contacts        |
| Asn61  | Val85            | 0.5076         | long-range contacts        |
| Thr64  | Thr110           | 0.50252        | long-range contacts        |
| Tyr28  | Ala44            | 0.50321        | Medium-range contacts      |

“Table I.13 (cont’d)”

| A47D <i>Tpv</i> sHSP14.3 Mutant Specific Contacts |           |         |                       |
|---|-----------|---------|-----------------------|
| Residue 1   | Residue 2 | P-score | Contacts evaluation   |
| Tyr28   | Met46     | 0.5123  | Medium-range contacts |
| Gln35   | Ser57     | 0.51068 | Medium-range contacts |
| Glu43   | Asn59     | 0.50577 | Medium-range contacts |
| Asn61   | Val82     | 0.52742 | Medium-range contacts |
| Asn61   | Lys84     | 0.53153 | Medium-range contacts |
| Val62   | Asp79     | 0.52849 | Medium-range contacts |
| Thr64   | Val82     | 0.54837 | Medium-range contacts |
| Ile65   | Val82     | 0.54797 | Medium-range contacts |
| Ser74   | Lys90     | 0.51947 | Medium-range contacts |
| Ser74   | Pro92     | 0.50263 | Medium-range contacts |
| Tyr86   | Asp99     | 0.50602 | Medium-range contacts |
| Tyr86   | Ile108    | 0.50382 | Medium-range contacts |
| Ile108  | Val121    | 0.50308 | Medium-range contacts |
| Met1  | Phe7      | 0.50348 | Short-range contacts  |
| Asp47   | Lys55     | 0.50639 | Short-range contacts  |
| Ala67   | Thr75     | 0.52873 | Short-range contacts  |
| Glu68   | Ser74     | 0.53092 | Short-range contacts  |
| Gln80   | Val88     | 0.54236 | Short-range contacts  |
| Lys90   | Asp99     | 0.50709 | Short-range contacts  |

**Table I. 14. List of residues forming contacts with the residue A47 in WT *Tpv* sHSP14.3 and D47 in A47D *Tpv* sHSP14.3 Mutant. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| WT <i>Tpv</i> sHSP14.3 Contacts at Position 47 |           |         |                     |
|--|-----------|---------|---------------------|
| Residue 1                                      | Residue 2 | P-score | Contacts evaluation |
| Ala47  | Val56     | 0.51622 | long-range contacts |
| Ala47  | Ile100    | 0.53501 | long-range contacts |
| Ala47  | Leu109    | 0.66634 | long-range contacts |
| Ala47  | Ile108    | 0.68044 | long-range contacts |
| Ala47  | Asn53     | 0.70492 | long-range contacts |
| Ala47  | Gly107    | 0.74787 | long-range contacts |
| Ala47  | Ile54     | 0.76054 | long-range contacts |
| Ala47  | Ala102    | 0.50309 | long-range contacts |
| Ala47  | Pro92     | 0.50486 | long-range contacts |
| Ala47  | Leu91     | 0.5216  | long-range contacts |
| Ala47  | Glu106    | 0.52557 | long-range contacts |

| A47D <i>Tpv</i> sHSP14.3 Mutant Contacts at Position 47 |           |         |                      |
|---|-----------|---------|----------------------|
| Residue 1   | Residue 2 | P-score | Contacts evaluation  |
| Asp47   | Ile100    | 0.53476 | long-range contacts  |
| Asp47   | Gly107    | 0.72999 | long-range contacts  |
| Asp47   | Ile108    | 0.67179 | long-range contacts  |
| Asp47   | Leu109    | 0.64868 | long-range contacts  |
| Asp47   | Asn53     | 0.71386 | Short-range contacts |
| Asp47   | Ile54     | 0.74981 | Short-range contacts |
| Asp47   | Lys55     | 0.50639 | Short-range contacts |
| Asp47   | Val56     | 0.51413 | Short-range contacts |

**Table I. 15. List of the specific contacts which are present in the WT *Tpv* sHSP14.3, but not for the E45G *Tpv* sHSP14.3 mutant and vice versa. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| WT <i>Tpv</i> sHSP14.3 Specific Contacts |           |         |                       |
|--|-----------|---------|-----------------------|
| Residue 1                                | Residue 2 | P-score | Contacts evaluation   |
| Ile27                                    | Ala67     | 0.5001  | long-range contacts   |
| Tyr28                                    | Ile123    | 0.51044 | long-range contacts   |
| Pro29                                    | Ile123    | 0.50904 | long-range contacts   |
| Pro30                                    | Ile123    | 0.53963 | long-range contacts   |
| Val31                                    | Ile123    | 0.50164 | long-range contacts   |
| Thr32                                    | Glu122    | 0.5108  | long-range contacts   |
| Thr32                                    | Ile123    | 0.51263 | long-range contacts   |
| Leu33                                    | Glu122    | 0.51057 | long-range contacts   |
| Leu33                                    | Val121    | 0.52608 | long-range contacts   |
| Tyr34                                    | Val121    | 0.5154  | long-range contacts   |
| Ser38                                    | Thr115    | 0.50279 | long-range contacts   |
| Ser38                                    | Met113    | 0.53874 | long-range contacts   |
| Val41                                    | Glu124    | 0.52839 | long-range contacts   |
| Leu42                                    | Ile123    | 0.51384 | long-range contacts   |
| Ala47                                    | Ala102    | 0.50309 | long-range contacts   |
| Ala47                                    | Glu106    | 0.52557 | long-range contacts   |
| Phe49                                    | Leu109    | 0.508   | long-range contacts   |
| Lys51                                    | Met113    | 0.50842 | long-range contacts   |
| Ile54                                    | Ser105    | 0.50174 | long-range contacts   |
| Ile54                                    | Tyr104    | 0.51169 | long-range contacts   |
| Ile54                                    | Ile108    | 0.51304 | long-range contacts   |
| Ile54                                    | Leu109    | 0.51623 | long-range contacts   |
| Val56                                    | Gly107    | 0.51962 | long-range contacts   |
| Ser57                                    | Ser101    | 0.50227 | long-range contacts   |
| Ser57                                    | Lys103    | 0.50486 | long-range contacts   |
| Val58                                    | Glu106    | 0.50057 | long-range contacts   |
| Val58                                    | Ile100    | 0.53535 | long-range contacts   |
| Lys60                                    | Ser105    | 0.50049 | long-range contacts   |
| Lys60                                    | Ala102    | 0.50136 | long-range contacts   |
| Asn61                                    | Pro92     | 0.54997 | long-range contacts   |
| Leu63                                    | Ile123    | 0.5108  | long-range contacts   |
| Ile65                                    | Ile123    | 0.50447 | long-range contacts   |
| Ile65                                    | Asp99     | 0.51814 | long-range contacts   |
| Ala67                                    | Gln98     | 0.50843 | long-range contacts   |
| Asn61                                    | Gln80     | 0.51041 | Medium-range contacts |
| Ala67                                    | Tyr86     | 0.54075 | Medium-range contacts |
| Ala67                                    | Val85     | 0.55527 | Medium-range contacts |
| Glu68                                    | Val85     | 0.56551 | Medium-range contacts |
| Thr75                                    | Leu91     | 0.50275 | Medium-range contacts |
| Thr75                                    | Pro92     | 0.51105 | Medium-range contacts |
| Thr75                                    | Lys87     | 0.51489 | Medium-range contacts |
| Ala102                                   | Lys114    | 0.50025 | Medium-range contacts |
| Met12                                    | Val20     | 0.50916 | Short-range contacts  |
| Ile13                                    | Val20     | 0.62073 | Short-range contacts  |
| Glu45                                    | Lys52     | 0.59449 | Short-range contacts  |
| Arg69                                    | Gln80     | 0.50341 | Short-range contacts  |
| Arg69                                    | Asp79     | 0.58868 | Short-range contacts  |
| Arg71                                    | Ile78     | 0.5772  | Short-range contacts  |
| Tyr77                                    | Tyr86     | 0.50164 | Short-range contacts  |

“Table I.15 (cont’d)”

| <b>E45G <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|--|------------------|----------------|----------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Pro29  | Leu91            | 0.50274        | long-range contacts        |
| Pro29  | Pro92            | 0.5143         | long-range contacts        |
| Thr32  | Asn61            | 0.50977        | long-range contacts        |
| Tyr34  | Val111           | 0.50449        | long-range contacts        |
| Gln35  | Thr64            | 0.50472        | long-range contacts        |
| Ser38  | Ser66            | 0.50528        | long-range contacts        |
| Leu40  | Gln80            | 0.51286        | long-range contacts        |
| Leu40  | Val82            | 0.51935        | long-range contacts        |
| Leu40  | Leu109           | 0.53763        | long-range contacts        |
| Val41  | Val82            | 0.50548        | long-range contacts        |
| Val41  | Ser101           | 0.50058        | long-range contacts        |
| Leu42  | Arg81            | 0.56357        | long-range contacts        |
| Glu43  | Gln80            | 0.50294        | long-range contacts        |
| Glu43  | Tyr104           | 0.51373        | long-range contacts        |
| Asn53  | Val93            | 0.50828        | long-range contacts        |
| Lys55  | Pro92            | 0.50874        | long-range contacts        |
| Val56  | Tyr86            | 0.51109        | long-range contacts        |
| Asn61  | Val85            | 0.50645        | long-range contacts        |
| Thr64  | Thr110           | 0.5019         | long-range contacts        |
| Phe26  | Ala47            | 0.50696        | Medium-range contacts      |
| Tyr28  | Ala44            | 0.50127        | Medium-range contacts      |
| Tyr28  | Met46            | 0.5123         | Medium-range contacts      |
| Gln35  | Ser57            | 0.51195        | Medium-range contacts      |
| Glu43  | Asn59            | 0.51051        | Medium-range contacts      |
| Asn61  | Val82            | 0.52951        | Medium-range contacts      |
| Asn61  | Lys84            | 0.5324         | Medium-range contacts      |
| Val62  | Asp79            | 0.52994        | Medium-range contacts      |
| Thr64  | Val82            | 0.54947        | Medium-range contacts      |
| Ile65  | Val82            | 0.54906        | Medium-range contacts      |
| Ser74  | Lys90            | 0.52073        | Medium-range contacts      |
| Ser74  | Pro92            | 0.50215        | Medium-range contacts      |
| Tyr86  | Asp99            | 0.5073         | Medium-range contacts      |
| Tyr86  | Ile108           | 0.50394        | Medium-range contacts      |
| Ile108   | Val121           | 0.50284        | Medium-range contacts      |
| Met1   | Phe7             | 0.50375        | Short-range contacts       |
| Ala67  | Thr75            | 0.53343        | Short-range contacts       |
| Glu68  | Ser74            | 0.53578        | Short-range contacts       |
| Gln80  | Val88            | 0.53924        | Short-range contacts       |
| Lys90  | Asp99            | 0.5067         | Short-range contacts       |



**Table I. 16. List of residues forming contacts with the residue E45 in WT *Tpv* sHSP14.3 and G45 in E45G *Tpv* sHSP14.3 mutant. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| <b>WT <i>Tpv</i> sHSP14.3 Contacts at Position 45</b> |                  |                |                           |
|---|------------------|----------------|---------------------------|
| <b>Residue 1</b>                                      | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| Glu45   | Ala102           | 0.51271        | long-range contacts       |
| Glu45   | Thr110           | 0.56208        | long-range contacts       |
| Glu45   | Gly107           | 0.65385        | long-range contacts       |
| Glu45   | Ile108           | 0.67842        | long-range contacts       |
| Glu45   | Leu109           | 0.69609        | long-range contacts       |
| Glu45   | Leu63            | 0.53312        | Medium-range contacts     |
| Glu45   | Thr64            | 0.58421        | Medium-range contacts     |
| Glu45   | Asn53            | 0.58865        | Short-range contacts      |
| Glu45   | Lys52            | 0.59449        | Short-range contacts      |
| Glu45   | Val56            | 0.65335        | Short-range contacts      |
| Glu45   | Lys55            | 0.65732        | Short-range contacts      |
| Glu45   | Lys51            | 0.68094        | Short-range contacts      |
| Glu45   | Ile54            | 0.71667        | Short-range contacts      |

| <b>E45G <i>Tpv</i> sHSP14.3 Mutant Contacts at Position 45</b> |                  |                |                           |
|--|------------------|----------------|---------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| Gly45  | Ala102           | 0.53836        | long-range contacts       |
| Gly45  | Gly107           | 0.67649        | long-range contacts       |
| Gly45  | Ile108           | 0.69964        | long-range contacts       |
| Gly45  | Leu109           | 0.73117        | long-range contacts       |
| Gly45  | Thr110           | 0.54441        | long-range contacts       |
| Gly45  | Leu63            | 0.56739        | Medium-range contacts     |
| Gly45  | Thr64            | 0.56361        | Medium-range contacts     |
| Gly45  | Lys51            | 0.59396        | Short-range contacts      |
| Gly45  | Asn53            | 0.59379        | Short-range contacts      |
| Gly45  | Ile54            | 0.74631        | Short-range contacts      |
| Gly45  | Lys55            | 0.58879        | Short-range contacts      |
| Gly45  | Val56            | 0.68869        | Short-range contacts      |

**Table I. 17. List of the specific contacts which are present in the WT *Tpv* sHSP14.3, but not for the E43V *Tpv* sHSP14.3 mutant and vice versa. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| WT <i>Tpv</i> sHSP14.3 Specific Contacts |           |         |                     |
|--|-----------|---------|---------------------|
| Residue 1                                | Residue 2 | P-score | Contacts evaluation |
| Tyr28                                    | Ile123    | 0.51044 | long-range contacts |
| Pro29                                    | Ile123    | 0.50904 | long-range contacts |
| Pro29                                    | Ala67     | 0.52662 | long-range contacts |
| Pro29                                    | Ser66     | 0.55308 | long-range contacts |
| Pro29                                    | Thr115    | 0.55447 | long-range contacts |
| Pro30                                    | Pro92     | 0.50072 | long-range contacts |
| Pro30                                    | Ser66     | 0.5037  | long-range contacts |
| Pro30                                    | Lys90     | 0.52096 | long-range contacts |
| Pro30                                    | Ala67     | 0.53841 | long-range contacts |
| Pro30                                    | Thr115    | 0.55958 | long-range contacts |
| Pro30                                    | Lys114    | 0.60024 | long-range contacts |
| Val31                                    | Ile123    | 0.50164 | long-range contacts |
| Val31                                    | Val88     | 0.51539 | long-range contacts |
| Val31                                    | Lys70     | 0.51951 | long-range contacts |
| Val31                                    | Thr64     | 0.51967 | long-range contacts |
| Val31                                    | Thr115    | 0.53056 | long-range contacts |
| Val31                                    | Arg112    | 0.55198 | long-range contacts |
| Val31                                    | Leu91     | 0.56154 | long-range contacts |
| Thr32                                    | Glu122    | 0.5108  | long-range contacts |
| Thr32                                    | Ile123    | 0.51263 | long-range contacts |
| Thr32                                    | Ser66     | 0.5243  | long-range contacts |
| Thr32                                    | Leu109    | 0.52675 | long-range contacts |
| Thr32                                    | Lys90     | 0.54477 | long-range contacts |
| Thr32                                    | Leu91     | 0.54602 | long-range contacts |
| Thr32                                    | Glu122    | 0.55415 | long-range contacts |
| Thr32                                    | Ile123    | 0.57476 | long-range contacts |
| Leu33                                    | Leu90     | 0.50829 | long-range contacts |
| Leu33                                    | Glu122    | 0.51057 | long-range contacts |
| Leu33                                    | Val121    | 0.52608 | long-range contacts |
| Tyr34                                    | Val89     | 0.5068  | long-range contacts |
| Tyr34                                    | Val58     | 0.51045 | long-range contacts |
| Tyr34                                    | Val121    | 0.5154  | long-range contacts |
| Tyr34                                    | Leu109    | 0.5287  | long-range contacts |
| Tyr34                                    | Leu63     | 0.53396 | long-range contacts |
| Gln35                                    | Val88     | 0.5099  | long-range contacts |
| Gln35                                    | Gly107    | 0.5103  | long-range contacts |
| Gln35                                    | Leu63     | 0.51646 | long-range contacts |
| Gln35                                    | Val62     | 0.53364 | long-range contacts |
| Gln35                                    | Lys87     | 0.5337  | long-range contacts |
| Gln35                                    | Leu109    | 0.5395  | long-range contacts |
| Leu40                                    | Ile108    | 0.50669 | long-range contacts |
| Val41                                    | Tyr86     | 0.50203 | long-range contacts |
| Val41                                    | Gly107    | 0.50254 | long-range contacts |
| Val41                                    | Glu124    | 0.52839 | long-range contacts |
| Leu42                                    | Ile123    | 0.51384 | long-range contacts |
| Ala47                                    | Ala102    | 0.50309 | long-range contacts |
| Lys51                                    | Lys103    | 0.50842 | long-range contacts |
| Ile54                                    | Ser105    | 0.50174 | long-range contacts |

“Table I.17 (cont’d)”

| WT <i>Tpv</i> sHSP14.3 Specific Contacts |           |         |                       |
|--|-----------|---------|-----------------------|
| Residue 1                                | Residue 2 | P-score | Contacts evaluation   |
| Ile54                                    | Tyr104    | 0.51169 | long-range contacts   |
| Ile54                                    | Leu109    | 0.51623 | long-range contacts   |
| Ile54                                    | Thr110    | 0.52215 | long-range contacts   |
| Lys55                                    | Thr110    | 0.51446 | long-range contacts   |
| Val56                                    | Lys114    | 0.51826 | long-range contacts   |
| Val56                                    | Gly107    | 0.51962 | long-range contacts   |
| Ser57                                    | Ser101    | 0.50227 | long-range contacts   |
| Ser57                                    | Lys103    | 0.50486 | long-range contacts   |
| Val58                                    | Glu106    | 0.50057 | long-range contacts   |
| Val58                                    | Ile100    | 0.53535 | long-range contacts   |
| Lys60                                    | Ser105    | 0.50049 | long-range contacts   |
| Lys60                                    | Ala102    | 0.50136 | long-range contacts   |
| Asn61                                    | Thr110    | 0.52452 | long-range contacts   |
| Val62                                    | Thr110    | 0.51521 | long-range contacts   |
| Val62                                    | Ile100    | 0.53919 | long-range contacts   |
| Ile65                                    | Asp99     | 0.51814 | long-range contacts   |
| Ser66                                    | Asp99     | 0.55428 | long-range contacts   |
| Ala67                                    | Gln98     | 0.50843 | long-range contacts   |
| Lys87                                    | Lys114    | 0.50455 | long-range contacts   |
| Tyr28                                    | Ala47     | 0.52942 | Medium-range contacts |
| Pro30                                    | Leu42     | 0.52328 | Medium-range contacts |
| Pro30                                    | Met46     | 0.59694 | Medium-range contacts |
| Pro30                                    | Ala47     | 0.59702 | Medium-range contacts |
| Val31                                    | Ala47     | 0.51406 | Medium-range contacts |
| Thr32                                    | Ala47     | 0.51    | Medium-range contacts |
| Thr32                                    | Met46     | 0.53653 | Medium-range contacts |
| Leu33                                    | Glu45     | 0.51847 | Medium-range contacts |
| Leu42                                    | Asn61     | 0.51032 | Medium-range contacts |
| Glu43                                    | Lys60     | 0.51996 | Medium-range contacts |
| Ser66                                    | Val89     | 0.50728 | Medium-range contacts |
| Glu68                                    | Gln80     | 0.54192 | Medium-range contacts |
| Glu68                                    | Val85     | 0.56551 | Medium-range contacts |
| Thr75                                    | Lys87     | 0.51489 | Medium-range contacts |
| Tyr86                                    | Ser105    | 0.50354 | Medium-range contacts |
| Lys87                                    | Gly107    | 0.50808 | Medium-range contacts |
| Ala102                                   | Lys114    | 0.50025 | Medium-range contacts |
| Arg112                                   | Glu124    | 0.50503 | Medium-range contacts |
| Met12                                    | Val20     | 0.50916 | Short-range contacts  |
| Thr32                                    | Leu40     | 0.61297 | Short-range contacts  |
| Leu33                                    | Asp39     | 0.54576 | Short-range contacts  |
| Tyr34                                    | Ala44     | 0.50735 | Short-range contacts  |
| Gln35                                    | Glu43     | 0.53681 | Short-range contacts  |
| Glu43                                    | Phe49     | 0.50552 | Short-range contacts  |
| Arg69                                    | Gln80     | 0.50341 | Short-range contacts  |
| Val76                                    | Lys87     | 0.56952 | Short-range contacts  |
| Tyr77                                    | Tyr86     | 0.50164 | Short-range contacts  |
| Ile78                                    | Val85     | 0.53982 | Short-range contacts  |
| Val93                                    | Ile100    | 0.50618 | Short-range contacts  |
| Lys114                                   | Ile123    | 0.54036 | Short-range contacts  |

“Table I.17 (cont’d)”

| E43V <i>Tpv</i> sHSP14.3 Mutant Specific Contacts |           |         |                       |
|---|-----------|---------|-----------------------|
| Residue 1   | Residue 2 | P-score | Contacts evaluation   |
| Ser38   | Lys114    | 0.53547 | long-range contacts   |
| Leu40   | Gln80     | 0.51138 | long-range contacts   |
| Leu40   | Val82     | 0.50955 | long-range contacts   |
| Leu40   | Lys87     | 0.50415 | long-range contacts   |
| Val41   | Val82     | 0.51083 | long-range contacts   |
| Val41   | Leu91     | 0.50099 | long-range contacts   |
| Leu42   | Arg81     | 0.55602 | long-range contacts   |
| Val43   | Asp79     | 0.50399 | long-range contacts   |
| Val43   | Tyr104    | 0.5155  | long-range contacts   |
| Ala44   | Ile100    | 0.51618 | long-range contacts   |
| Ala44   | Arg112    | 0.50337 | long-range contacts   |
| Glu45   | Val111    | 0.50981 | long-range contacts   |
| Val56   | Tyr86     | 0.55849 | long-range contacts   |
| Ser57   | Val85     | 0.55848 | long-range contacts   |
| Lys60   | Val85     | 0.51018 | long-range contacts   |
| Asn61   | Val85     | 0.50967 | long-range contacts   |
| Ala67   | Gly107    | 0.50232 | long-range contacts   |
| Gln35   | Ser57     | 0.5231  | Medium-range contacts |
| Val43   | Asn59     | 0.51312 | Medium-range contacts |
| Lys51   | Arg69     | 0.50485 | Medium-range contacts |
| Lys52   | Lys70     | 0.5494  | Medium-range contacts |
| Lys60   | Gln80     | 0.55146 | Medium-range contacts |
| Lys60   | Arg81     | 0.57271 | Medium-range contacts |
| Lys60   | Val82     | 0.55449 | Medium-range contacts |
| Lys60   | Asp83     | 0.52356 | Medium-range contacts |
| Asn61   | Asp79     | 0.51232 | Medium-range contacts |
| Asn61   | Val82     | 0.60391 | Medium-range contacts |
| Asn61   | Lys84     | 0.5437  | Medium-range contacts |
| Val62   | Ile78     | 0.53649 | Medium-range contacts |
| Val62   | Asp79     | 0.59862 | Medium-range contacts |
| Thr64   | Val82     | 0.53947 | Medium-range contacts |
| Ile65   | Val82     | 0.56241 | Medium-range contacts |
| Ser66   | Tyr86     | 0.54391 | Medium-range contacts |
| Ser74   | Lys90     | 0.50251 | Medium-range contacts |
| Ser74   | Pro92     | 0.50763 | Medium-range contacts |
| Lys84   | Ala102    | 0.5073  | Medium-range contacts |
| Val89   | Gly107    | 0.50514 | Medium-range contacts |
| Lys90   | Gly107    | 0.50441 | Medium-range contacts |
| Lys90   | Ile108    | 0.51253 | Medium-range contacts |
| Pro92   | Ile108    | 0.50357 | Medium-range contacts |
| Met1  | Phe7      | 0.51751 | Short-range contacts  |
| Ile5  | Met12     | 0.54567 | Short-range contacts  |
| Lys6  | Met12     | 0.53101 | Short-range contacts  |
| Lys6  | Ile13     | 0.51992 | Short-range contacts  |
| Ser24   | Val31     | 0.51887 | Short-range contacts  |
| Ile27   | Leu33     | 0.59418 | Short-range contacts  |
| Ser66   | Ser74     | 0.53269 | Short-range contacts  |
| Ser66   | Thr75     | 0.54666 | Short-range contacts  |
| Ala67   | Ser74     | 0.51589 | Short-range contacts  |
| Ala67   | Thr75     | 0.59418 | Short-range contacts  |
| Glu68   | Ser74     | 0.61845 | Short-range contacts  |
| Gln80   | Val88     | 0.55216 | Short-range contacts  |
| Lys90   | Asp99     | 0.50006 | Short-range contacts  |
| Pro92   | Asp99     | 0.50127 | Short-range contacts  |

**Table I. 18. List of residues forming contacts with the residue E43 in WT *Tpv* sHSP14.3 and V43 in E43V *Tpv* sHSP14.3 mutant.** Probability score is the probability of the two residues in contact within the 8 Å distance.

| <b>WT <i>Tpv</i> sHSP14.3 Specific Contacts at Position 43</b> |                  |                |                           |
|--|------------------|----------------|---------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| Glu43  | Met113           | 0.52659        | long-range contacts       |
| Glu43  | Lys114           | 0.53158        | long-range contacts       |
| Glu43  | Val88            | 0.53291        | long-range contacts       |
| Glu43  | Lys103           | 0.54939        | long-range contacts       |
| Glu43  | Gly107           | 0.56368        | long-range contacts       |
| Glu43  | Lys87            | 0.56456        | long-range contacts       |
| Glu43  | Ser101           | 0.5933         | long-range contacts       |
| Glu43  | Arg112           | 0.61792        | long-range contacts       |
| Glu43  | Ala102           | 0.64331        | long-range contacts       |
| Glu43  | Ile108           | 0.66926        | long-range contacts       |
| Glu43  | Val111           | 0.68235        | long-range contacts       |
| Glu43  | Thr110           | 0.71199        | long-range contacts       |
| Glu43  | Leu109           | 0.72799        | long-range contacts       |
| <b>Glu43</b>   | <b>Lys60</b>     | 0.51996        | Medium-range contacts     |
| Glu43  | Ser66            | 0.60563        | Medium-range contacts     |
| Glu43  | Leu63            | 0.62875        | Medium-range contacts     |
| Glu43  | Val62            | 0.64148        | Medium-range contacts     |
| Glu43  | Ile65            | 0.67923        | Medium-range contacts     |
| Glu43  | Val58            | 0.68334        | Medium-range contacts     |
| Glu43  | Thr64            | 0.7137         | Medium-range contacts     |
| Glu43  | Ser57            | 0.74022        | Medium-range contacts     |
| Glu43  | Lys55            | 0.78242        | Medium-range contacts     |
| Glu43  | Val56            | 0.8133         | Medium-range contacts     |
| <b>Glu43</b>   | <b>Phe49</b>     | 0.50552        | Short-range contacts      |
| Glu43  | Asp50            | 0.53931        | Short-range contacts      |
| Glu43  | Asn53            | 0.62552        | Short-range contacts      |
| Glu43  | Lys52            | 0.69664        | Short-range contacts      |
| Glu43  | Lys51            | 0.7532         | Short-range contacts      |
| Glu43  | Ile54            | 0.75365        | Short-range contacts      |

| <b>E43V <i>Tpv</i> sHSP14.3 Mutant Specific Contacts at position 43</b> |                  |                |                           |
|---|------------------|----------------|---------------------------|
| <b>Residue 1</b>  | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| <b>Val43</b>  | <b>Asp79</b>     | 0.50399        | long-range contacts       |
| Val43   | Lys87            | 0.58311        | long-range contacts       |
| Val43   | Val88            | 0.54553        | long-range contacts       |
| Val43   | Ser101           | 0.57846        | long-range contacts       |
| Val43   | Ala102           | 0.6203         | long-range contacts       |
| Val43   | Lys103           | 0.51698        | long-range contacts       |
| <b>Val43</b>  | <b>Tyr104</b>    | 0.5155         | long-range contacts       |
| Val43   | Gly107           | 0.59253        | long-range contacts       |
| Val43   | Ile108           | 0.70505        | long-range contacts       |
| Val43   | Leu109           | 0.7724         | long-range contacts       |
| Val43   | Thr110           | 0.74516        | long-range contacts       |
| Val43   | Val111           | 0.75305        | long-range contacts       |
| Val43   | Arg112           | 0.65753        | long-range contacts       |

“Table I.18 (cont’d)”

| E43V <i>Tpv</i> sHSP14.3 Mutant Specific Contacts at Position 43 |           |         |                       |
|--|-----------|---------|-----------------------|
| Residue 1  | Residue 2 | P-Score | Contact evaluation    |
| Val43  | Met113    | 0.53695 | long-range contacts   |
| Val43  | Lys114    | 0.52442 | long-range contacts   |
| Val43  | Lys55     | 0.76668 | Medium-range contacts |
| Val43  | Val56     | 0.80095 | Medium-range contacts |
| Val43  | Ser57     | 0.76235 | Medium-range contacts |
| Val43  | Val58     | 0.68765 | Medium-range contacts |
| Val43  | Asn59     | 0.51312 | Medium-range contacts |
| Val43  | Val62     | 0.64833 | Medium-range contacts |
| Val43  | Leu63     | 0.68398 | Medium-range contacts |
| Val43  | Thr64     | 0.70431 | Medium-range contacts |
| Val43  | Ile65     | 0.69615 | Medium-range contacts |
| Val43  | Ser66     | 0.646   | Medium-range contacts |
| Val43  | Asp50     | 0.56509 | Short-range contacts  |
| Val43  | Lys51     | 0.71896 | Short-range contacts  |
| Val43  | Lys52     | 0.64813 | Short-range contacts  |
| Val43  | Asn53     | 0.63646 | Short-range contacts  |
| Val43  | Ile54     | 0.77553 | Short-range contacts  |

**Table I. 19. List of the specific contacts which are present in the WT *Tpv* sHSP14.3, but not for the Y34FG48E *Tpv* sHSP14.3 mutant and vice versa.** Probability score is the probability of the two residues in contact within the 8 Å distance.

| WT <i>Tpv</i> sHSP14.3 Specific Contacts |           |         |                     |
|--|-----------|---------|---------------------|
| Residue 1                                | Residue 2 | P-score | Contacts evaluation |
| Ile27                                    | Ala67     | 0.5001  | long-range contacts |
| Val31                                    | Ile123    | 0.50164 | long-range contacts |
| Thr32                                    | Ile123    | 0.51263 | long-range contacts |
| Tyr34                                    | Val89     | 0.5068  | long-range contacts |
| Tyr34                                    | Val121    | 0.5154  | long-range contacts |
| Gln35                                    | Gly107    | 0.5103  | long-range contacts |
| Asp39                                    | Glu124    | 0.55505 | long-range contacts |
| Asp39                                    | Ile123    | 0.55527 | long-range contacts |
| Leu40                                    | Glu122    | 0.53661 | long-range contacts |
| Val41                                    | Gly107    | 0.50254 | long-range contacts |
| Val41                                    | Glu124    | 0.52839 | long-range contacts |
| Leu42                                    | Ile123    | 0.51384 | long-range contacts |
| Ala47                                    | Ala102    | 0.50309 | long-range contacts |
| Phe49                                    | Leu109    | 0.508   | long-range contacts |
| Ile54                                    | Ser105    | 0.50174 | long-range contacts |
| Ile54                                    | Tyr104    | 0.51169 | long-range contacts |
| Ile54                                    | Ile108    | 0.51304 | long-range contacts |
| Ile54                                    | Leu109    | 0.51623 | long-range contacts |
| Lys55                                    | Ile108    | 0.5146  | long-range contacts |
| Val56                                    | Gly107    | 0.51962 | long-range contacts |
| Ser57                                    | Ser101    | 0.50227 | long-range contacts |
| Ser57                                    | Lys103    | 0.50486 | long-range contacts |
| Val58                                    | Glu106    | 0.50057 | long-range contacts |
| Val58                                    | Ile100    | 0.53535 | long-range contacts |
| Lys60                                    | Ser105    | 0.50049 | long-range contacts |

“Table I.19 (cont’d)”

| <b>WT <i>Tpv</i> sHSP14.3 Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>                                | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Lys60   | Ala102           | 0.50136        | long-range contacts        |
| Asn61   | Ala102           | 0.5715         | long-range contacts        |
| Val62   | Ile100           | 0.53919        | long-range contacts        |
| Leu63   | Ile123           | 0.5108         | long-range contacts        |
| Ile65   | Ile123           | 0.50447        | long-range contacts        |
| Ile65   | Lys90            | 0.51142        | long-range contacts        |
| Ile65   | Asp99            | 0.51814        | long-range contacts        |
| Ser66   | Asp99            | 0.55428        | long-range contacts        |
| Ala67   | Gln98            | 0.50843        | long-range contacts        |
| Leu42   | Asn61            | 0.51032        | Medium-range contacts      |
| Asn61   | Gln80            | 0.51041        | Medium-range contacts      |
| Ile65   | Tyr77            | 0.5043         | Medium-range contacts      |
| Ser66   | Val88            | 0.50573        | Medium-range contacts      |
| Ser66   | Val89            | 0.50728        | Medium-range contacts      |
| Glu68   | Val85            | 0.56551        | Medium-range contacts      |
| Thr75   | Lys87            | 0.51489        | Medium-range contacts      |
| Thr75   | Val88            | 0.59051        | Medium-range contacts      |
| Ala102  | Lys114           | 0.50025        | Medium-range contacts      |
| Arg112  | Glu124           | 0.50503        | Medium-range contacts      |
| Met12   | Val20            | 0.50916        | Short-range contacts       |
| Ile13   | Val20            | 0.62073        | Short-range contacts       |
| Val76   | Lys87            | 0.56952        | Short-range contacts       |
| Tyr77   | Tyr86            | 0.50164        | Short-range contacts       |
| Ile78   | Val85            | 0.53982        | Short-range contacts       |

| <b>Y34FG48E <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|--|------------------|----------------|----------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Pro29  | Leu91            | 0.50074        | long-range contacts        |
| Pro29  | Pro92            | 0.51462        | long-range contacts        |
| Thr32  | Asn61            | 0.50581        | long-range contacts        |
| <b>Phe34</b>   | <b>Ser66</b>     | 0.51349        | long-range contacts        |
| <b>Phe34</b>   | <b>Lys114</b>    | 0.50579        | long-range contacts        |
| Ser38  | Ser66            | 0.50114        | long-range contacts        |
| Ser38  | Val82            | 0.51439        | long-range contacts        |
| Ser38  | Lys114           | 0.50501        | long-range contacts        |
| Asp39  | Lys84            | 0.50919        | long-range contacts        |
| Leu40  | Val82            | 0.54706        | long-range contacts        |
| Leu40  | Thr115           | 0.50471        | long-range contacts        |
| Val41  | Val82            | 0.52846        | long-range contacts        |
| Leu42  | Arg81            | 0.60611        | long-range contacts        |
| Glu43  | Gln80            | 0.52518        | long-range contacts        |
| Glu43  | Asn120           | 0.51403        | long-range contacts        |
| Glu43  | Val121           | 0.50652        | long-range contacts        |
| Lys52  | Thr115           | 0.51928        | long-range contacts        |
| Lys55  | Pro92            | 0.50012        | long-range contacts        |
| Val56  | Tyr86            | 0.50285        | long-range contacts        |
| Lys60  | Val85            | 0.50168        | long-range contacts        |
| Asn61  | Val85            | 0.54416        | long-range contacts        |
| Gln80  | Leu109           | 0.51841        | long-range contacts        |
| Phe26  | Ala47            | 0.5052         | Medium-range contacts      |
| Tyr28  | Ala44            | 0.50557        | Medium-range contacts      |

“Table I.19 (cont’d)”

| Y34FG48E <i>Tpv</i> sHSP14.3 Mutant Specific Contacts |           |         |                       |
|---|-----------|---------|-----------------------|
| Residue 1   | Residue 2 | P-score | Contacts evaluation   |
| Tyr28   | Met46     | 0.51466 | Medium-range contacts |
| Gln35   | Ser57     | 0.50513 | Medium-range contacts |
| Glu43   | Asn59     | 0.50579 | Medium-range contacts |
| Lys60   | Val82     | 0.51747 | Medium-range contacts |
| Asn61   | Val82     | 0.56622 | Medium-range contacts |
| Asn61   | Lys84     | 0.5437  | Medium-range contacts |
| Val62   | Asp79     | 0.513   | Medium-range contacts |
| Thr64   | Val82     | 0.5785  | Medium-range contacts |
| Ile65   | Val82     | 0.59319 | Medium-range contacts |
| Ala67   | Val82     | 0.53415 | Medium-range contacts |
| Arg69   | Arg81     | 0.53462 | Medium-range contacts |
| Tyr86   | Asp99     | 0.50799 | Medium-range contacts |
| Tyr86   | Ile108    | 0.50572 | Medium-range contacts |
| Ile108  | Val121    | 0.50388 | Medium-range contacts |
| Met1  | Phe7      | 0.53522 | Short-range contacts  |
| Tyr2  | Thr9      | 0.55012 | Short-range contacts  |
| Thr3  | Thr9      | 0.50116 | Short-range contacts  |
| Ile5  | Met12     | 0.55706 | Short-range contacts  |
| Lys6  | Met12     | 0.58457 | Short-range contacts  |
| Lys6  | Ile13     | 0.55351 | Short-range contacts  |
| Ala67   | Thr75     | 0.50125 | Short-range contacts  |
| Lys70   | Asp79     | 0.54338 | Short-range contacts  |
| Arg71   | Asp79     | 0.50942 | Short-range contacts  |
| Glu72   | Ile78     | 0.56184 | Short-range contacts  |
| Gln80   | Val88     | 0.53592 | Short-range contacts  |
| Lys90   | Asp99     | 0.50567 | Short-range contacts  |

**Table I. 20. List of residues forming contacts with the residues Y34, G48 in WT *Tpv* sHSP14.3 and F34, E48 in Y34FG48E *Tpv* sHSP14.3 double mutant. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| WT <i>Tpv</i> sHSP14.3 Contacts at Position 34 and 48 |               |         |                      |
|---|---------------|---------|----------------------|
| Residue 1   | Residue 2     | P-Score | Contact evaluation   |
| <b>Tyr34</b>  | <b>Val89</b>  | 0.5068  | long-range contacts  |
| Tyr34   | Val58         | 0.51045 | long-range contacts  |
| <b>Tyr34</b>  | <b>Val121</b> | 0.5154  | long-range contacts  |
| Tyr34   | Leu109        | 0.5287  | long-range contacts  |
| Tyr34   | Leu63         | 0.53396 | long-range contacts  |
| Tyr34   | Val88         | 0.53549 | long-range contacts  |
| Tyr34   | Lys87         | 0.54221 | long-range contacts  |
| Tyr34   | Ile65         | 0.56108 | long-range contacts  |
| Gly48   | Ile108        | 0.53602 | long-range contacts  |
| Gly48   | Ile54         | 0.66209 | long-range contacts  |
| Tyr34   | Ala44         | 0.50735 | Short-range contacts |
| Tyr34   | Glu43         | 0.62804 | Short-range contacts |
| Tyr34   | Val41         | 0.74313 | Short-range contacts |
| Tyr34   | Leu42         | 0.74691 | Short-range contacts |
| Tyr34   | Leu40         | 0.75437 | Short-range contacts |
| Gly48   | Leu109        | 0.5333  | Short-range contacts |
| Gly48   | Gly107        | 0.67628 | Short-range contacts |



“Table I.20 (cont’d)”

| Y34FG48E <i>Tpv</i> sHSP14.3 Mutant Contacts at Position 34 and 48 |           |         |                      |
|--|-----------|---------|----------------------|
| Residue 1  | Residue 2 | P-Score | Contact evaluation   |
| Phe34  | Val58     | 0.52197 | long-range contacts  |
| Phe34  | Leu63     | 0.53983 | long-range contacts  |
| Phe34  | Ile65     | 0.54878 | long-range contacts  |
| Phe34  | Ser66     | 0.51349 | long-range contacts  |
| Phe34  | Lys87     | 0.56374 | long-range contacts  |
| Phe34  | Val88     | 0.55543 | long-range contacts  |
| Phe34  | Leu109    | 0.50415 | long-range contacts  |
| Phe34  | Lys114    | 0.50579 | long-range contacts  |
| Glu48  | Gly107    | 0.63914 | long-range contacts  |
| Glu48  | Ile108    | 0.51566 | long-range contacts  |
| Glu48  | Leu109    | 0.51782 | long-range contacts  |
| Phe34  | Leu40     | 0.77504 | Short-range contacts |
| Phe34  | Val41     | 0.76626 | Short-range contacts |
| Phe34  | Leu42     | 0.76225 | Short-range contacts |
| Phe34  | Glu43     | 0.63923 | Short-range contacts |
| Phe34  | Ala44     | 0.50906 | Short-range contacts |
| Glu48  | Ile54     | 0.67028 | Short-range contacts |

**Table I. 21. List of the specific contacts which are present in the WT *Tpv* sHSP14.3, but not for the G48E1108K *Tpv* sHSP14.3 mutant and vice versa.** Probability score is the probability of the two residues in contact within the 8 Å distance.

| WT <i>Tpv</i> sHSP14.3 Specific Contacts |           |         |                       |
|--|-----------|---------|-----------------------|
| Residue 1                                | Residue 2 | P-score | Contacts evaluation   |
| Ile27                                    | Ala67     | 0.5001  | long-range contacts   |
| Val31                                    | Ile123    | 0.50164 | long-range contacts   |
| Val41                                    | Glu124    | 0.52839 | long-range contacts   |
| Leu42                                    | Ile123    | 0.51384 | long-range contacts   |
| Ala47                                    | Ala102    | 0.50309 | long-range contacts   |
| Phe49                                    | Leu109    | 0.508   | long-range contacts   |
| Ile54                                    | Ser105    | 0.50174 | long-range contacts   |
| Ile54                                    | Tyr104    | 0.51169 | long-range contacts   |
| Ile54                                    | Leu109    | 0.51623 | long-range contacts   |
| Lys55                                    | Ile108    | 0.5146  | long-range contacts   |
| Val56                                    | Gly107    | 0.51962 | long-range contacts   |
| Ser57                                    | Ser101    | 0.50227 | long-range contacts   |
| Ser57                                    | Lys103    | 0.50486 | long-range contacts   |
| Val58                                    | Glu106    | 0.50057 | long-range contacts   |
| Val58                                    | Ile100    | 0.53535 | long-range contacts   |
| Lys60                                    | Ser105    | 0.50049 | long-range contacts   |
| Lys60                                    | Ala102    | 0.50136 | long-range contacts   |
| Leu63                                    | Ile123    | 0.5108  | long-range contacts   |
| Ile65                                    | Ile123    | 0.50447 | long-range contacts   |
| Ile65                                    | Asp99     | 0.51814 | long-range contacts   |
| Ala67                                    | Gln98     | 0.50843 | long-range contacts   |
| Leu42                                    | Asn61     | 0.51032 | Medium-range contacts |
| Asn61                                    | Gln80     | 0.51041 | Medium-range contacts |
| Ile65                                    | Tyr77     | 0.5043  | Medium-range contacts |
| Ser66                                    | Val88     | 0.50573 | Medium-range contacts |
| Ser66                                    | Val89     | 0.50728 | Medium-range contacts |

“Table I.21 (cont’d)”

| <b>WT <i>Tpv</i> sHSP14.3 Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>                                | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Thr75   | Lys87            | 0.51489        | Medium-range contacts      |
| Ala102  | Lys114           | 0.50025        | Medium-range contacts      |
| Arg112  | Glu124           | 0.50503        | Medium-range contacts      |
| Met12   | Val20            | 0.50916        | Short-range contacts       |
| Ile13   | Val20            | 0.62073        | Short-range contacts       |
| Tyr77   | Tyr86            | 0.50164        | Short-range contacts       |
| Ile78   | Val85            | 0.53982        | Short-range contacts       |

| <b>G48E1108K <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|---|------------------|----------------|----------------------------|
| <b>Residue 1</b>  | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Ser38   | Ser66            | 0.5019         | long-range contacts        |
| Lys60   | Val88            | 0.50406        | long-range contacts        |
| Thr32   | Asn61            | 0.5066         | long-range contacts        |
| Gln35   | Thr64            | 0.50181        | long-range contacts        |
| Val56   | Tyr86            | 0.51061        | long-range contacts        |
| Lys55   | Pro92            | 0.50426        | long-range contacts        |
| Leu42   | Arg81            | 0.56136        | long-range contacts        |
| Leu40   | Gln80            | 0.51012        | long-range contacts        |
| Val41   | Val82            | 0.50312        | long-range contacts        |
| Leu40   | Val82            | 0.51756        | long-range contacts        |
| Thr64   | Thr110           | 0.5033         | long-range contacts        |
| Pro29   | Leu91            | 0.50365        | long-range contacts        |
| Pro29   | Pro92            | 0.51593        | long-range contacts        |
| Leu40   | Leu109           | 0.50026        | long-range contacts        |
| Leu33   | Lys108           | 0.50761        | long-range contacts        |
| Leu40   | Thr115           | 0.50405        | long-range contacts        |
| Ser38   | Lys114           | 0.50369        | long-range contacts        |
| Tyr34   | Val111           | 0.50339        | long-range contacts        |
| Arg69   | Arg81            | 0.50214        | Medium-range contacts      |
| Tyr86   | Asp99            | 0.50957        | Medium-range contacts      |
| Lys108  | Val121           | 0.53961        | Medium-range contacts      |
| Ala67   | Val82            | 0.50801        | Medium-range contacts      |
| Val76   | Leu91            | 0.50414        | Medium-range contacts      |
| Tyr28   | Ala44            | 0.50457        | Medium-range contacts      |
| Glu43   | Asn59            | 0.50468        | Medium-range contacts      |
| Val62   | Asp79            | 0.51115        | Medium-range contacts      |
| Ile65   | Val82            | 0.56702        | Medium-range contacts      |
| Tyr28   | Met46            | 0.5148         | Medium-range contacts      |
| Thr64   | Val82            | 0.55593        | Medium-range contacts      |
| Phe26   | Ala47            | 0.50787        | Medium-range contacts      |
| Asn61   | Val82            | 0.52903        | Medium-range contacts      |
| Gln35   | Ser57            | 0.5075         | Medium-range contacts      |
| Tyr86   | Lys108           | 0.50086        | Medium-range contacts      |
| Asn61   | Lys84            | 0.53095        | Medium-range contacts      |
| Met1  | Phe7             | 0.50406        | Short-range contacts       |
| Phe49   | Lys55            | 0.50079        | Short-range contacts       |
| Glu72   | Ile78            | 0.52521        | Short-range contacts       |
| Ala67   | Thr75            | 0.50162        | Short-range contacts       |
| Gln80   | Val88            | 0.54007        | Short-range contacts       |
| Lys70   | Asp79            | 0.52988        | Short-range contacts       |
| Lys90   | Asp99            | 0.50707        | Short-range contacts       |

**Table I. 22. List of residues forming contacts with the residues G48, I108 in WT *Tpv* sHSP14.3 and E48, K108 in G48E108K *Tpv* sHSP14.3 Mutant.** Probability score is the probability of the two residues in contact within the 8 Å distance.

| WT <i>Tpv</i> sHSP14.3 Contacts at Position 48,108 |           |         |                       |
|--|-----------|---------|-----------------------|
| Residue 1  | Residue 2 | P-score | Contacts evaluation   |
| Gly48  | Leu109    | 0.5333  | long-range contacts   |
| Gly48  | Ile108    | 0.53602 | long-range contacts   |
| Gly48  | Gly107    | 0.67628 | long-range contacts   |
| Ile108   | Glu122    | 0.53133 | Medium-range contacts |
| Ile108   | Ile123    | 0.68024 | Medium-range contacts |
| Ile108   | Glu124    | 0.7122  | Medium-range contacts |
| Gly48  | Ile54     | 0.66209 | Short-range contacts  |

| G48E108K <i>Tpv</i> sHSP14.3 Mutant Specific Contacts at Position 48 and 108 |           |         |                       |
|--|-----------|---------|-----------------------|
| Residue 1  | Residue 2 | P-score | Contacts evaluation   |
| Glu48  | Gly107    | 0.63914 | long-range contacts   |
| Glu48  | Lys108    | 0.56462 | long-range contacts   |
| Glu48  | Leu109    | 0.51887 | long-range contacts   |
| Lys108   | Val121    | 0.53961 | Medium-range contacts |
| Lys108   | Glu122    | 0.59337 | Medium-range contacts |
| Lys108   | Ile123    | 0.7081  | Medium-range contacts |
| Lys108   | Glu124    | 0.74104 | Medium-range contacts |
| Tyr86  | Lys108    | 0.50086 | Medium-range contacts |
| Glu48  | Ile54     | 0.6692  | Short-range contacts  |

**Table I. 23. List of the specific contacts which are present in the WT *Tpv* sHSP14.3, but not for the K87E *Tpv* sHSP14.3 mutant and vice versa.** Probability score is the probability of the two residues in contact within the 8 Å distance.

| WT <i>Tpv</i> sHSP14.3 Mutant Specific Contacts |           |         |                     |
|---|-----------|---------|---------------------|
| Residue 1                                       | Residue 2 | P-score | Contacts evaluation |
| Ile27   | Ala67     | 0.5001  | long-range contacts |
| Val31   | Ile123    | 0.50164 | long-range contacts |
| Gln35   | Gly107    | 0.5103  | long-range contacts |
| Asp39   | Glu124    | 0.55505 | long-range contacts |
| Asp39   | Ile123    | 0.55527 | long-range contacts |
| Leu40   | Glu122    | 0.53661 | long-range contacts |
| Val41   | Gly107    | 0.50254 | long-range contacts |
| Val41   | Glu124    | 0.52839 | long-range contacts |
| Leu42   | Val89     | 0.51088 | long-range contacts |
| Leu42   | Ile123    | 0.51384 | long-range contacts |
| Ala44   | Lys87     | 0.50482 | long-range contacts |
| Ala47   | Ala102    | 0.50309 | long-range contacts |
| Ala47   | Pro92     | 0.50486 | long-range contacts |
| Phe49   | Leu109    | 0.508   | long-range contacts |
| Ile54   | Ser105    | 0.50174 | long-range contacts |
| Ile54   | Tyr104    | 0.51169 | long-range contacts |
| Ile54   | Ile108    | 0.51304 | long-range contacts |
| Ile54   | Leu109    | 0.51623 | long-range contacts |

“Table I.23 (cont’d)”

| <b>WT <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|--|------------------|----------------|----------------------------|
| <b>Residue 1</b>                                       | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Ile54  | Lys87            | 0.52954        | long-range contacts        |
| Lys55  | Ile108           | 0.5146         | long-range contacts        |
| Val56  | Gly107           | 0.51962        | long-range contacts        |
| Ser57  | Lys103           | 0.50486        | long-range contacts        |
| Lys60  | Ser105           | 0.50049        | long-range contacts        |
| Lys60  | Ala102           | 0.50136        | long-range contacts        |
| Leu63  | Ile123           | 0.5108         | long-range contacts        |
| Ile65  | Ile123           | 0.50447        | long-range contacts        |
| Ile65  | Lys90            | 0.51142        | long-range contacts        |
| Ala67  | Gln98            | 0.50843        | long-range contacts        |
| Lys87  | Lys114           | 0.50455        | long-range contacts        |
| Leu42  | Asn61            | 0.51032        | Medium-range contacts      |
| Asn61  | Gln80            | 0.51041        | Medium-range contacts      |
| Ser66  | Val88            | 0.50573        | Medium-range contacts      |
| Ser66  | Val89            | 0.50728        | Medium-range contacts      |
| Thr75  | Leu91            | 0.50275        | Medium-range contacts      |
| Thr75  | Pro92            | 0.51105        | Medium-range contacts      |
| Thr75  | Lys87            | 0.51489        | Medium-range contacts      |
| Thr75  | Val88            | 0.59051        | Medium-range contacts      |
| Tyr86  | Ser105           | 0.50354        | Medium-range contacts      |
| Lys87  | Asp99            | 0.50756        | Medium-range contacts      |
| Lys87  | Gly107           | 0.50808        | Medium-range contacts      |
| Ala102   | Lys114           | 0.50025        | Medium-range contacts      |
| Arg112   | Glu124           | 0.50503        | Medium-range contacts      |
| Met12  | Val20            | 0.50916        | Short-range contacts       |
| Ile13  | Val20            | 0.62073        | Short-range contacts       |
| Val76  | Lys87            | 0.56952        | Short-range contacts       |
| Tyr77  | Tyr86            | 0.50164        | Short-range contacts       |

| <b>K87E <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|--|------------------|----------------|----------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Pro29  | Leu91            | 0.50422        | long-range contacts        |
| Pro29  | Pro92            | 0.51773        | long-range contacts        |
| Thr32  | Asn61            | 0.51218        | long-range contacts        |
| Gln35  | Thr64            | 0.50256        | long-range contacts        |
| Ser38  | Ser66            | 0.50526        | long-range contacts        |
| Ser38  | Lys114           | 0.50409        | long-range contacts        |
| Leu40  | Thr115           | 0.50319        | long-range contacts        |
| Leu42  | Arg81            | 0.55687        | long-range contacts        |
| Glu43  | Gln80            | 0.51857        | long-range contacts        |
| Glu43  | Val121           | 0.50483        | long-range contacts        |
| Lys52  | Thr115           | 0.50529        | long-range contacts        |
| Lys55  | Pro92            | 0.50012        | long-range contacts        |
| Val56  | Tyr86            | 0.50015        | long-range contacts        |
| Ser57  | Val85            | 0.51052        | long-range contacts        |
| Lys60  | Glu87            | 0.56505        | long-range contacts        |
| Phe26  | Ala47            | 0.51051        | Medium-range contacts      |
| Tyr28  | Ala44            | 0.50503        | Medium-range contacts      |
| Tyr28  | Met46            | 0.51738        | Medium-range contacts      |
| Gln35  | Ser57            | 0.51044        | Medium-range contacts      |
| Glu43  | Asn59            | 0.50546        | Medium-range contacts      |

“Table I.23 (cont’d)”

| <b>K87E <i>Tpv</i> sHSP14.3 Mutant Specific Contacts</b> |                  |                |                            |
|--|------------------|----------------|----------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-score</b> | <b>Contacts evaluation</b> |
| Lys52  | Arg71            | 0.50161        | Medium-range contacts      |
| Asn61  | Val82            | 0.52875        | Medium-range contacts      |
| Asn61  | Lys84            | 0.50732        | Medium-range contacts      |
| Val62  | Asp79            | 0.52201        | Medium-range contacts      |
| Thr64  | Val82            | 0.52051        | Medium-range contacts      |
| Ile65  | Val82            | 0.54007        | Medium-range contacts      |
| Tyr86  | Asp99            | 0.50316        | Medium-range contacts      |
| Ile108   | Val121           | 0.50151        | Medium-range contacts      |
| Met1   | Phe7             | 0.50338        | Short-range contacts       |
| Ala67  | Thr75            | 0.50174        | Short-range contacts       |
| Lys70  | Asp79            | 0.52121        | Short-range contacts       |
| Glu72  | Ile78            | 0.53359        | Short-range contacts       |
| Lys90  | Asp99            | 0.50271        | Short-range contacts       |

**Table I. 24. List of residues forming contacts with the residues K87 in WT *Tpv* sHSP14.3 and E87 in K87E *Tpv* sHSP14.3 Mutant. Probability score is the probability of the two residues in contact within the 8 Å distance.**

| <b>WT <i>Tpv</i> sHSP14.3 Specific Contacts at Position 87</b> |                  |                |                           |
|--|------------------|----------------|---------------------------|
| <b>Residue 1</b>   | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| Lys87  | Lys114           | 0.50455        | long-range contacts       |
| Lys87  | Asp99            | 0.50756        | Medium-range contacts     |
| Lys87  | Arg112           | 0.58           | long-range contacts       |
| Lys87  | Ile108           | 0.59706        | long-range contacts       |
| Lys87  | Gly107           | 0.50808        | Medium-range contacts     |
| Lys87  | Tyr104           | 0.51959        | Medium-range contacts     |
| Lys87  | Val111           | 0.55979        | Long-range contacts       |
| Lys87  | Thr110           | 0.60462        | Medium-range contacts     |
| Lys87  | Ser101           | 0.62813        | Medium-range contacts     |
| Lys87  | Lys103           | 0.67389        | Medium-range contacts     |
| Lys87  | Leu109           | 0.68074        | Medium-range contacts     |
| Lys87  | Ala102           | 0.6947         | Medium-range contacts     |

| <b>K87E <i>Tpv</i> sHSP14.3 Mutant Specific Contacts at position 87</b> |                  |                |                           |
|---|------------------|----------------|---------------------------|
| <b>Residue 1</b>  | <b>Residue 2</b> | <b>P-Score</b> | <b>Contact evaluation</b> |
| Glu87   | Ser101           | 0.61722        | Medium-range contacts     |
| Glu87   | Arg112           | 0.55173        | long-range contacts       |
| Glu87   | Ala102           | 0.68581        | Medium-range contacts     |
| Glu87   | Lys103           | 0.6922         | Medium-range contacts     |
| Glu87   | Tyr104           | 0.52687        | Medium-range contacts     |
| Glu87   | Ile108           | 0.58205        | Medium-range contacts     |
| Glu87   | Leu109           | 0.67506        | Medium-range contacts     |
| Glu87   | Thr110           | 0.61923        | Medium-range contacts     |
| Glu87   | Val111           | 0.53889        | long-range contacts       |



## CURRICULUM VITAE

### PERSONAL INFORMATION

Surname, Name: Zabcı, Sema  
Nationality: Turkish (TC)  
Date and Place of Birth: 1 May 1989, Malatya  
Marital Status: Single  
Phone: +90 5067810980  
email: sema.zabc@yahoo.com

### EDUCATION

| Degree      | Institution                          | Year of Graduation |
|-------------|--------------------------------------|--------------------|
| MS          | METU Biological Sciences             | 2015               |
| BS          | METU Biological Sciences             | 2008               |
| High School | Atatürk Anadolu High School, Malatya | 2007               |

### FOREIGN LANGUAGES

Advanced English

### PUBLICATIONS

#### Published Abstracts

A. Rafiq, S. Zabcı, Y. Kaçer, S. Kocabiyik, Temperature dependent variations in oligomer status of an archeal small heat shock protein, TpvHSP14.3.

"BioMicroWorld Abstract Book", (2017), p.335.

Semra Kocabiyik, Sema Zabcı, Tülay Y. Swenson, Gene Expression analysis of thermoacidophilic archaeon *Thermoplasma volcanium* GSS1 under mild- and severe-pH stress. "BioMicroWorld 2017 Book of Abstracts", (2017), p.321.

Semra Kocabiyik, Sema Zabcı, Alpha crystallin domain mutations increase holdase capacity of the archeal small heat shock protein, Tpv HSP 14.3, *New Biotechnology*, Volume 33, Supplement, 2016, Pages S103-S104, ISSN 1871-6784,

<https://doi.org/10.1016/j.nbt.2016.06.1083>.

Zabcı, S., & Kocabiyik, S. (2016). Effect of N-terminal extension of small heat shock protein, Tpv HSP 14.3 on its molecular chaperone activity. *New BIOTECHNOLOGY*, 33(Supplement), S98.

<https://doi.org/10.1016/j.nbt.2016.06.1065>

Zabcı, S., Celebi, M., & Kocabiyik, S. (2015). Gene expression profile of *Thermoplasma volcanium* GSS1 under mild and severe oxidative stress. *FEBS JOURNAL*, 282, 75. Retrieved from

<http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip&db=edswsc&AN=000362570601068&site=eds-live>