Characterization of high temperature requirement serine protease and its binding partners

By

Saujanya Acharya
[LIFE09201204002]

Tata Memorial Centre (TMC),
Advanced Centre for Treatment, Research and Education in Cancer (ACTREC),
Kharghar, Navi Mumbai – 410210.

A thesis submitted to the
Board of Studies in Life Sciences
In partial fulfillment of requirements

For the Degree of

DOCTOR OF PHILOSOPHY

Of

HOMI BHABHA NATIONAL INSTITUTE

May 2020

Homi Bhabha National Institute

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Ms. Saujanya Acharya entitled "**Characterization of high temperature requirement serine protease and its binding partners**" and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Chairperson–Dr. Rajiv Sarin — Date: 19-05-2020

Guide/Convener –Dr. Kakoli Bose — Date: 19-05-2020

External Examiner–Dr.Arun Shukla — Date: 19-05-2020

Member –Dr. Sorab Dalal — Date: 19-05-2020

Member –Dr. Neelam Shirsat — Date: 19-05-2020

Member -Dr. Samir Maji — Date: 19-05-2020

Invitee - Prof. S.V. Chiplunkar — Date: 19-05-2020

Invitee -Dr. Mukesh Kumar — Date: 19-05-2020

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

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

Date: 19-05-2020
Place: Navi Mumbai

Dr. Kakoli Bose, Guide

STATEMENT BY AUTHOR

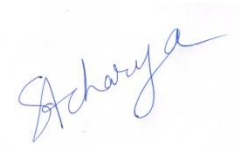
This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

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

Saujanya Acharya

DECLARATION

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



Saujanya Acharya

List of Publications arising from the thesis

Journal:

Published

- **Saujanya Acharya**, Shubhankar Dutta, Sucheta Chopra & Kakoli Bose. "**Identification of a distal allosteric ligand binding pocket in HtrA3**". *Biochemical and Biophysical Research Communications, 2019 July, https://doi.org/10.1016/j.bbrc.2019.07.005*
- **Saujanya Acharya**, Shubhankar Dutta & Kakoli Bose. "**A distinct concerted mechanism of structural dynamism defines activity of human serine protease HtrA3**". *Biochemical Journal, 2020 January, https://doi.org/10.1042/BCJ20190706*

Book/Book Chapter:

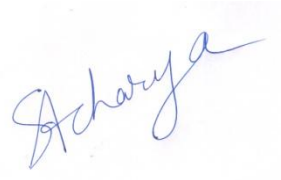
- **Acharya S**, Kuppili RR, Chaganti LK**,** Bose K. "**Proteases in Apoptosis: Protocols and Methods**". *Springer*, 2017, DOI: 10.1007/978-3-319-19497-4

Conference/Symposium

Poster Presentations:

- ***1st Global Cancer Genomics Consortium Symposium (GCGC).*** November 19-21, Advanced Centre for Treatment Research and Education in Cancer (ACTREC), Navi Mumbai.
- ***International Symposium on Conceptual Advances in Cellular Homeostasis Regulated by Proteases and Chaperones the Present, the Future and Impact on Human Diseases*.** Organized by ACTREC, Tata Memorial Centre, Kharghar, Navi-Mumbai, and Maharashtra, India. December 3-6, 2013.

- **Biophysics-Pashchim V** meeting organized by Tata Institute of Fundamental Research, Mumbai, India.
- **Biophysics-Pashchim VI** meeting organized by Tata Institute of Fundamental Research, Mumbai, India.
- ***A Conference of new ideas in Cancer- Challenging dogmas***. TMC Platinum Jubilee Celebrations. 26th February to 28th February 2016, Mumbai, India.
- ***Indo-US conference on Advances in Enzymology: Implications in Health, Disease and Therapeutics*** held at TMC-ACTREC on 15th- 19th January 2017. **Second best poster award.**
- **31st Annual Symposium of "The Protein Society"** at Montreal, Canada, 24th - 27th July, 2017.
- ***Indo-US Conference on Sculpting the Future of Medicine-Gateway to Post Proteogenenomic Era*** held at TMC-ACTREC on 8th – 11th December 2018.
- All **National Research Scholar's meet (NRSM)** at ACTREC, Mumbai, held annually since 2012 to 2018.



Saujanya Acharya

Dedicated to my beloved parents and my best friend Pramathesh

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to all the people who have been my pillars of support throughout these years and have helped in the completion of this work directly or indirectly.

Firstly, I would like to acknowledge my mentor, Dr. Kakoli Bose, for her guidance and encouragement. It was an invaluable learning curve under her mentorship. Working in her laboratory not only helped me immensely in improving execution, critical analysis and documentation of my research work but also tipped my research interests in new directions.

My heartfelt appreciation goes to my doctoral committee members, Dr. Rajiv Sarin, Dr. Sorab, Dr. Neelam Shirsat and Dr. Samir Maji. I owe them a debt of gratitude for their patience and concern when the "going got tough". Their valuable suggestions and insight helped steer my work towards a meaningful conclusion.

I am extremely thankful to all the present and past members of Bose lab. My seniors, Late Dr. Lalith Kumar Chaganti, Dr. Nitu Singh, Dr. Raja Reddy Kuppili who welcomed me with open hearts, made me feel comfortable in a new environment. Over the years they guided me in honing my technical skills. My lovely juniors, Ajay, Raghupathi, Rashmi, Aasna, Sucheta, Rucha, Shubham, Shubhankar and Jitesh, have made this journey merry and delightful to say the least. I thank them sincerely for partaking of my smiles and sorrows alike. I express my warmest thanks to my trainees Pooja, Shweta, Mriganka, Swathi, Shashank, Amruta, Deep and Abhilasha. For every little thing I taught them, I learnt much more in return. I am also thankful to Mr. P.L Padale and Ms. Snehal Pandav for managing lab resources smoothly and helping with official paperwork which can often become daunting.

I would like to extend my personal gratitude to all my batchmates for their unflinching support, both intellectually and emotionally.

Special mention needs to be made of our 'next-door neighbours and Prasanna lab members' for reviving the spirit of "scientific soirees" in the PG room.

Most importantly, I feel honoured to thank those without whose support my Ph.D. journey could not have begun or ended. I thank my parents, my sibling and my best friend for constantly stressing over my career more than I do and going through my professional ups and downs like they are their very own. Needless to say, over the years their command over scientific jargons have become impressive and terms like "DC committee meeting", "manuscript submission", "thesis" etc. evoke the same sense of reverence and veneration as does in the minds of every Ph.D. student.

Last but not the least, I express my gratitude to my institute Advanced Centre for Treatment, Research and Education in Cancer for providing state of the art facilities round the clock. I am also very grateful towards the Sam-Mistry travel fund and HBNI international travel fund for supporting my first international conference.

Even when all the cogs of the wheel (hopefully I have covered them all) are in place, pivotal journeys of life do not begin without that inexplicable big push called "divine providence" and smaller nudges called "travel", "coffee" and "Netflix". So, thank you, your contribution has not gone unnoticed.

Qui Audet Adipiscitur!

(Who dares, wins.) *Saujanya Acharya*

CONTENTS

Synopsis

Homi Bhabha National Institute

SYNOPSIS OF Ph. D. THESIS

1. **Name of the Student: Saujanya Acharya**
2. **Name of the Constituent Institution: Tata Memorial Centre, ACTREC**
3. **Enrolment No. : LIFE09201204002, 1/10/12**
4. **Title of the Thesis: Characterization of high temperature requirement serine protease and its binding partners**

SYNOPSIS

Title of the project: Characterization of high temperature requirement serine protease and its binding partners

Introduction:

The HtrA (High temperature requirement protease A) proteins are categorized as a unique family of serine proteases that are found in a wide variety of organisms including bacteria and humans [1]. While, DegP or HtrA is the first member of this family to be identified in *Escherichia coli*, its human homologs include four members HtrA 1-4. Although, the prokaryotic HtrAs are well known for their chaperone and protease activities, their mammalian counterparts are also implicated as modulators of apoptosis and are associated with a wide variety of diseases such as arthritis, neurodegenerative diseases and cancer [1] Among the HtrA homologs, structures of

HtrA1 and HtrA2 have been solved, a low-resolution structure is available for HtrA3 while no structure for HtrA4 is currently available [2, 3]. All HtrA proteases share conserved C-terminal domains but subtle conformational changes and variations lead to differences in their specificity and may contribute to their multi-tasking ability[1].

HtrA2, the most extensively studied member of this family is a unique serine protease confirmed to be a positive regulator of apoptosis. It has been shown that the serine protease HtrA2 is able to induce apoptosis in human cells both in a caspase-independent (possibly through its serine protease activity) as well as caspase-dependent manner by binding to BIR2 domain of XIAP and disrupting its interaction with the activated caspase-9[4-6]. However, the intricacies of the mechanism involved in each of the above phenomenon still elude us. Apart from IAPs, several other important cellular proteins have been found to be targets of HtrA2, many of them are anti-apoptotic such as Pea-15 and Hax1. Although a few substrates are known, several others are yet to be identified which will highlight its physiological significance in the cell as well as help decipher the mechanism of action. With an aim at understanding the molecular mechanism, we carried out extensive *in silico* studies and identified a putative binding partner DUSP9 (Dual specificity protein phosphatase 9) and validated our findings through wet lab studies.

HtrA3, the third and one of the least characterized members of this family thought to be involved in both intrinsic as well non-classical cell death pathways is down-regulated in ovarian and endometrial cancers [7-10] and hence is proposed to be a tumor suppressor. HtrA3 was initially identified as a pregnancy-related protease. It is proposed that HtrA3 is an inhibitor of trophoblast invasion during placental development [9, 11]. In ovarian cancer, the decrease in expression of HtrA3 gene was correlated with the degree of malignancy and increasing grades of the tumor. Moreover, in endometrial cancers, a statistically significant negative correlation between the

levels of Tgf-β1 and HtrA3 proteins has been demonstrated [9]. It is therefore possible that it may participate in Tgf-β signaling pathway in cancer tissues although the molecular mechanism of this function is not known. Apart from suggestions that this phenomenon can be attributed to HtrA3 proteolytic activity, the mechanism remains elusive. These reports indicate that HtrA3 may be involved in carcinogenesis, though the molecular basis of this is not known. HtrA3 has been shown to promote cytotoxicity of chemotherapeutic drugs in lung cancer cell lines [7, 12]. This function is dependent on its proteolytic activity. Moreover, it has been proposed that HtrA3 may modulate cell death through a mechanism involving proteolysis of factors crucial to cell survival which prompts the investigation of HtrA3 binding partners and substrates. The described findings also show a possibility of using HtrA3 as a drug target in cancer therapy, in conjunction with chemotherapy. However any effort in this direction needs to be preceded by characterization of HtrA3 structure and function. The mechanism of regulation of its enzymatic activity also needs to be clarified. Till date, HtrA3 has been poorly characterized from the biochemical point of view. Its mechanism of activation has not been elucidated in detail and its known physiological substrates are also limited. It is also not known whether PDZ domain of HtrA3 plays a role in regulation of HtrA3 activity similar to the model of regulation proposed for HtrA2.

It is interesting to note that despite the functional similarity it is supposed to share with HtrA2, there is a distinct structural difference in terms of additional signal peptide, Insulin-like growth factor (IGF) binding and Kazal like domain (KI) present at the N-terminus, functions of which are yet unknown [1]. HtrA3 is synthesized as a 453 amino acid precursor protein, which consists of a signal peptide (1-17 residues), an IGFBP N-terminal (21-77 residues), a Kazal-like (64-128 residues) domain, a serine protease domain or SPD (175-340 residues), and one PDZ (359-444

residues) domain**.** The active-site residue Ser305 along with His191 and Asp227 form the catalytic triad. It has two isoforms, long and short produced by alternative mRNA splicing, where the latter differs from the former by the lack of PDZ domain and presence of a seven-amino acid long sequence (APSLAVH). On apoptotic stimulus, autocatalytic cleavage of the long isoform of HtrA3 occurs through removal of the N-terminal region spanning IGFBP and KI domains followed by its migration from the mitochondrial compartment to the cytosol in a process termed 'maturation'. This 'mature' active protein then triggers the onset of apoptosis. The crystal structure of the substrate-unbound HtrA3 (3.27 Å) represents a pyramidal trimer; where the orthosteric pocket is positioned 33 Å from the base of this pyramidal structure. The crystal structure does not define the important loops clearly, especially the regulatory loop LA and the loop L3 which is required for substrate specificity/ binding due to low electron density (high B factor). Therefore, despite providing an overview of the three dimensional structure of HtrA3, this substrate-unbound structure fails to elucidate the structural reorientations and conformational dynamics involved in HtrA3 activation. Our aim was to perform an in depth biophysical and biochemical analyses of the protein including functional studies, secondary/tertiary structural properties of the different domains and determination of its oligomerization status.

Hypothesis:

Our central hypothesis is that understanding the structure-function correlation of HtrA proteins, -2 and -3 along with identification of their substrates/binding partners will help delineating their mechanism of action and hence subsequently lead toward devising therapeutic strategies against diseases they are associated with.

Results:

Objective 1: Characterization of structural properties and conformational stability of HtrA3 and the role of different domains in protein function and specificity

Expression and purification of wild type HtrA3, its domains and variants

The constructs used for this study are described as follows. Mature HtrA3 (amino acids 129-453), active-site mutant (S305A), N-SPD (without the PDZ domain), SPD domain alone, N-terminal deleted SPD-PDZ, PDZ were used to study the contribution of each of the domains. A few N-terminal deleted variants of the wild type protein, Δ134 HtrA3 (amino acids 134-453) and Δ143 HtrA3 (amino acids 144-453) were generated to understand the importance of the of N-terminal region in HtrA3. The trimeric architecture of HtrA3 is stabilized through intermolecular Vander Waal's interactions by three crucial aromatic residues (F140, F142 and F255) which form a 'triple lock' that stabilizes the trimer. Therefore, to understand the importance of these 'Phenylalanine-lock' residues as well, we made HtrA3 F142D, HtrA3 F142A, HtrA3 F255D single mutations. These proteins were expressed in bacterial strain BL21 DE3 with C/N-terminal His_6-tag and purified using two steps- Ni-NTA affinity chromatography and Gel filtration (FPLC) chromatography to get rid of additives as well as impurities.

Secondary and tertiary structural properties of HtrA3 mutants and domains

Mature HtrA3, Δ134 HtrA3 variant, active-site mutant HtrA3 (S305A), monomeric mutant HtrA3 (F142D) as well as its domains SPD and PDZ, showed both alpha-helical and beta-sheet characteristics as expected. HtrA3 F142D, HtrA3 SPD and HtrA3 PDZ showed a decrease in α-helical and increase in β-sheet characteristics compared to the wild type protein, which is

suggestive of significant changes in their secondary structural properties. Fluorescence emission studies between 310-400 nm demonstrated that mature HtrA3 showed almost no change in emission maxima (~ 342 nm) compared to its active-site mutant (S305A) suggesting that the mutation did not perturb the structure of the protease. Comparative emission scans of urea denatured HtrA3 S305A (emission maxima ~ 348 nm) and the native protein (emission maxima ~ 342 nm) showed a blue shift of the emission maxima by 6 nm for the native protein thereby demonstrating that the protein has a well folded tertiary structure. Thermal stability of the variants demonstrate that T_m (melting temperature) of HtrA3 S305A and HtrA3 PDZ are 63°C and 46°C respectively suggesting that the removal of the N and SPD domains has a destabilizing effect on the protease.

Role of N-terminal region and the PDZ domain in oligomerization

HtrA3 wild type and active site mutant (S305A) with intact N-terminal region form trimers. PDZ domain alone also formed trimers according to our observation. HtrA3 N-SPD, HtrA3 SPD-PDZ, HtrA3 F142D and HtrA3 F255D are monomers according to our data. HtrA3 F142A exists as a mixed population of trimers (~30%) and monomers (~70%) indicating that the weak interaction between the trimers due to the mutation might lead them to dissociate. Surprisingly, HtrA3 N-SPD variant, despite its intact N-terminal region is a monomer. This hints towards the role of the PDZ domain in oligomerization. Therefore, these results point towards the combinatorial effect of the 'triple lock' (F140, 142, 255), comprising residues from the N-terminus and SPD as well as the PDZ domain in maintaining protein stability and consequently the oligomeric ensemble of HtrA3.

Role of oligomerization, individual domains and critical residues in protease activity

Substrate catalysis for all these variants and mutants was monitored over a time period of 2 hr using a generic serine protease substrate β-casein. Mature HtrA3, its variants (Δ134 HtrA3, Δ143 HtrA3) as well as HtrA3 N-SPD cleaved β-casein but HtrA3 F142D, HtrA3 F142A, HtrA3 F255D, HtrA3 SPD, HtrA3 SPD-PDZ were completely inactive or showed minimal activity even at high substrate concentrations highlighting the importance of the N-terminal region comprising critical 'triple lock' residues in trimerization and hence in HtrA3 activity. Mature HtrA3 and Δ134 HtrA3 had comparable activity suggesting the deletion did not affect the active-site conformation of the protease. Additionally, Δ143 HtrA3 and HtrA3 N-SPD exhibited activity much less than the wild type. These observations suggest the role of HtrA3 PDZ, the 'triple lock' phenylalanine residues in maintaining structural integrity, and consequently optimal catalytic activation of HtrA3.

Role of temperature in protease activity

Protease assays of HtrA3 were performed where HtrA3 variants were incubated with 6 μg of β-casein in the temperature range of 30ºC to 60ºC for 10 mins and 30 mins. The proteolytic products were then analyzed using SDS-PAGE gels. We observed that HtrA3 completely processes substrate at all temperatures except at 30ºC when incubated for 10 mins. However, with subsequent increase in temperature beyond 50ºC, we observed significant degradation of HtrA3, which might be due to protease destabilization as also observed by thermal denaturation studies.

HtrA3 cleaves beta casein at four specific sites

Proteolytically cleaved fragments of β-casein were resolved on SDS-PAGE gel and four cleaved fragments (between ~5-24 kDa) of β- casein were isolated and further subjected to N-terminal sequencing by Edman degradation. Upon analysis four preferred cleavage sites on β-casein were identified: 15A-16R, 71Q-72S, 138Q-139S and 170V-171M. Our observations suggest that HtrA3, similar to the other human HtrA homologs, has a strong preference for aliphatic residues at P1 (A, V, and Q) position. At the P2 position, aliphatic (L) and polar (S, T) residues are preferred, again similar to HtrA1 and HtrA2 (L). P3 and P4 positions are occupied by aliphatic (L, A), polar residues (T, Q), non polar (P) and acidic (E) residues, which also matches with that of HtrA1 & 2 (A, L, P, Q). For the P1' positions, we observed that HtrA3 preferred aliphatic (M) and polar amino acids (S, R), similar to other family members. While, P2' position was occupied by acidic (E), aromatic (F) and aliphatic (L) residues, P3' position had aliphatic (L, V), polar (T) as well as non polar (P) residues. At P4', HtrA3 showed a preference for aliphatic (L), aromatic (Y), non polar (P) and acidic (E) residues. Comparatively, HtrA1 and HtrA2 showed preference for similar residues at P1-P4 and P1'- P4' sites suggesting that they have similar (but not identical) substrate specificities

Mechanism of HtrA3 activation involves allosteric signal propagation

FITC-labeled generic serine protease substrate, β-casein was used to study the steady-state kinetic parameters to determine whether HtrA3 exhibits allostery. A wide range of substrate concentration (0-10 μM) was used to measure the initial rates of substrate cleavage which was then plotted as a function of the corresponding substrate concentrations. This plot of reaction velocities were fitted with modified Michaelis-Menten equation incorporating Hill's coefficient suggesting cooperative substrate binding behaviour.

Comparing the enzymatic parameters of HtrA3 and its homolog HtrA2, we observed that V_{max} for HtrA2 exceeded that of HtrA3 ~ 2 folds. Moreover, the catalytic efficiency and substrate turnover rate of HtrA2 was also more than that of HtrA3 [5, 13]. However, Hill constant and $K_{0.5}$ value for both HtrA2 (2.8±0.2 and HtrA2 $K_{0.5}$ 2.3±0.1 μM) and HtrA3 were not significantly different suggesting similar cooperativity and substrate binding affinity [4, 6, 13, 14]. Overall, our observations prove the existence of an allosteric activation of HtrA3 at the molecular level.

Identification of Selective Binding Pocket and important interacting residues using docking study

The modeled HtrA3 protein was subjected to SiteMap analysis and four putative binding sites were identified. Among them Site 1, designated as SBP_HtrA3, was chosen as the best site according to the score. Scoring was done based on the following parameters:

a) Number of available hydrogen donor and acceptor groups available for interaction

b) Size of the pocket depending on its capability to accommodate the heptameric peptides generated from the peptide library analysis

Site 1 was then chosen for further docking and MD simulations. Docking was performed with chain A of HtrA3 trimer model and heptamers from the library were generated as described in **section 3.2.5** in **chapter 3**. Using this method, the best ranked β-casein peptide **^{131}AMVKLYC^{137}**, was taken as a positive control. A second library was generated by mutating each residue of **^{131}AMVKLYC^{137}** with all the 20 amino acids one at a time and the best ranked peptides were chosen. An 11 mer peptide (LNTPLNNPKNN), which did not match the consensus pattern, was used as the negative control. P200 and R202 from the LB loop of the SPD and Q403, D404, G405 from the PDZ were the most common residues from SBP_HtrA3 involved in the protein-

peptide interactions. These residues are located far from the catalytic pocket of HtrA3 and further highlights the role of SBP_HtrA3 in HtrA3 allostery [14].

MD simulation analysis shows conformational reorientation of the catalytic pocket of HtrA3 in the presence of peptides

RMSD (Root Mean Squared Deviation) and RMSF (Root Mean Squared Fluctuation) plots of HtrA3-GMMMRFC and HtrA3-AMVKLYC highlighted largest fluctuations in L3 and L1 loop regions (260-305) ranging from 2.2 Å-2.5 Å and 1.8 Å in L2 loop when compared to unbound HtrA3. This suggests that loops L1, L2 and L3 may together regulate allostery by relaying information from the PDZ domain, towards the active site pocket. Hinge region in both the peptide bound complexes showed a deviation of about 1.2 Å from that of unbound HtrA3. Comparative analysis with unbound HtrA3 also showed that there were no significant overall SPD domain movements (175-340), however PDZ domain showed movements ranging from 0.5 Å to 1.9 Å which could be due to the movements seen in PDZ helices, namely H6 and H7. Critical assessment of the bound complexes demonstrated deviation in flexible regions that contain the catalytic triad, namely, LB (188-202) and LC loops (223-227) harbouring H191 and D227 respectively.

Both the peptide-bound complexes were subjected to atomic distance analysis to understand the conformational changes in the catalytic triad. The analysis showed that distances between nitrogen (Nδ1) atom of H191 and oxygen (Oγ) atom of S305 decreased by 1.8 Å in HtrA3-GMMMRFC complex, and 1.1 Å in HtrA3-AMVKLYC complex when compared to unbound HtrA3. However, the distance between nitrogen (Nε2) atom of H191 and oxygen (Oδ2) atom of D227 increased by 2 Å for HtrA3-GMMMRFC, and 1.7 Å for HtrA3-AMVKLYC with

reference to the unbound HtrA3. Apart from this, overall angular distance between nitrogen (Nε2) atom of H191, oxygen (Oγ) atom of S305 and oxygen (Oδ2) atom of D227 increased by 11.3 deg in HtrA3-GMMMRFC complex and 8.3 deg in HtrA3-AMVKLYC complex with respect to unbound HtrA3. Moreover, conformational changes were also observed in the relative orientation of oxyanion-hole residues, namely Y302, G303 and N304. The aromatic ring of Y302 exhibited an anti-clockwise flip moving towards the imidazole ring of H191. This further disrupted the interaction between G260 and D296, resulting in an inward L1 loop movement towards the orthosteric pocket leading to its subsequent stabilization. MD simulation analysis further demonstrated that LC loop that comprises D227 residue of the catalytic triad moved in close proximity to the active site pocket upon peptide binding by disrupting its Van der Waals interaction with β11 strands. All these movements synergistically shift the LB loop away from the proximal region of SPD thereby bringing the catalytic triad residues within optimal distance to facilitate the assembly of a well-formed catalytic pocket. Overall, post peptide binding in SBP_HtrA3, the reorientation of the catalytic triad residues along with formation of a functional oxyanion hole at the active site makes it compact and stable for further substrate binding and catalysis.

PDZ domain is indispensable for optimal catalytic activation in HtrA3

To understand whether allosteric activation of HtrA3 requires any direct involvement of the C-terminal PDZ domain, we determined the enzymatic parameters for the PDZ lacking variant (HtrA3 N-SPD). Reaction velocities were plotted as a function of substrate concentrations and fitted with modified Michaelis–Menten equation that accommodates Hill's cooperativity parameters. HtrA3 N-SPD cleaved β-casein with steady state kinetic parameters of $K_{0.5}$ =1.04±0.06 and V_{max} = 2.90±0.26 x 10^{-13} M.s^{-1}. On comparison of the kinetic parameters, the

V_{max}, k_{cat} and $k_{cat}/K_{0.5}$ of this variant were significantly less compared to the wild type mature HtrA3 by ~6.7 x 10^3, 3.6 x 10^3 and 2.8 x 10^3 fold respectively (For mature HtrA3 V_{max} = 1.95±0.20 x 10^{-9} M.s^{-1}, k_{cat} = 0.56±0.07 x 10^{-3} s^{-1} and $k_{cat}/K_{0.5}$= 0.42±0.03 x 10^3 $M^{-1}.s^{-1}$) [14]. However, the K_m was slightly decreased by 1.2 fold (for mature HtrA3 $K_{0.5}$ =1.31±0.06). The drastic decrease in maximum velocity, substrate turnover rates and enzyme catalysis hints toward the presence of a malformed oxyanion hole. Additionally, since β-casein is a C-terminal binding substrate, deletion of PDZ domain possibly provides greater accessibility to the active-site. This in turn is reflected in the slight increase in substrate affinity for this variant. Additionally, we reported that the cleavage of β-casein by HtrA3 N-SPD represents a sigmoidal curve having a Hill coefficient of 1.53±0.03 that is less than that of the mature HtrA3 [14]. The absence of PDZ also results in significant decrease in the catalytic efficiency, which might be due to a malformed oxyanion hole as observed previously in the case of its homolog HtrA2[6].

Peptide docking to β-casein peptide (AMVKLYC) [14] with HtrA3 N-SPD at the allosteric site [14]resulted in a low docking score of -5.194 kcal/mole. Moreover, the comparative molecular dynamics (MD) simulation study of the unbound and bound HtrA3 N-SPD showed no significant reorientation of the oxyanion hole residues to form a pocket conducive for substrate cleavage. These observations provide a solid basis for the decrease in catalytic efficiency for this HtrA3 variant. Furthermore, exhibition of residual allostery in this ΔPDZ variant might be due to the binding of the β-casein peptide to a part of the allosteric pocket encompassing the residues in the SPD domain.

The conserved GLGF motif in HtrA3 does not participate in substrate binding and allostery

'G-Φ-G-Φ motif' is a highly conserved canonical binding site in PDZ domains where Φ represents hydrophobic residues. The Gly residue in the first position is variable, whereas the second and fourth residues are hydrophobic. YIGV represents this sequence in HtrA2, while FIGI stands for the same in HtrA3 [3, 5]. It was reported that the FIGI groove mutant (FIAI or HtrA3 G358A), cleaved beta casein indicating that this groove may not be necessary for initial substrate binding unlike other HtrAs. Additionally we wanted to understand if this groove played any role in allosteric regulation of HtrA3 activity. The steady-state kinetic parameters of HtrA3 G358A showed cooperativity and substrate binding affinity (Hill constant of 2.93±0.006 and 1.07±0.32 μM respectively) comparable with mature HtrA3.

Overall, our findings highlight that the FIGI groove may not be necessary for initial substrate binding and allostery in HtrA3. We further believe that the novel alternate pocket (SBP_HtrA3), as shown by our *in silico* studies, may be the site for substrate binding.

Summary:

Our study focuses on delineating the mode of activation of HtrA3 as well as understanding the regulatory switch that drives HtrA3 activation**.** *For the first time,* we have underlined the roles of both the N-terminus and the PDZ in holding the structure in a stable oligomeric ensemble and proper active-site formation. This study also reports allostery in modulating HtrA3 enzyme activation. Allosteric modulators have the potential to develop into important drug targets owing to specificity, diversity and less toxicity. Therefore, our findings, besides providing new insights into its structure-function coordination, also highlight an excellent approach to regulate HtrA3 activity efficiently in strategies devised against diseases it is associated with.

Objective 2: Identifying and validating a putative binding partner of HtrA2

Identification of putative binding partner of HtrA2 through *in-silico* studies

The binding motif of the known binding partners of HtrA2 (based on extensive literature review) was taken into consideration and a peptide library was generated by substituting each position of the sequence with all possible amino acids. A peptide library containing 2 crore 41 lacs 41 thousand peptide sequences was generated and putative 30 sequences were filtered out. Standalone pBLAST was used to look for all possible binding peptide patterns and isolate homologs from the human proteome. Using a set of filtering parameters such as the E-value (Expect value), sub-cellular localization etc. search was narrowed down to a few proteins from which we selected DUSP9 (Dual specificity protein phosphatase 9) for validation through wet lab studies. The E-value is indicative of the error quotient in the hits obtained and a lower E-value is desirable. Since the crystal structure of the full length protein was not available, homology modeling of DUSP9 was used to generate a full length model. This model was assessed using ProSA and SAVES server. Using an *in silico* tool/server Cluspro 2.0, a blind docking approach was used to dock HtrA2 and DUSP9 to identify the critical interface residues.

Protein expression and purification

We obtained the DUSP9 clone in pBLUESCRIPT mammalian expression vector from Origene and subcloned into pGEX2T, pMALc5E, pET28a, p3xFLAG-CMV-10 vectors with GST (Glutathione S transferase), MBP (Maltose binding protein), His (Histidine) and FLAG tags respectively. The MBP-tagged DUSP9 was expressed in RosettapLysS. Similarly, GST and His tagged DUSP9 were expressed in RosettapLysS cells. For reverse pull down studies, His-tagged DUSP9 was expressed in RosettapLysS cells. The FLAG tagged construct was over-expressed in HEK 293 cells for Co-IP.

Protein-protein interaction studies

Pull down studies were performed with cell lysate expressing MBP-DUSP9 and purified HtrA2 S174A (active site mutant). This experimental setup was applied to the GST-DUSP9 construct as well. Our data showed positive interaction between HtrA2 and DUSP9. Western blot was performed to further confirm these initial findings. The positive results obtained goes on to highlight that we successfully identified a novel binding partner of HtrA2 based on the *in-silico* approach designed by our lab. To confirm these interactions in a cellular background, Co-IP was performed by using DUSP9 FLAG-tagged construct as the bait and endogenous HtrA2 as the prey.

Characterizing the binding interface

We identified 4 patches of residues each from the interacting interface of both HtrA2 and DUSP9. All these mutants were generated and pull down studies were done with a few mutants so as to determine the residues crucial for interaction. Cleavage assays with wild type HtrA2 and DUSP9 was done and the products of the proteolytic cleavage were isolated for further analysis using N-term sequencing to identify the sites of cleavage.

Summary:

We identified DUSP9 as a putative binding partner of HtrA2 through *in silico* studies and validated the interaction using *in vitro* studies. The intricacies of the interaction are currently being studied.

Significance of the study:

1. This study *is the first report* on the allosteric behaviour of HtrA3 on substrate binding. Allostery, also observed in its homologs, may be instrumental in helping execute the specific functions it performs in a cell.
2. *For the first time,* we have underlined the roles of both the N-terminus and the PDZ in holding the structure in a stable oligomeric ensemble and proper active-site formation and provided a comprehensive illustration of the mode of HtrA3 activation. This opens up possibilities of designing appropriate peptidomimetics or small molecule analogs to favourably manipulate HtrA3 functions for disease intervention.
3. We also identified DUSP9 as a novel binding partner of HtrA2 and validated this interaction. Understanding this interaction might help elucidate the significance of this interaction in the cellular milieu.

Publications in Refereed Journal:

a. Published:

- **Saujanya Acharya**, Shubhankar Dutta, Sucheta Chopra & Kakoli Bose. "**Identification of a distal allosteric ligand binding pocket in HtrA3**". *Biochemical and Biophysical Research Communications, 2019 July, https://doi.org/10.1016/j.bbrc.2019.07.005*
- **Saujanya Acharya**, Shubhankar Dutta, Sucheta Chopra & Kakoli Bose. "**A distinct concerted mechanism of structural dynamism defines activity of human serine protease HtrA3**". *Biochemical Journal, 2020 January, https://doi.org/10.1042/BCJ20190706*

Book/Book Chapter:

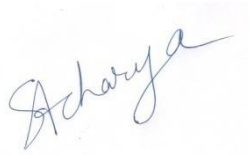
- **Acharya S**, Kuppili RR, Chaganti LK**,** Bose K. "**Proteases in Apoptosis: Protocols and Methods**". *Springer*, 2017, DOI: 10.1007/978-3-319-19497-4

Conference/Symposium

Poster Presentations:

- ***1st Global Cancer Genomics Consortium Symposium (GCGC).*** November 19-21, Advanced Centre for Treatment Research and Education in Cancer (ACTREC), Navi Mumbai
- ***International Symposium on Conceptual Advances in Cellular Homeostasis Regulated by Proteases and Chaperones the Present, the Future and Impact on Human Diseases*.** Organized by ACTREC, Tata Memorial Centre, Kharghar, Navi-Mumbai, and Maharashtra, India. December 3-6, 2013
- **Biophysics-Pashchim V** meeting organized by Tata Institute of Fundamental Research, Mumbai, India
- **Biophysics-Pashchim VI** meeting organized by Tata Institute of Fundamental Research, Mumbai, India
- ***A Conference of new ideas in Cancer- Challenging dogmas***. TMC Platinum Jubilee Celebrations. 26th February to 28th February 2016, Mumbai, India.
- ***Indo-US conference on Advances in Enzymology: Implications in Health, Disease and Therapeutics*** held at TMC-ACTREC on 15th- 19th January 2017. Second best poster award
- **31st Annual Symposium of "The Protein Society"** at Montreal, Canada, 24th - 27th July, 2017.

- ***Indo-US Conference on Sculpting the Future of Medicine-Gateway to Post Proteogenenomic Era*** held at TMC-ACTREC on 8th – 11th December 2018.
- **All National Research Scholar's meet (NRSM)** held at ACTREC, Mumbai, held annually since 2012 to 2018



Signature of Student Date: 31/07/2019

Doctoral Committee:

S. No.	Name	Designation	Signature	Date
1.	Dr. Rajiv Sarin	Chairman		1.8.19
2.	Dr. Kakoli Bose	Guide & Convener		2/8/19
3.	Dr. Neelam Shirsat	Member		2/8/19
4.	Dr. Sorab Dalal	Member		2/8/19
5.	Dr. Samir Maji	Member		2/8/19

Forwarded Through:

6/8/19

Dr. S.V. Chiplunkar
Chairperson,
Academic &Training Program, ACTREC

Prof. Dr. S. V. Chiplunkar)
Chairperson, Academic &
Training Programme , ACTREC

7/8/19

Dr. S. D. Banavali
Dean, Academics
T.M.C.

PROF. S. D. BANAVALI, MD
DEAN (ACADEMICS)
TATA MEMORIAL CENTRE
MUMBAI – 400 012.

List of figures

List of tables

CHAPTER 1

Introduction

1. Introduction

The balance between cell proliferation and cell death is carried out by an orderly multistep process called apoptosis (or programmed cell death) [15-17]. It plays a critical role in the development; homeostasis of normal tissues and at the same time is crucial for purging ageing and impaired cells [16]. Disruption of this fine balance is instrumental in the onset of many diseases. Inhibition of apoptosis is a hallmark of cancer and autoimmune diseases, whereas uncontrolled cell death promotes several neurodegenerative disorders such as Parkinson's and Alzheimer's [18-20].

Apoptosis involves two major apoptotic pathways: intrinsic (or mitochondrial mediated) and extrinsic (or death receptors mediated) where activation of caspases is central to both the pathways [21, 22]. Though caspases were believed to be the key effectors responsible for apoptosis, research over the years clearly indicate that cell death can occur in a caspase independent manner as well [21, 23]. Recently, HtrAs (high temperature requirement protease A), have emerged as key players in the apoptotic pathway due to their ability to initiate a proteolytic cascade in apoptosis [24-29].

HtrA family proteins are a family of unique homo-oligomeric and ATP-independent serine proteases (S1, chymotrypsin family) that are conserved from prokaryotes to humans. They are described as potential modulators of programmed cell death, mitochondrial homeostasis, unfolded protein response and cell growth [1]. Given their important biological roles, human HtrAs have also been implicated in pathogenicity such as arthritis, cancer and neurodegenerative disorders [30-39]. Till date, four human HtrA homologs (HtrA1-HtrA4) have been identified with strong evidence of their involvement in pro-apoptotic activity [13, 40].

This thesis focuses mainly on HtrA3, one of the lesser known members of the human HtrAs and also addresses a few key questions yet unanswered for HtrA2.

HtrA3, a mitochondrial pro-apoptotic serine protease, was first identified as a pregnancy related serine protease (PRSP) and plays an important role in regulating trophoblast invasion during placentation [11, 41-43]. Research in the past few years has linked this protein to cancer development due to its involvement in apoptosis and cell signaling [7, 9, 44, 45]. HtrA3 is synthesized as a 453 amino acid precursor protein, localized in the inter-membrane space (IMS) of mitochondria [2, 7]. Upon apoptotic stimulation, the protein undergoes a process called 'maturation' where a portion of the N-terminal region gets cleaved and the 36 kDa mature protein is released from mitochondria into the cytosol [46]. Domain wise, HtrA3 consists of a signal peptide (1-17 residues), an IGFBP N-terminal (21-77 residues), a Kazal-like (64-128 residues) domain, a serine protease domain (175-340 residues), and one PDZ (359-444 residues) domain. The active-site residue Ser305 along with His191 and Asp227 form the catalytic triad in this protease [1].

The crystal structure of the substrate-unbound HtrA3 (3.27 Å) represents a trimeric pyramidal architecture where the N-terminal region of each chain come together to form the top and PDZ domains constitute the base of the pyramid [2, 47]. The core serine protease domains (SPD) house the orthosteric pocket which is buried 33 Å above the base of the pyramid. Moreover, the structure highlights three phenylalanine residues (F140, F142 and F255) from each monomeric chain that 'lock' together and might be crucial for stabilization of the trimer. Despite providing an overview of HtrA3 spatial organization, the crystal structure does not satisfactorily define the important loops, especially the regulatory loop LA and the specificity/substrate binding loop L3

due to low electron density. Therefore, this substrate-unbound structure fails to decipher the structural reorganizations and intrinsic conformational dynamics that lead to HtrA3 activation.

Dysfunction of HtrA3 has been correlated with pre-eclampsia and oncogenesis and it has recently emerged as a potential tumor suppressor [7, 12, 44]. Given its role as a potential therapeutic target, it becomes imperative that its mode of activation and regulation be elucidated in detail to be able to modulate its functions with desired effects for disease intervention.

Here, we performed biochemical, biophysical, functional enzymolozy and *in silico* studies of 'mature' HtrA3, its different domains (individually and in various combinations) as well as its mutants to delineate the contribution of each structural component of the enzyme in defining oligomerization state, stability, substrate specificity and allosteric properties. Our studies highlight the importance of N-terminal region, PDZ domain as well as the phenylalanine 'lock' residues in oligomerization and protease activity. They also provide a detailed understanding of the regulatory switch driving HtrA3 activation with the prospect of exploiting this mechanism for devising therapeutic strategies against diseases it is associated with.

HtrA2, an extensively studied member of the HtrA protease family is also a mitochondrial pro-apoptotic serine protease with multitasking ability. Just like its homolog HtrA3, it is also involved in several critical biological functions such as protein quality control, unfolded protein response (UPR), cell growth, apoptosis, arthritis, cancer and neurodegenerative disorders [26, 32, 33, 48, 49]. It is expressed as a 49 kDa pro-enzyme in the mitochondrial membrane. Upon maturation, the N-terminal 133 residues get cleaved and the 36 kDa mature protease is released from mitochondria into the cytosol where it promotes apoptosis by binding (via the N-terminal AVPS motif) and cleaving inhibitor of apoptosis (IAP) proteins [3, 28]. This relieves the inhibitory effect of IAPs on caspases and triggers the onset of apoptosis. Besides, HtrA2 is also

known to induce the caspase independent apoptotic pathway by binding and regulating many anti-apoptotic proteins such as FLIP, Ped/Pea15 and Hax1 which may or may not involve its serine protease activity [50-53].

Though there is quite some information on its pro-apoptotic property, its complex mechanism of action is still not completely understood due to the limited number of known natural substrates/ binding partners [54]. Therefore to gain new insights into HtrA2 function, we performed a comprehensive *in silico* study to identify a putative binding partner/substrate of HtrA2 and Dual specificity phosphatase 9 (DUSP9) was identified using this approach. To validate this finding through wet lab studies, we performed various interaction and enzymatic studies to specifically understand HtrA2-DUSP9 interaction. These studies also helped understand the substrate recognition and specificity of HtrA2 in addition to shedding some light on its not so well known biological functions.

CHAPTER 2

Review of literature

2.1 Programmed cell death or Apoptosis

Apoptosis or "dropping off" (*Greek)* is a highly selective, metabolic, genetically encoded and evolutionarily selected death pathway [15, 55, 56]. It is a consequence of a series of precisely regulated events that include shrinkage of cell, DNA fragmentation, blebbing of the plasma membrane, condensation of chromatin and appearance of intracellular inclusions or 'apoptotic bodies' [16, 57]. These bodies are eventually rapidly engulfed and removed by phagocytic cells that recognize "eat-me" signals displayed on the outer surface of the apoptotic cell. Since it is a highly selective process involving a series of precisely regulated events it remains one of the most investigated processes in biological research [58]. This is also the reason why it still remains the most desirable target mechanism for the induction of cell death in tumor cells without damage to normal cells [57-60]. Deregulation of apoptosis is also responsible for several other pathological conditions including autoimmune diseases, neurodegenerative disorders, stroke, acquired immunodeficiency syndrome (AIDS) and pre-eclampsia [20, 61, 62].

The classical apoptotic pathway include two major pathways, extrinsic and intrinsic cell death path way. While, the extrinsic pathway is mainly cytoplasmic and involves death receptors, the intrinsic pathway originates in the mitochondria [17, 22, 58, 63]. Both pathways converge to activate a cascade of caspases leading to specific cell death.

2.1.1 The apoptotic pathways

The intrinsic cell death pathway

The intrinsic mitochondrial pathway is dependent on the mitochondrial membrane permeabilization (MMP) which can affect both the inner and outer mitochondrial membranes. It is characterized by the rapid release of cytochrome c form the mitochondrial intermembrane space into the cytosol [17, 29].

Many stress inducing stimuli, pro-apoptotic molecules and secondary messengers act on mitochondria to induce membrane permeability know as upstream regulators of intrinsic pathway [64]. Stimuli for intrinsic pathway include calcium overload, endoplasmic reticulum stress, DNA damage, altered redox potential, growth factor deprivation, oncoproteins, viral virulence factors and chemotherapeutic agents [22]. Regulatory molecules in this pathway are the BCL-2 proteins, which regulate mitochondrial membrane potential (MOMP). The BCL-2 proteins are characterized as follows [65, 66]:

Table 2.1 The BCL-2 family of proteins in apoptosis

BCL-2 proteins	Functional roles
BCL-2, BCL-W, BCL-xL, a-1 and MCL-1	Anti-apoptotic
BAX, BAK, and BOX	Pro-apoptotic
BH3 only proteins (BID, BIM, BAD, etc.)	Pro-apoptotic

BAX and BAK form oligomeric complexes in the outer membrane of the mitochondria resulting in membrane permeabilization. This releases cytochrome-c and other anti-apoptotic proteins such as HtrA2 and Smac/Diablo into the cytosol. Cytochrome c binds apoptotic protease-activating factor 1 (APAF1) monomer, leading to formation of heptameric wheel-like structure called the apoptosome. The apoptosome later recruits and activates initiator procaspase-9 by cleaving it (**Fig. 2.1**).

Other proteins like Smac/Diablo and HtrA2 move into the cytosol from the mitochondria and cleave 'Inhibitor of Apoptosis Proteins' (IAPs) and relieve their inhibition on caspase -9. Consecutively, active caspase-9 in turn activates other executioner caspases, caspase-3, -6 and -7

[27, 28, 64, 67]. These caspases cleave key substrates in the cell to produce many of the cellular and biochemical events of apoptosis.

The extrinsic cell death pathway

The extrinsic or death receptor pathway starts with ligands binding to the death receptors which are members of the tumor necrosis factor (TNF) receptor gene super family [52, 68]. These receptors also comprise a cytoplasmic domain of about 80 amino acids known as 'death domain'. Some of the death ligand-receptor pairs are as follows [69-71]:

Table 2.2 A list of death ligands and their receptors

Ligands	Receptors
FasL	FasR
TNF-α	TNFR1
TRAIL-L1 (TNF-related apoptosis-inducing ligand-receptor 1)	TRAIL-R1
Apo3L	DR3
Apo2L	DR4
Apo2L	DR5

The extrinsic signaling cascade is characterized with the death ligands binding their receptors resulting in receptor trimerization. These receptors then interact with the cytosolic adapter proteins such as TNF receptor-associated death domain (TRADD) and Fas-associated death domain (FADD) through homotypic interactions by their respective death domains (DD) [72, 73]. This interaction recruits upstream procaspase-8 to these preassembled adaptor molecules to form a death-inducing signaling complex (DISC) [19, 70]. Oligomerization of caspase-8 upon

DISC formation drives its auto-catalytic activation through self-cleavage. Thus, activated caspase-8 induces apoptosis by activating other downstream or executioner caspases such as caspase-3 and -7 **(Fig. 2.1)** [23].

The intrinsic and extrinsic pathways interconnect with the proapoptotic protein Bid, a member of the BH3 family of proteins [66]. Caspase-8 of the extrinsic pathway processes BH3 Interacting Domain Death Agonist (Bid) to tBid which translocates to the mitochondria and interacts with BAX and BAK, thereby promoting the release of cytochrome c from the mitochondria.

Other players of the apoptotic pathway

While, caspases are well known as the key players responsible for the proteolytic cascades in apoptosis, many studies over the years suggest that cell death can occur in a programmed fashion in absence of caspase activation as well. Some of these emerging players of apoptosis include proteases, such as cathepsins, Granzymes and HtrAs (high temperature requirement protease A), which are involved in the initiation and execution of the apoptosis [58]. These proteases are capable of triggering mitochondrial dysfunction with subsequent caspase activation and cell death.

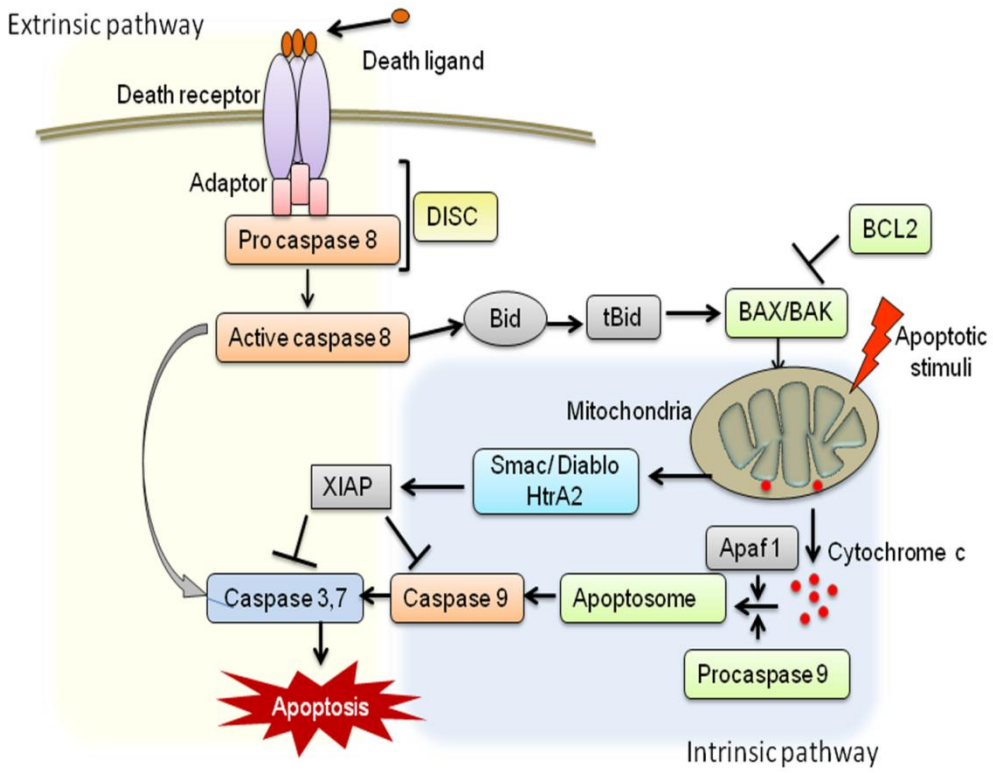


Figure 2.1 Apoptotic pathways. *A) The extrinsic apoptotic pathway initiated by the ligation of death receptors with their cognate ligands with final activation of caspases 3 and 7, leading to apoptosis. B) Intrinsic apoptotic pathway leading to release of various proteins from the mitochondrial intermembrane space (IMS) promoting caspase activation and apoptosis.*

2.2 Human HtrA proteases

HtrAs (high-temperature requirement protease A), belong to the family of S1B class of serine proteases first identified in *E. coli* as a temperature-induced, envelope-associated serine protease. While the prokaryotic HtrAs (e.g. DegP, DegQ and DegS) exhibit dual temperature-dependent chaperone-protease activity, their eukaryotic counterparts have evolved to undertake more intricate cellular functions like growth, signaling, apoptosis etc [74-77]. Human HtrA proteins have been implicated in pathogenicity such as arthritis, cancer and neurodegenerative disorders [78] . These proteases are known for their structural complexity, which also results in their functional diversity [79]. Despite diversity in functions and low sequence identity, they share a common basic structure **(Fig. 2.2)** which includes an N-terminal region, a serine protease domain (SPD) and one or more C-terminal post-synaptic density protein, Drosophila disc large tumor suppressor, zonula occludens-1 protein (PDZ) domain arranged in a pyramidal assembly ranging from trimers to large 24-mer oligomers [1, 80, 81]. The N-terminal region is variable among the family members and contains signaling as well as regulatory elements; the catalytic domain is represented by the chymotrypsin-like SPD while the PDZ is normally recruited for mediating protein-protein interactions. The variations in their structural elements across the family, along with conformational dynamics and oligomeric assembly lead to subtle differences in their functionality, specificity and pathophysiology [1]. Recently, human HtrAs have drawn attention as key effectors in multiple pathways of programmed cell death and are described as potential modulators of chemotherapy-induced cytotoxicity hence making them therapeutically important.

The four human HtrAs are: HtrA1 (L56, PRSS11), HtrA2 (Omi), HtrA3 (PRSP) and HtrA4. Among the HtrA homologs, HtrA1 and HtrA2 are well characterized, while very little information on HtrA3 and HtrA4 are currently available [1].

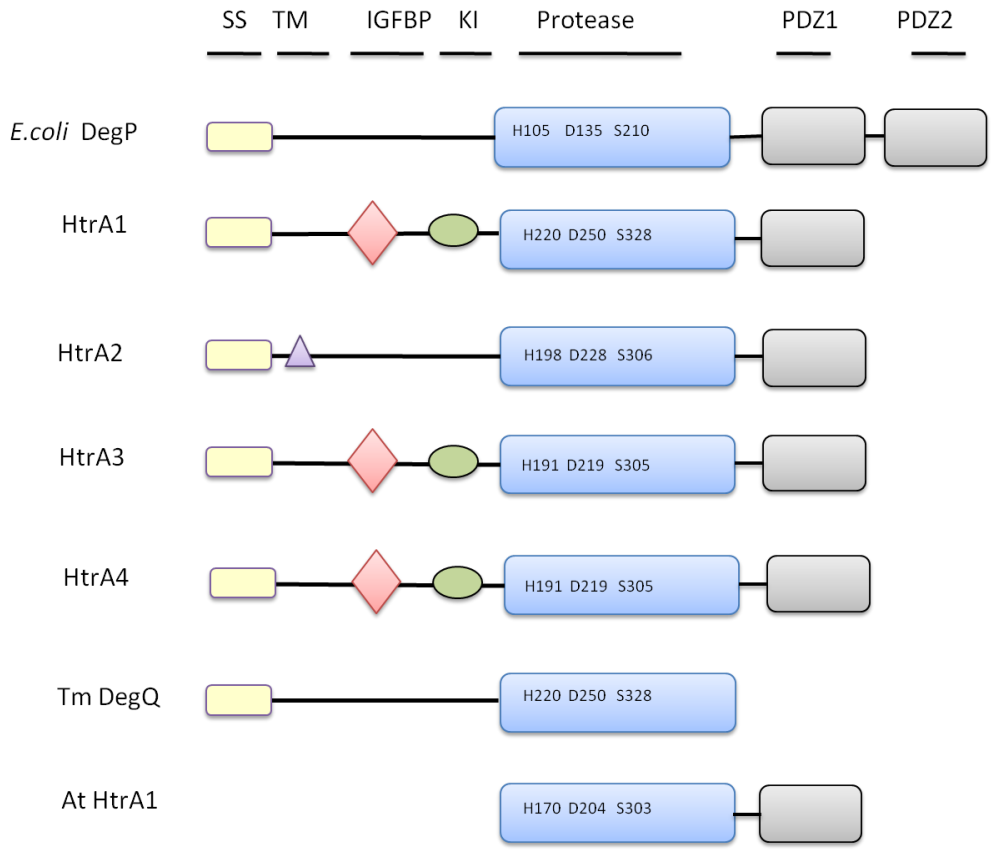


Figure 2.2 Schematic representation of domain organization of HtrA family of proteases from different species. *The protease domains are in blue rectangles, PDZ1 and PDZ2 domains are in grey rectangles, SS (signal sequence) in yellow rectangles, TM (transmembrane domain) in purple triangle, IGFBP (Insulin-like growth factor binding) in pink diamonds, and KI (Kazal protease inhibitor domain) in green oval. Here, E. coli: Escherichia coli, Tm: Thermotoga maritima, At: Arabidopsis thaliana and HtrA: high temperature requirement protease A*

General mode of activation of HtrA proteins

Catalytic triad in HtrA proteins

A catalytic triad is a collection of three amino acids that are arranged in a specific conformation in three dimensional space so as to make the active site environment conducive for substrate

catalysis [82]. In serine proteases, residues serine, histidine, and aspartate form the catalytic triad. A series of concerted non-covalent interactions among these residues occur during substrate binding and catalysis. In general, the histidine acts as a proton acceptor, thereby increasing the nucleophilicity of the active site serine [83]. On the other hand, aspartate moiety aids in this process through several hydrogen and electrostatic bonds with the histidine residue. These observations hint toward the dynamic behaviour of the active site where rearrangement of side chains of catalytic triad residues is essential for substrate binding and subsequent hydrolysis. The reactive hydroxyl (–OH) group of the serine acts as a nucleophile which attacks the carbonyl carbon of the scissile peptide bond of the substrate subsequently leading to its cleavage. The active site pocket harbours several other residues that aid in the process of substrate recognition and binding. Their physicochemical properties and stereochemical arrangements are critical toward determining substrate specificity and affinity.

Oxyanion hole in HtrA proteins

An oxyanion hole is made of positively charged residues which neutralizes the negative charge on the tetrahedral catalytic intermediate thereby promoting proteolysis by reducing energy of activation [82]. The proper formation of the oxyanion hole is extremely important for substrate catalysis. In HtrA proteins, for example HtrA2, substrate binding at the binding pocket(YIGV) leads to a series of conformational changes at the PDZ-protease interface, leading to flipping of a phenylalanine (Phe) towards histidine (His) of the catalytic triad, which is essential for proper oxyanion hole formation and optimal substrate catalysis [5, 6, 84].

2.3 HtrA3

HtrA3 or PRSP was first identified in mouse where it was observed to be uniquely regulated at the time of implantation [46]. Eventually it was also identified in humans to be involved in maintenance of ovarian function. HtrA3 gene is located on the human chromosome 14p16.1 and consists of 10 exons **(Fig. 2.3)**. The overall intron/exon structure of HtrA3 is very well conserved in evolution [41, 46].

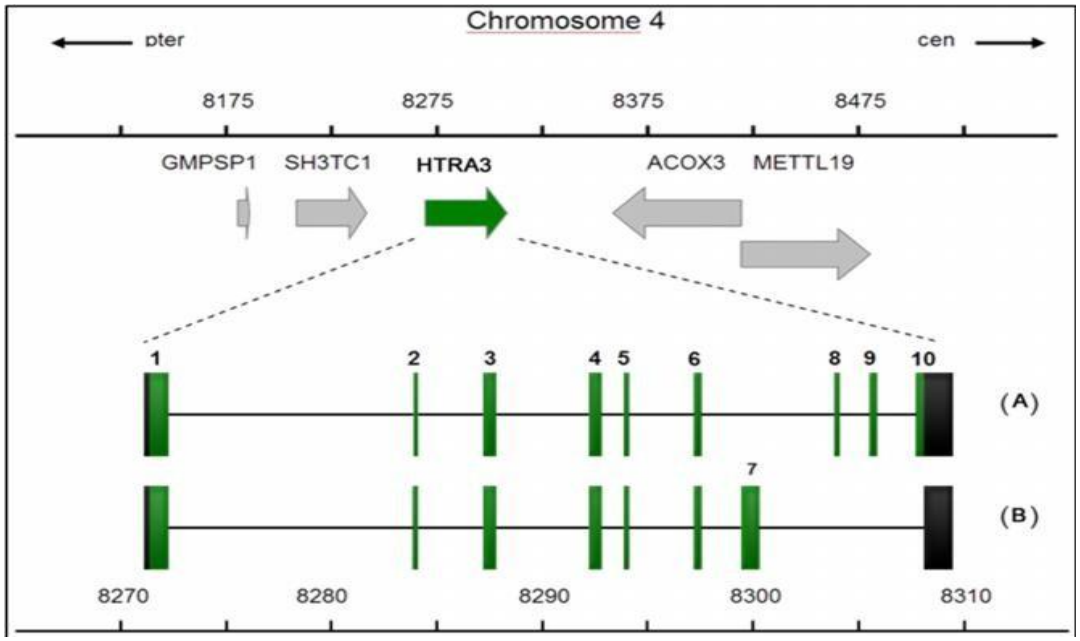


Figure 2.3 Location of Human HtrA3 on chromosome 4. *Sizes in kilobases are represented by the numbers. Exons are represented by green boxes while the black boxes represent untranslated regions. (A) indicates exons in the long isoform while (B) indicates exons in short isoform of HTRA3 mRNA.*

HtrA3 exits as two isoforms, long (HtrA3-L) and short, generated by alternative splicing. The long isoform is a product of exons 1-10 except exon 7, while the short isoform utilizes exons 1-7. The long splice variant is a 453 amino acid protein (49 kDa) coded by a ~2.8 Kb mRNA, while the short form is a 357 amino acid polypeptide (39 kDa) coded by a ~2.2 Kb mRNA **(Fig. 2.3)**. The two isoforms are identical except that the C-terminal PDZ domain is absent in the short form and is replaced by a stretch of seven amino acids (APSLAVH) which is completely different

from the corresponding (DWKKRFI) residues in HtrA3-L. Across species, HtrA3 is well conserved between human, rat and mouse. The mouse and rat mRNA have identical bases at 79% and 87% respectively while at the protein level, both show a similarity of 95% [85].

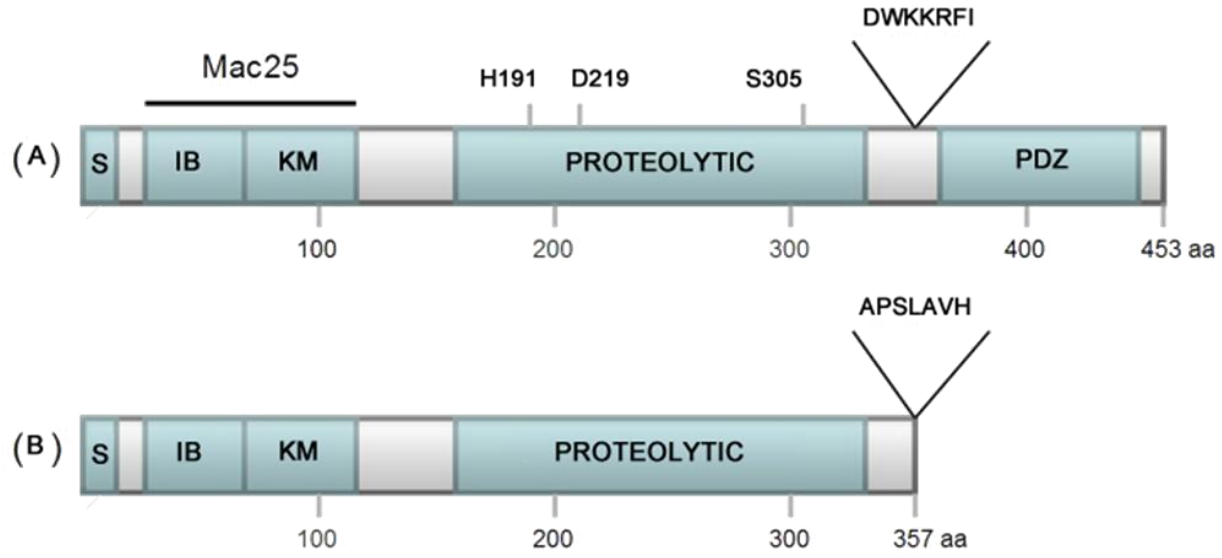


Figure 2.4 The two isoforms of human HtrA3: A) long and B) short. *S: signal peptide, IB: insulin-like growth factor binding domain, KM: Kazal type serine protease inhibitor motif, PROTEOLYTIC: trypsin-like catalytic domain, PDZ: PDZ domain, DWKKRFI/APSLAVH: amino acid residues which differ in the two HtrA3 variants, H191 D219 S305: catalytic triad residues where numbers indicate the position of the residue.*

Tissue expression of HtrA3 is seen in both adult and fetal heart, ovary, uterus placenta and bladder and is localized in the cytoplasm of these cells. It is primarily expressed in uterine endometrium and is involved in human placentation. Research over the last few years prove that deregulation of HtrA3 is often seen in many diseases including pre-eclampsia and cancers [7, 8, 43, 78]. However, till date HtrA3 has been poorly characterized from its biochemical and biophysical point of view. Given its important biological functions and it potential to be a

therapeutic agent, a detailed account of HtrA3 structure, function and regulation needs to be studied.

2.3.1 Structural organization

Domain wise, HtrA3 contains an N-terminal region, an insulin/insulin-like growth factor binding domain (IGFBP), a Kazal-like protease-inhibitor domain (KI), a serine protease domain (SPD) (175-340 aa), a hinge region and a PDZ domain (359-444 aa). The short signal peptide, IGFBP and KI motifs of the N-terminus are together referred to as the Mac25 (21-128 aa) and is required for the localization of protein in the mitochondria. The presence of IGFBP motif suggests that it can be secreted and may be involved in the regulation of the IGF system [41] [86]. The role of the SPD is to cleave various target proteins, while PDZ domain mediates protein-protein interactions and binds preferentially to the C-terminus of target proteins and stabilizes the interaction. It is also speculated to be involved in regulating enzyme activity; however there are hardly any reports that prove this hypothesis unambiguously. The regulatory loop regions are namely L1 (296-307 residues), L2 (324-327 residues), LA (163-168 residues), LB (189-202 residues) and LC (223-227 residues) [2, 47, 87].

The first available crystal structure obtained by X-ray crystallography was that of the PDZ domain of HtrA3 in complex with a peptide ligand ($FGRWV_{COOH}$). The PDZ fold comprised of five β strands (β1–β5) surrounded by two α-helices (α1, α3) at the top and a stretch of two short β-strands at the N- and C-termini (**Fig. 2.5**).

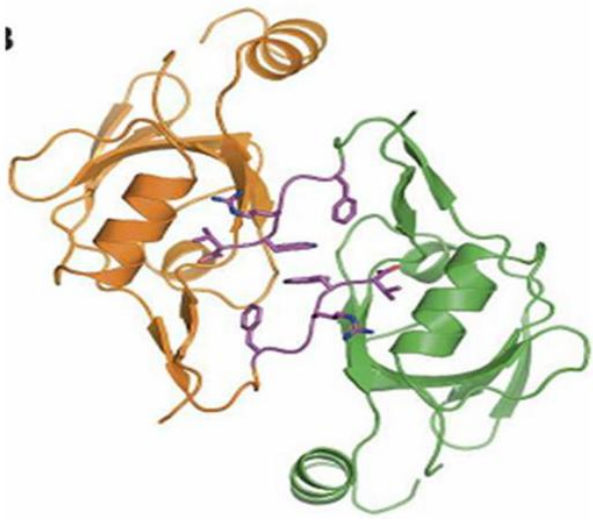


Figure 2.5 The crystal structure of HtrA3 PDZ dimer (2P3W). *The two PDZ domains are colored in orange and green respectively while the pentapeptide ligand chains bound to the PDZ domains through triglycine linkers are colored in purple.*

The structure of the full length mature protein was later solved through X-ray crystallography albeit with low resolution (3.27 Å) (**Fig. 2.6**). The structure resembles the trimeric pyramidal structures of human HtrA1 and HtrA2, however, is neither as 'open' as that of HtrA1 nor as 'closed' as that of HtrA2. The PDZ domain of chain A alone was visible in the X-ray structure, hence this chain was used to model the two other PDZ domains of chain B and C. The protein is a pyramidal trimer (109 kDa) where each monomer comprises 8 α-helices and 20 β-strands and the orthosteric pocket is buried 33Å from the base of the pyramid. The L1, LB and LC loops combine topologically to block access to the active site pocket. The trimer is also stabilized by the interaction of three phenylalanines (F140, F142, and F255) from each monomer, forming a 'lock' between monomers similar to other HtrA proteins (**Fig. 2.7**).

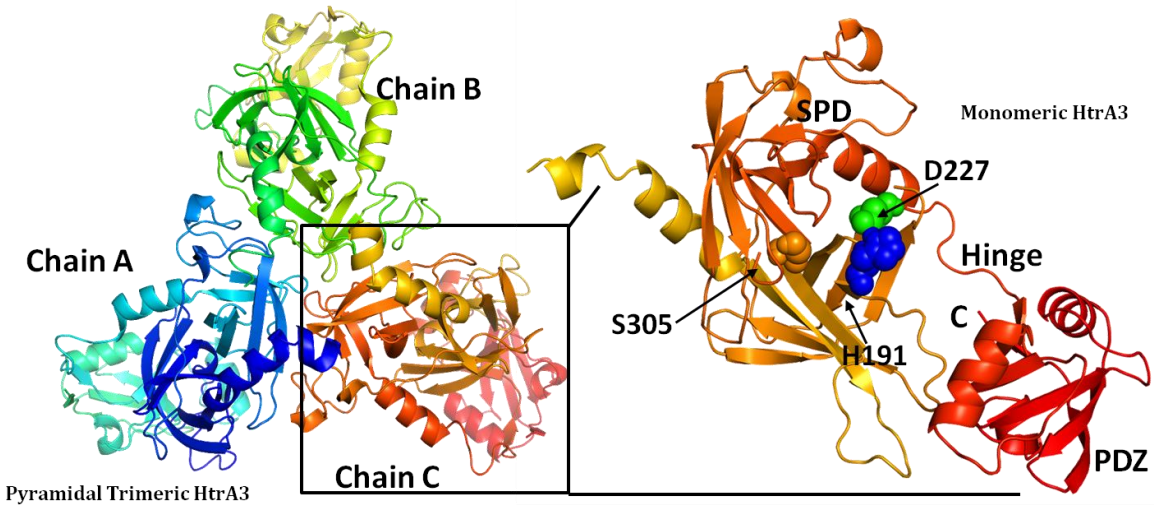


Figure 2.6 The modeled structure of HtrA3.*Using the x-ray crystal structure as the template (4RI0), an energy minimized loop-filled model of HtrA3 trimer showing domain organization for all the three chains (A,B and C) was generated. Each chain has an N-terminal region (inset, coloured gold), serine protease domain (SPD), PDZ domain and a hinge region connecting the SPD and PDZ.*

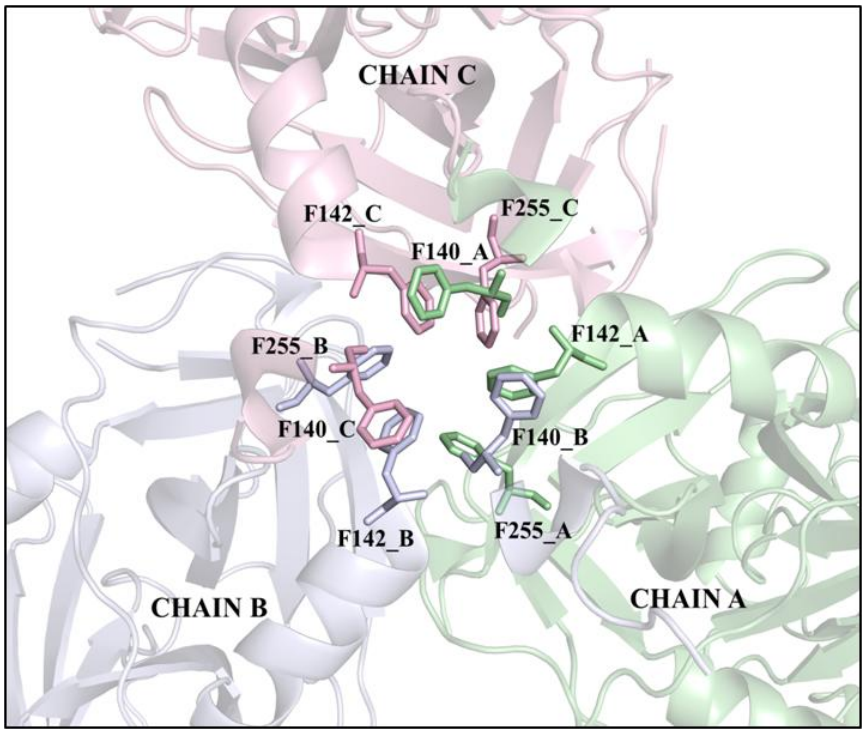


Figure 2.7 The phenylalanine 'lock' residues of the HtrA3 trimer. *The ribbon diagram of the crystal structure of HtrA3 (PDB accession number 4RIO) shows the three phenylalanine*

residues F140, F142 and F255 of each chain substituted by the amino acids as indicated in the figure in stick models. The image was generated using PyMOL.

2.3.2 Biological functions

2.3.2.1 HtrA3 negatively regulates TGF-β signaling

HtrA3 binds to members of the TGF-β protein family (TGF-β1, TGF- β2,) BMP4 (bone morphogenetic protein 4), and GDF5 (Growth differentiation factor 5), and suppresses signal transduction mediated by these extracellular cytokines [45, 88]. During embryogenesis, HtrA3 secreted by decidual cells inhibit TGF-β signaling and restrict trophoblast invasion during placental development [45]. In skeletal tissues, HtrA3 inhibits the development of articular chondrocytes by negatively regulating BMPs, GDFs and TGF-β1 signaling. However, the mechanism of TGF-β signaling inhibition by HtrA3 still remains elusive although there is some evidence that the phenomenon may depend on its proteolytic activities.

2.3.2.2 HtrA3 is associated with placental development

The levels of HtrA3 expression peaks during post implantation and placentation [11, 42, 89]. HtrA3 is secreted form decidual cells of the maternal endometrium where it negatively regulates trophoblast invasion. It was observed that on inhibition of HtrA3 activity, the trophoblast invasion significantly increases. It is believed that HtrA3 provides a chemokine environment that attracts invading extravillous trophoblast cells and controls its invasion [90].

2.3.2.3 HtrA3 and apoptosis

HtrA3 functions as a pro-apoptotic protein in the intrinsic, mitochondria-mediated apoptotic pathway [24, 58]. Upon action of apoptosis-inducing agents HtrA3 undergoes autoproteolysis resulting in removal of the N-terminal domain. Subsequently, the processed HtrA3 is released into the cytosol where it promotes apoptotic cell death. This translocation to the cytosol was observed to be accompanied by the release of pro-apoptotic factors, HtrA2 and cytochrome c,

and was correlated with a significant increase of cell death [58]. This mode of action resembles HtrA2 which also is released from mitochondria upon action of stress stimuli and promotes cell death. However, contrary to HtrA2, very little is known about molecular events following the release of HtrA3 to cytoplasm. So far it has been shown that Bcl-2 overexpression attenuates the HtrA3-mediated apoptosis. It was also recently reported to bind XIAP [44]. Therefore specific cytoplasmic substrates of HtrA3 need to be identified to understand how it regulates this pathway Apart from the classical cell death pathways, HtrA3 has also been reported to induce the non-classical cell death pathways through its proteolytic activity [78]. However, not much is known and substrates need to be identified to elucidate this process.

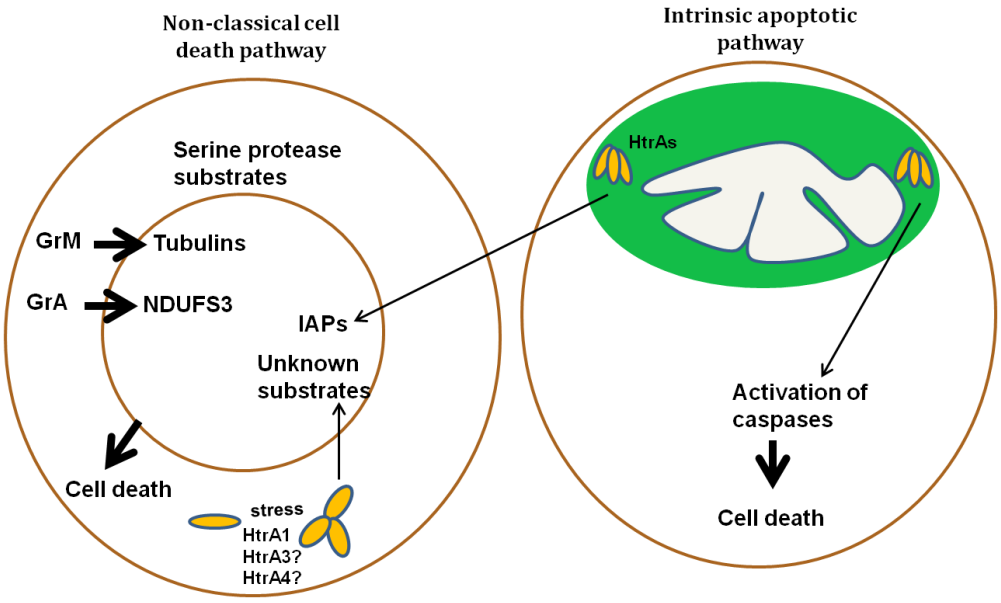


Figure 2.8 HtrA3 and apoptosis. *Model proposing the role of HtrA3 in both classical and non-classical cell death pathways. In the intrinsic pathway (right), HtrA3 is activated on apoptotic stress to undergo maturation and move to the cytosol where it relieves the inhibition of IAPs on caspases and brings about cell death. In the non-classical death pathway (left), HtrA3 is proposed to bring about cell death by cleaving a range of substrates most of which are yet unknown.*

2.3.3 Pathological involvement

2.3.3.1 HtrA3 in cancers

Due to the role HtrA3 plays in apoptosis and cell signaling, this protein has often been linked with cancer development [78, 91]. Downregulation of HtrA3 has been observed in several cancer cell lines and tumors such as ovarian, endometrial and lung cancers and it is suggested that the protease may act as a tumor suppressor (**Table 2.3**).

Table2.3: HtrAs in cancers.

Protein	Downregulated in
HtrA3	B and T cell lymphoblastic leukemia, ovarian and endometrial cancers, lung cancers

In ovarian cancer the decrease in HtrA3 expression coincided with the degree of malignancy [8, 92-94]. In endometrial cancers the low HtrA3 levels correlated with increasing grades of the tumor. It was also observed that this was accompanied with a reciprocal level of TGF-β1suggesting that HtrA3 regulated TGF-β1 signaling pathway might be the molecular basis of this type of cancer.

It has recently been shown that HtrA3 is down regulated in lung tumors from smokers and it correlates with an increased frequency of methylation within the first exon of the gene and also correlated with a poorer response to chemotherapy [7, 91]. I*n vitro* studies further confirmed that cigarette smoke induced methylation and silencing the HtrA3 gene, which could be the cause of chemoresistance in smoking-related lung cancer [12].

It has been demonstrated that HtrA3 proteolytic activity sensitizes lung cancer cells to chemotherapeutic drugs, e.g. cisplatin and etoposide [7]. Re-expression of catalytically active HtrA3 in these lung cancer cell lines lacking endogenous HtrA3 expression suppressed tumor cell invasiveness and induced death on administering these cytotoxic agents.

HtrA3 was also shown to be up-regulated in some other cancers, including seminoma and several hematologic malignancies. However, there is no functional explanation of this selective behaviour in these cancers till date [77].

2.3.3.2 HtrA3 in arthritis

Regulation of TGF-β signaling is required to keep articular chondrocytes in their normal undifferentiated state. It was reported that the expression of HtrA3 increases substantially in articular cartilage and suppresses TGF-β pathway which probably leads to terminal differentiation of chondrocytes leading to osteoarthritis like phenotype in humans [88, 95]. This regulation of the signaling pathway may be dependent on the proteolytic activity of this protein.

2.3.3.3 HtrA3 in pre-eclampsia

Several studies have revealed the involvement of HtrA3 in embryo implantation and development of placenta. Adequate invasion of the maternal decidua by extravillous trophoblast is of critical importance for implantation and placentation. Interstitial extravillous trophoblast cells (EVTs) migrate, invading deeper layers of decidua and are important for proper formation of the decidual arteries and for sufficient blood flow into the placenta. Defective trophoblast invasion results in poor maternal-fetal circulation and placental hypoxia while excessive invasion of trophoblasts into the endometrium is associated with invasive choriocarcinoma [42, 96, 97]. Therefore, the invasion of EVTs is tightly regulated to ensure a successful pregnancy. It has been shown that HtrA3 expression increases dramatically in the endometrial glands and decidual cells when the endometrium prepares for implantation. Though HtrA3 is expressed in most trophoblast cells, which indicates that it is required for trophoblast differentiation, expression was not detected in the invading interstitial trophoblast cells suggesting that it functions as an inhibitor of trophoblast invasion [11]. More recent studies have demonstrated that the decidual

cells secrete HtrA3 and negatively regulate trophoblast invasion ensuring a proper development of the placenta. Placental HtrA3 levels are maximal in the first trimester of pregnancy, and then decrease at 13-14 weeks of gestation which is associated with an increase of placental oxygen at this stage of pregnancy [43, 89]. Abnormally high levels of maternal serum HtrA3 at 13-14 weeks of gestation were observed in pre-eclampsia and hence HtrA3 could be considered as a potential molecular marker in early diagnostics of this disorder.

2.3.4 Mechanism of activation

Like its homologs, HtrA3 has been reported to be allosterically regulated. The general mechanism in HtrAs include a series of conformational changes which is relayed by the regulatory loops (L3, LD, L1 and L2), resulting in the remodeling of active-site for efficient substrate catalysis [14]. The active site in HtrA3 has the catalytic serine residue (Ser305) positioned far from His191 for an efficient proton transfer suggesting the need for similar structural reorientations. Conformational dynamics analyses show that upon substrate binding, the regulatory loops (L1, L2, and L3) of HtrA3 undergo conformational changes that result in a catalytically favorable active-site. The PDZ domain is also believed to play an important role in mediating these changes.

In terms of substrate specificity, HtrA3 protease is similar to HtrA1 and HtrA2 and preferentially cleaves peptide bonds formed by the carboxyl groups of hydrophobic, aliphatic amino acids. Its activity is highly stimulated by temperature as is typically seen in temperature dependent HtrA proteins.

2.4 HtrA2

Human HtrA2/Omi is a mitochondrial serine protease which is produced as a 49 kDa (458 amino acids) proenzyme in the mitochondrial intermembrane space and has the basic function of a

proteolytic enzyme that removes damaged and denatured proteins at elevated temperatures (**Fig. 2.9**) [3, 58]. The proenzyme is processed in the mitochondria to its mature form through the removal of the first 133 amino acids, after which it translocates to the cytoplasm with an exposed 'AVPS' tetra-peptide motif. HtrA2 is normally not secreted and resides in the IMS of the mitochondria under normal conditions. This mammalian enzyme shares sequence and structural homology with the Escherichia coli serine endoprotease HtrA/DegP. The proteolytic activity of recombinant HtrA2 is significantly elevated upon heat shock [98].

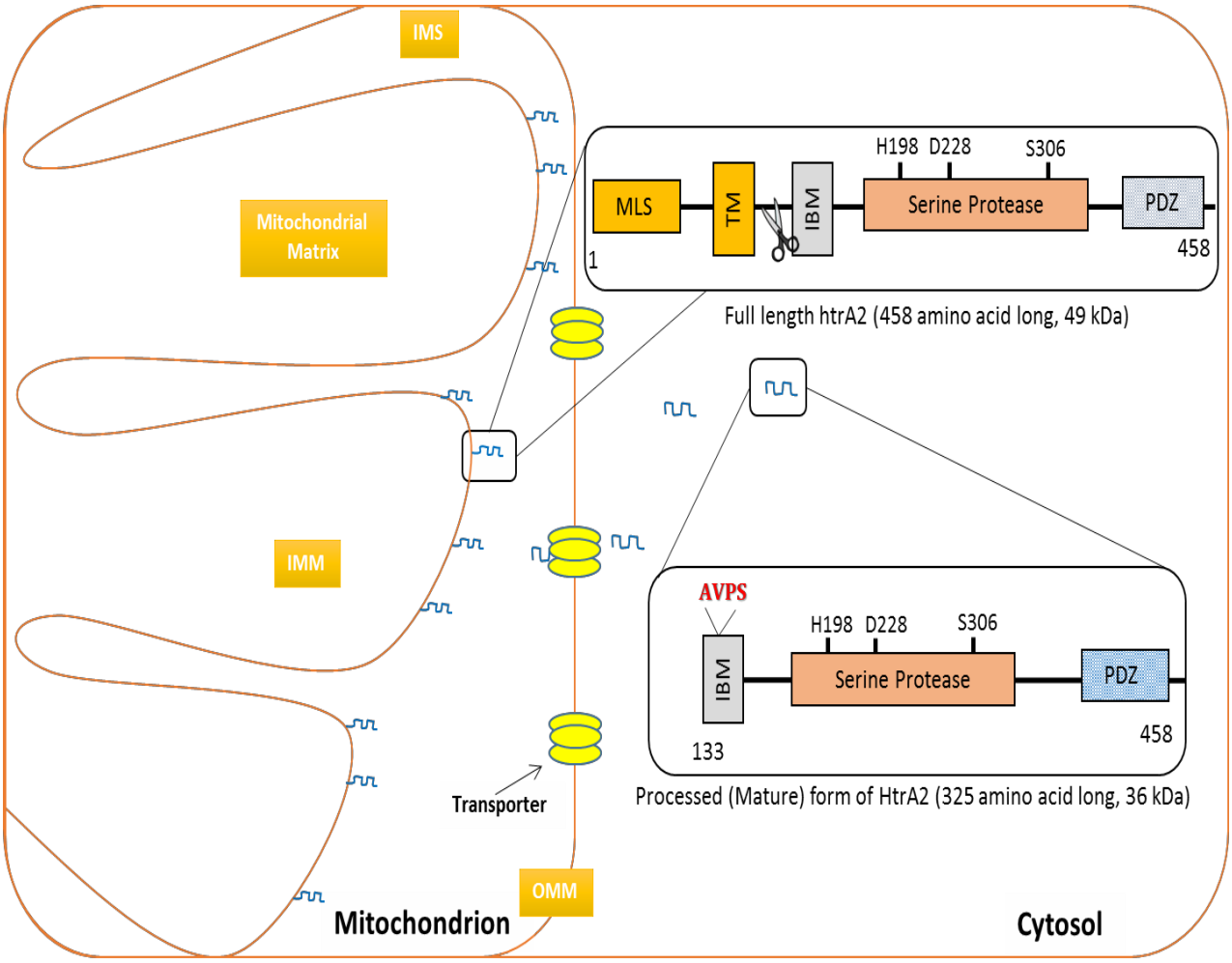


Figure 2.9 ***Schematic representation of full length HtrA2 highlighting its maturation process.*** *Full length HtrA2 (1-458 aa) comprising a mitochondrial localization signal and a*

transmembrane domain (1-133 aa) is localized in the mitochondria. Maturation occurs as a result of cellular stress that removes first 133 amino acids and exposes a tetra peptide IBM motif (AVPS) in its N-terminus. The catalytic triad comprising His 198, D228 and S306 has been shown in the figure Abbreviations: MLS Mitochondrial localization sequence; TM Transmembrane domain; IBM IAP binding motif; PDZ Post synaptic of 95kDa, Disc large, Zonula occludens; aa, amino acid. The full length HtrA2 has been shown as light blue 'coil-like" structure attached to mitochondrial inner membrane.

2.4.1 Structural organization

HtrA2 is coded by a gene with eight exons and has a chromosomal location of 2p13.1 ***(Fig. 2.10).***

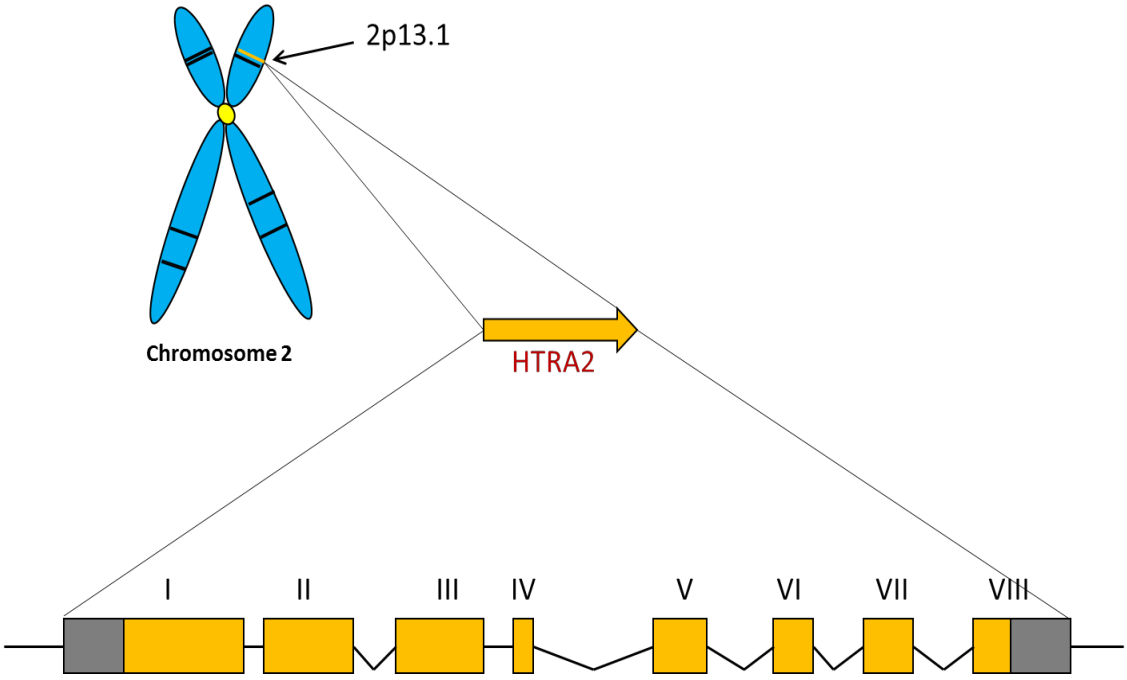


Figure 2.10 Localization and schematic organization of the HTRA2 gene. *The HTRA2 gene is present on chromosome 2p13.1. The magnified view of the location shows the exons marked I to VIII (orange).*

The X- ray crystallographic structure of mature, inactive, substrate unbound form of HtrA2 has a resolution of 2.1Å (**PDB: 1LCY**) [3, 6]. Like its homologs, it has a trimeric pyramidal architecture with the N-terminal regions of each monomeric chain forming the top and PDZ

domains forming the base of the pyramid. The serine protease domains reside 25 Å above the base of the pyramid. Each chain comprises 7 α-helices and 19 β-strands and contains several loops, which are named according to the chymotrypsin nomenclature- LA (residues 170-174), L1 (302-306), L2 (323-329), L3 (275-295), and LD (259-266) (**Fig. 2.11**).

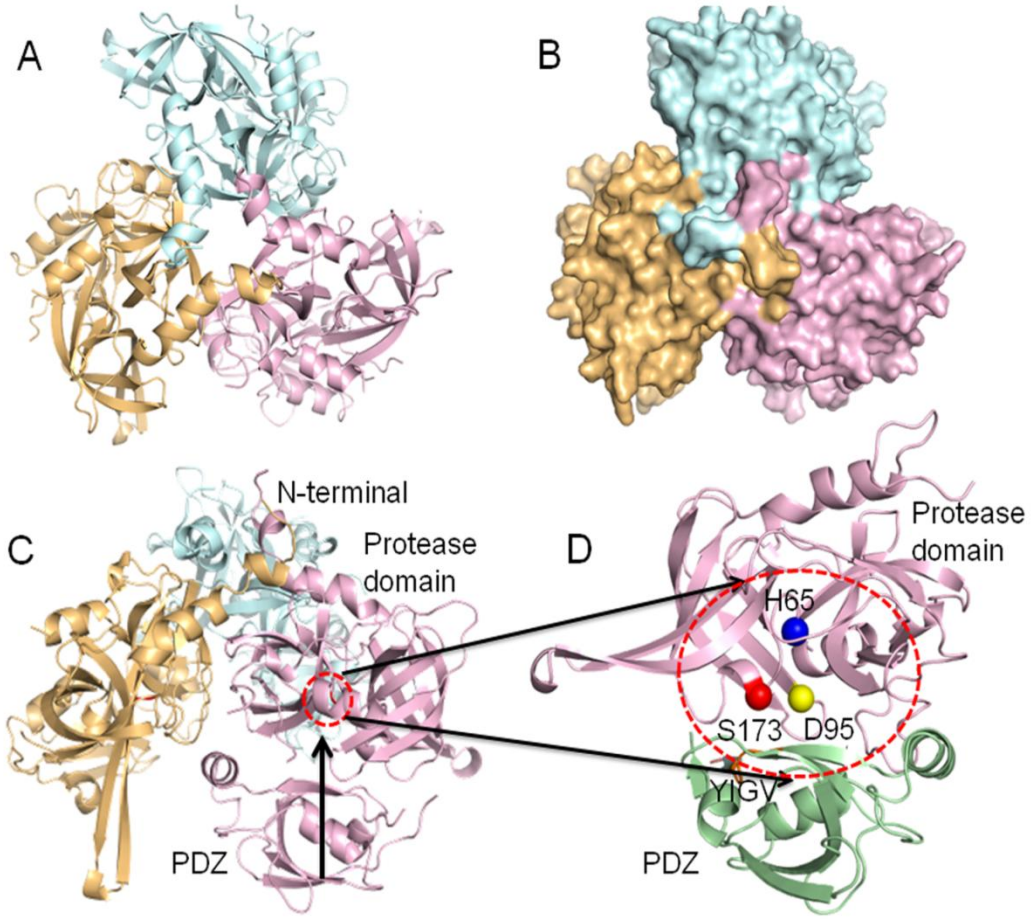


Figure 2.11 Crystal structure of HtrA2. *A) Schematic representation of HtrA2 trimer (PDB entry 1LCY). Each monomeric subunit is represented in blue, orange and pink. B) Surface representation of the HtrA2 structure. C) side view of trimeric HtrA2. Trimerization is mediated exclusively by the serine protease domain. The N-terminal IAP-binding tetrapeptide motif is located at the top of the pyramid, and the PDZ domain is at the base. D) Cartoon representation of HtrA2 monomer. Serine protease and PDZ domains are colored in pink and green respectively. The position of catalytic triad residues: H65 (blue), D95 (yellow) and S173 (red) are shown as spheres. The position of canonical peptide binding groove 'YIGV' is represented in orange. The figures are generated using PyMOL (DeLano Scientific, USA).*

The active site pocket and the catalytic triad residues (Ser306, His198 and Asp228) remain buried in the hydrophobic core of the serine protease domain among the regulatory loops. The trimeric assembly is mainly held together by the N-terminal regions of each monomer especially through the aromatic residues Tyr147, Phe149 and Phe256.

The G-Φ-G-Φ motif, where Φ denotes hydrophobic residues, is the canonical substrate binding motif in PDZ domains [4, 5]. This motif is represented by YIGV in HtrA2, and resides buried in the intimate interface between the PDZ and the protease domains. The PDZ and the protease domains make numerous contacts with each other through van der Waals interaction. The hydrophobic residues on strands β11 and β12 of the protease domain interact with the hydrophobic residues from strand β14 and helix α5 of the PDZ domain. This leads to the formation of a 'closed' conformation where the PDZ binding groove remains unavailable and also blocks passage of substrate molecules to the active site pocket.

2.4.2 Mechanism of activation

Allostery is a characteristic feature of many multidomain enzymes where an active functional conformation is achieved through ligand binding at a site distal to the catalytic pocket. There are two conformational states, 'T' (inactive or tensed) and 'R' (active or relaxed) which are always in a dynamic equilibrium. Binding of an allosteric activator changes the structure from 'T' to 'R'; while, an allosteric inhibitor does the reverse. The first reports on HtrA2 activation suggests that it has a complex allosteric mechanism of activation, according to which the 'YIGV' groove of PDZ interacts with substrate/modulator and induces a significant conformational change at the PDZ-protease interface which then opens up the active-site for further substrate catalysis [42]. However, recent reports have-re defined this activation model suggesting that a dual-regulatory switch is instrumental in HtrA2 activation. The complex allosteric mechanism involves a series

of conformational changes on ligand binding at the distal substrate binding pocket which is also accompanied by inter-molecular PDZ-protease movement for proper active-site (by allowing access to the catalytic site) and oxyanion hole formation that subsequently leads to further substrate catalysis. Further research in this field provides evidence of a distal non-canonical substrate binding site where binding causes the signal to be relayed with subsequent opening up of the YIGV [4-6]. The allosteric signal propagated by PDZ domain might lead to a disorder-to-order transition of important regulatory loops (L1, L2, L3, and LD) in protease domain. This intrinsic allosteric regulation mechanism is initiated by sensor loop L3 and subsequently a cascade of conformational changes occurs along L3→LD→L1/L2, enabling the remodeling and activation of the proteolytic site **(Fig. 2.12)**.

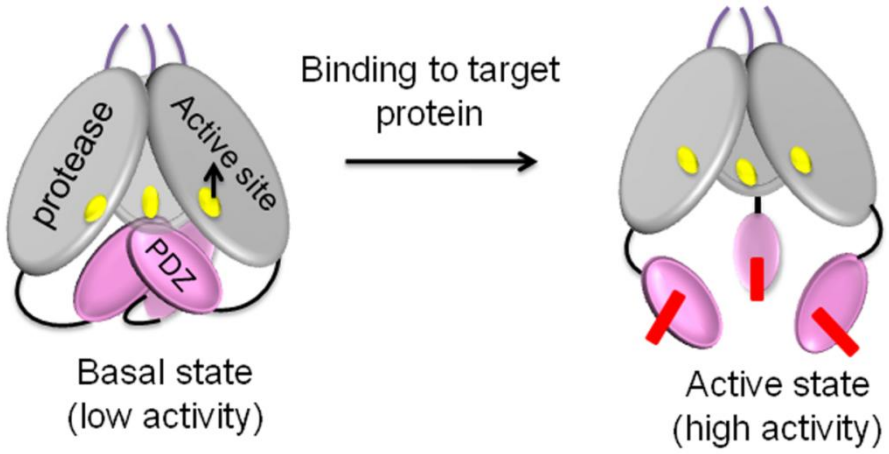


Figure 2.12 Proposed working model for HtrA2 activation. *In the basal state, the PDZ domains keep the protease activity of HtrA2 in check. Upon binding to the target proteins, the PDZ domains moves away from the protease to make the catalytic serine accessible to the substrate*

Apart from this classical model where events are coordinated by the C-terminal peptide binding, the new theory provides evidence that the N-terminal region is equally important in protease stabilization of HtrA2 trimer and catalytic activation through the N-terminal 'AVPS' tetrapeptide

motif. In this N-terminal mediated activation, protease activity of HtrA2 is also modulated by binding to IAPs, through its N-terminus. It was observed that HtrA2 incubated with XIAP showed increase proteolytic activity toward its substrates, H2-Opt peptide (synthetic substrate of HtrA2) and β-casein (generic substrate for all proteases). This phenomenon emphasizes involvement of multiple modes of HtrA2 activation and regulation, involving PDZ as well as other regions of the protein

2.4.3 Biological functions

2.4.3.1 Protein quality control

Aggregation of misfolded or damaged proteins poses a serious threat to the cell by obstructing essential biological processes. To prevent this, cells make molecular chaperones and proteases for protein quality control. HtrA2 is one such heat shock induced serine protease and alteration in its proteolytic activity leads to the accumulation of unfolded proteins in mitochondria, dysfunction of the mitochondrial respiration, generation of reactive oxygen species and also loss of mitochondrial competence [99-101]. For e.g., missense mutation of S276C, homologous deletion of the HtrA2 gene and loss of phophorylation of HtrA2 by serine/threonine kinase PINK1 lead to impairment of HtrA2 proteolytic activity and increased susceptibility to mitochondrial stress and neuronal death as a result of loss of its function in protein quality control [35, 36, 99, 101].

2.4.3.2 Apoptosis

HtrA2 regulates apoptosis through both the caspase dependent and independent pathways. The caspase dependent pathway initiates in the mitochondria while the independent pathway is dependent on its proteolytic activity [24-26, 102].

On apoptotic stimulus, HtrA2 undergoes maturation to expose the N-terminal tetrapeptide motif (AVPS) through which it interacts with the IAPs. XIAP (X-linked inhibitor of apoptosis protein), cIAP1 (cellular inhibitor of apoptosis protein-1), cIAP2, Apollon/BRUCE inhibit caspase-3, -7 and -9. Upon interaction with mature HtrA2, these inhibitors are cleaved thereby lifting the caspase inhibition and promoting the apoptotic signaling cascade through the caspases. Some reports suggest that HtrA2 PDZ may bind with a trimeric assembly of TNFR1 or Fas subsequently activating the 'death domain' present in the cytoplasmic region of each Fas monomer thereby initiating caspase-8-dependent apoptotic pathways [24, 26, 29, 50, 103].

In the caspase independent apoptotic pathways, proteolytic activity of HtrA2 was mainly involved. HS1-associated protein X-1(Hax-1) is mitochondrial antiapoptotic protein responsible for regulation of mitochondrial membrane potential. Proteolytic degradation of HAX-1 by HtrA2 induced a pro-apoptotic pathway initiated in the mitochondria without the release of HtrA2 into the cytosol [51]. Furthermore, cleavage of important cytoskeletal proteins such as actin, α-, β-tubulin, and vimentin by HtrA2 has been shown to culminate in apoptosis. Additionally, the cleavage of transcription factor p73, Bmf (an endogenous inhibitor of BAX) also leads to apoptosis through transcription of the proapoptotic BAX. Among other proteins, Wilms tumor suppressor protein (WT1), a transcriptional regulator of genes controlling apoptosis, cell growth and differentiation, when cleaved by HtrA2 causes subsequent apoptosis of human osteosarcoma cells treated with staurosporine [32, 103].

2.4.4 Pathological involvement

2.4.4.1 HtrA2 and cancers

HtrA2 shows variable expression in cancers. HtrA2 expression was observed to be down-regulated in ovarian cancer, metastatic prostate cancer and adult male germ cell tumor. In breast

cancer, HtrA2 expression was reciprocal to tumor staging. Down-regulation of HtrA2 expression reduced cell death mediated by integrin α7 (ITGA7) in prostate cancer cell lines. ITGA7 acts as a proapoptotic factor and enhances the HtrA2 protease activity both *in vitro* and *in vivo.* A similar inhibitory effect was exerted by the HtrA2 inactive variant [77, 78].

HtrA2 expression was up-regulated in lung adenocarcinoma, superficial or invasive transitional cell carcinoma of bladder, oligodendroglioma (brain) and squamous cell carcinoma of head and neck, B-cell acute lymphoblastic and T-cell lymphoblastic leukemia. HtrA2 expression was also up-regulated in Wilm's tumors compared to normal fetal kidney or clear cell sarcoma. Collectively altered expression levels of HtrA2 highlight its importance in cancer progression and provide a rationale for targeting this protein for cancer therapy

2.4.4.2 HtrA2 in neurodegenerative disorders

HtrA2 in Alzheimer's disease

Alzheimer's disease (AD) is characterized by the presence of aggregates of the amyloid β, a major element of neurotoxic deposits, in brains of AD patients. In normal conditions, HtrA2, via its chaperoning function prevents the aggregation of amyloid β-peptides (Aβ40 and Aβ42), keeping the peptide in monomeric state. HtrA2 is also involved in processing of mitochondrial amyloid precursor protein (APP), to generate a 28-kDa APP fragment that is apparently released in to the cytosol [104]. Disruption of the APP processing activity and subsequent accumulation of APP disrupts the basic functions of mitochondria and impairs energy metabolism. In a breakthrough study in mnd2 mice with a homozygous loss-of-function mutation (S276C) in the HtrA2 gene, accumulation of APP was seen leading to the death of the HtrA2 deficient neuronal cells. Presenilin-1 is a catalytic component of γ-secretase enzyme that is implicated in the

inherited forms of early onset AD [105]. It was validated through *in vitro* studies that presenilin mediates regulation of HtrA2 protease activity.

HtrA2 in Parkinson's disease

Parkinson's disease (PD) is the most common movement disorder that affects 1% of the population over the age of 65. It is characterized by an exclusive degeneration of dopamine releasing neurons of the substantia nigra pars compacta in the brain, and the presence of characteristic proteinaceous intracytoplasmic inclusions, known as Lewy-bodies in the affected brain areas. The first connection between HtrA2 dysfunction and PD came from the characterization of mnd2 mutant mice which displayed neurodegeneration, muscle wasting and death by 40 days of age [35, 36, 48, 106]. These mice showed a loss of catalytic activity of HtrA2 due to a S276C point mutation. Furthermore, studies with HtrA2 knockout mice displayed a neurodegenerative phenotype with parkinsonian features. These mice showed accumulations of unfolded proteins in mitochondria, a defective mitochondrial respiration and an enhanced production of reactive oxygen species in the brain tissue cells. Several studies suggest that defects in mitochondrial respiratory chain, impaired mitochondrial dynamics and mitochondrial trafficking, play a significant role in the mitochondrial dysfunction that takes place in neurodegenerative disorders. The loss of proteolytic activity in both mnd2 and HtrA2 knockout mice result in enhanced sensitivity to stress suggesting that this protease might play an essential role in the mitochondria-related stress signals.

2.5 Dual-specificity phosphatase 9 (DUSP9)

MAP kinases such as extracellular signal-regulated kinase (ERK), 1 c-Jun N-terminal kinase/stress-activated protein kinase (JNK/SAPK) and p38/RK/CSBP (p38) are responsible for a range of cellular responses including cell differentiation, proliferation, cell cycle regulation and

apoptosis. These MAPKs are activated by a range of growth factors as well as pro-inflammatory cytokines and cellular stress. The activated MAPKs in turn phosphorylate diverse target proteins in membrane or cytosolic fractions, nuclear transcription factors through various pathways leading to changes in cell function. The activation of MAPKs is reversible and involves phosphorylation of both tyrosine and threonine residues by dual specificity kinases. The DUSPs were shown to be involved in this process of regulating the MAPKs by dephosphorylating both threonine and tyrosine residues and rendering them inactive [107, 108].

DUSP9 gene is located on the X-chromosome and codes a 384-amino acid protein which has two distinct domains. The rhodanese-like domain at the N-terminal region serves as MAPK-binding domain which is followed by the catalytic phosphatase domain at the C-terminus **(Fig. 2.13)**. Crystal structure of the full length protein has not been solved yet; however, the structure of the catalytic domain has been solved at resolution of 2.7 Å (PDB: **2HXP**) [109, 110].

Human DUSP9 is expressed in kidney, placenta and adipose tissue with localization in the cytosolic fragment. Several studies have shown that DUSP9 modulates the insulin signaling pathway and is also implicated in cancers and placental development [111-113].

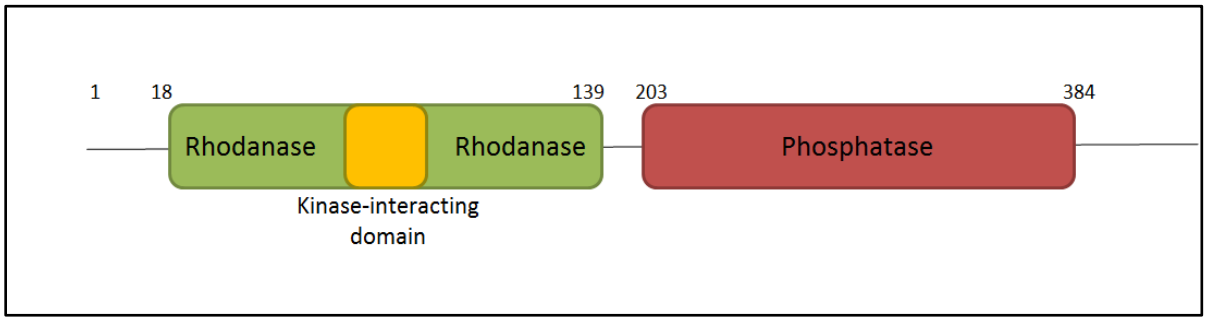


Figure 2.13 Schematic representation of DUSP9 domain organization. *DUSP9 consists of two prominent domains, the rhodanase domain which is involved in protein-protein interactions and a catalytic phosphatase domain. The numbers indicate the domain range.*

Aims and objectives

1. Characterization of structural properties and conformational stability of HtrA3 and the role of different domains in protein function and specificity

a) Role of different domains and critical residues in HtrA3 structure and function

b) Role of substrate binding and conformational dynamics in regulating mode of HtrA3 action

2. Identification and characterization of HtrA3-ligand interaction

a) Substrate specificity of HtrA3

b) Exploring XIAP as a binding partner and substrate of HtrA3

3. Identification and characterization of a novel binding partner/substrate of HtrA2

a) *In silico* prediction of a putative binding partner/substrate of HtrA2

b) Validate the findings through wet lab studies

CHAPTER 3

Materials & Methods

3.1 MATERIALS

3.1.1 Plasmids/Vectors

Bacterial vectors:

Plasmid / Vector	Tag	Antibiotic resistance	Source
pJET1.2/blunt	No tag	Ampicillin	Thermofischer Scientific (Massachusetts, USA)
pET 28a	N- and C- terminal His_6 tags	Kanamycin	Addgene (Cambridge, MA, USA)
pET 20b	C- terminal His_6 tag	Ampicillin	Addgene (Cambridge, MA, USA)
pMALc5E-TEV	N-terminal MBP tag (Maltose binding protein) with TEV protease site	Ampicillin	New England Biolabs (Ipswich, USA)
pGEX-4T	N-terminal GST tag (Glutathione S-transferase)	Ampicillin	Addgene (Cambridge, MA, USA)

Mammalian expression vectors:

Plasmid / Vector	Tag	Bacterial resistance	Source

c-Flag pcDNA3	C- terminal FLAG	Ampicillin	Addgene (Cambridge, MA, USA)
p3xFLAGCMV-10	N- terminal FLAG	Ampicillin	Addgene (Cambridge, MA, USA)

Plasmids containing gene of interest:

Plasmid	**Tag**	**Bacterial resistance**	**Source**
pDONR221-HtrA3 (cDNA in a gateway/master vector)	No tag	Kanamycin	DNASU (Arizona State University, USA)
pCDNA3-WT type HtrA2-Flag	C- terminal FLAG	Ampicillin	Dr. L.M. Martins (Leicester)
pCDNA3- HtrA2 S306A-Flag	C- terminal FLAG	Ampicillin	Dr. L.M. Martins (Leicester)
pBLUESCRIPT-DUSP9	No tag	Ampicillin	Origene (Rockville, USA)

3.1.2 Primers

Primers for gene amplification

HtrA3 variants	**Sequence (5'→3')**	**Orientation**
Mature HtrA3 (130-453)	CATATGCTCCACCAGCTGAGCAGCCCG	Forward
	GGATCCCATGACCACCTCAGGTGCGAG	Reverse
HtrA3 variant (135-453)	CATATG AGCCCGCGCTACAAGTTCAAC	Forward

	GGATCCCATGACCACCTCAGGTGCGATG	Reverse
HtrA3 variant (144-453)	CATATGGCTGACGTGGTGGAGAAGATC	Forward
	GGATCCCATGACCACCTCAGGTGCGATG	Reverse
HtrA3 N-SPD (130-340)	CATATGCTCCACCAGCTGAGCAGCCCG	Forward
	GGATCCGAACCGTGTGATGCGGTCTGAGGG	Reverse
HtrA3 SPD (175-340)	CATATGGGTTCTGGCTTCATCATGTCAG	Forward
	GGATCCGAACCGTGTGATGCGGTCTGAGGG	Reverse
HtrA3 SPD-PDZ (175-444)	CATATGGGTTCTGGCTTCATCATGTCAG	Forward
	GGATCCGAGGAGGTCGTCGTTCCCCCGC	Reverse
HtrA3 PDZ (359-444)	CATATGATACGGATGCGGACGATCACAC	Forward
	GGATCCGAGGAGGTCGTCGTTCCCCCGC	Reverse
DUSP9 (1-384)	CATATGAAGGTGGAAGCTGGGTCCGGCTG	Forward
	GAATTCCGTGGGGGCCAGCTCGAAGGCG	Reverse

Primers for gene site directed mutagenesis

HtrA3 variants	**Sequence (5'→3')**	**Orientation**
HtrA3 S305A (active site mutant)	CAACTACGGGAAC**GCT**GGGGGACCACTG	Forward
HtrA3 F142D	CGCTACAAGTTCAAC**GAC**ATTGCTGACGTGGTG	Forward
HtrA3 F142A	CGCTACAAGTTCAAC**GCT**ATTGCTGACGTGGTG	Forward

HtrA3 F255D	GCGGCCTGGGGAG**GAC**GTGGTGGCCATCG	Forward
HtrA3 G358A	GAAGCGCTTCATC**GCT**ATACGGATGCGG	Forward

For all these mutants, the reverse primers were their complementary sequences.

3.1.3 Kits used for cloning

Name	Purpose	Source
Plasmid miniprep kit	Plasmid isolation	Sigma chemicals (St. Louis, MO)
Gel extraction kit	DNA extraction from agarose gel	Sigma chemicals (St. Louis, MO)
pJET 2.1 blunt end cloning kit	Blunt end ligation	Fermentas (Waltham, Massachusetts, USA)
Quick change site directed mutagenesis kit	Site directed mutagenesis	Stratagene (Cedar Creek, TX, USA)

3.1.4 Host strains

Bacterial host strains used

Strain	Genotype	Source
***E. coli* DH5α (Cloning host)**	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	New England Biolabs (NEB), Ipswich, USA
***E. coli* BL21 (DE3) (Expression host)**	*fhuA2 [lon] ompT gal (λ DE3) [dcm]* Δ*hsdS*	NEB
***E. coli* Rosetta pLysS (Expression host)**	F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3) *pLysSRARE* (*Cam*^{R})	NEB

Mammalian cell line used

HEK293T cells (Human Embronic Kidney cells)

3.1.5 Reagents

3.1.5.1 Antibiotics used for clone selection

a) Kanamycin (Sigma Aldrich)

Stock concentration: 50 mg/ml, dissolved in water and stored at -20^{o}C.

Working concentration: 50 µg/ml, stored at 4^{o}C.

b) Chloramphenicol (Sigma Aldrich)

Stock concentration: 34 mg/ml, dissolved in ethanol and stored at -20^{o}C.

Working concentration: 34 µg/ml, stored at 4^{o}C.

b) Ampicillin (Sigma Aldrich)

Stock concentration: 100 mg/ml, dissolved in water and stored at -20^{o}C.

Working concentration: 100 µg/ml, stored at 4^{o}C.

All antibiotics were filter sterilized using 0.22 µm membranes.

3.1.5.2 Reagents for molecular cloning

a) Polymerase chain reaction (PCR) reagents (20 µL reaction)

Reaction components	**Final concentration**
DNA template	60 ng/µl
Forward Primer	125 ng/µl
Reverse Primer	125 ng/µl
dNTP's Mix	10 mM

MgCl2	50 mM
Pfu enzyme	1 U
Pfu reaction buffer (10X)	1X
PCR water	Volume to 20 µL

All reagents were stored at -20^{o}C.

b) 6X Gel loading dye for DNA (Storage at RT)

Xylene Cyanol FF	0.25% (w/v) (migrates at 4160 bp with TAE)
Bromophenol blue	0.25% (w/v) (migrates at 370 bp with TAE)
Glycerol	30% (w/v)

c) Enzymes and buffers for restriction digestion and ligation (Storage at -20^{o}C)

NdeI, BamHI, EcoRI, NheI, XhoI (NEB).

Fast Digest DpnI and 10X fast digest buffer (Fermentas).

T4 DNA ligase and 10X ligase buffer (NEB).

d) 0.5 M EDTA, pH 8

Note: pH adjusted by NaOH is essential for solubility.

e) GeneRuler 1 Kb DNA ladder *(ThermoFischer Scientific)*

f) Ethidium Bromide (EtBr)

Stock concentration: 10 mg/ml, Working concentration: 0.5 µg/ml

g) Liquid Nitrogen (for snap freezing)

3.1.5.3 Bacterial culture media

a) Luria-Bertani (LB) medium (for 1 l)

LB powder (*Himedia, Mumbai India*)	25 g
Milli Q (MQ)	Make up volume to 1l

Autoclaved and stored at RT (room temperature)

b) LB-Agar Plates (for 1 l)

LB agar powder (*Himedia, Mumbai India*)	35 g
Milli Q (MQ)	Make up volume to 1l

Autoclaved and cooled to about 55 °C before adding respective antibiotic.

The media was then poured into petri dishes (~25 ml/100 mm plate).

3.1.5.4 Reagents for protein expression, purification and estimation [Storage at 4°C unless otherwise mentioned]

a) IPTG (Isopropyl-D-thiogalactoside) *(Sigma)***:** 0.5 mM used for bacterial protein induction.

b) Imidazole *(Sigma)***:** gradient of 10-250 mM used for elution of His-tagged proteins.

c) Glutathione *(Sigma)***:** 30 mM Glutathione used for elution of His-tagged proteins.

d) Maltose *(Sigma)***:** 5-10 mM used for elution of MBP-tagged proteins.

e) Bradford Reagent (*BioRad*)**:** Used for protein estimation.

f) Stained and unstained protein marker (5-245 kDa) *(APS Lifetech)* (Storage at -20°C)

g) 5X SDS Loading dye (Storage at RT)

TrisCl (pH 6.8)	0.25 M
β-Mercaptoethanol	0.5 M
SDS	10%
Glycerol	10%

Bromophenol blue	0.25%

Makeup required volume using Milli Q water.

h) Staining solution (Storage at RT)

Methanol	50%
Glacial acetic acid	10%
Coomassie Brilliant Blue R250	0.1%

Makeup required volume Milli Q water.

i) Resolving gel composition

Percentage	6%	7%	10%	12%	15%
D/W (ml)	8.48	7.68	6.4	5.28	3.68
30% ACRYLAMIDE (ml)	3.2	4	4	6.4	8
1.5M Tris – 8.8pH(ml)	4	4	4	4	4
10% SDS (µl)	160	160	160	160	160
10% APS (µl)	160	160	160	160	160
TEMED (µl)	16	14	8	8	8

j) Stacking gel composition

Volume (ml)	2	3	4	5	6
D/W (ml)	1.4	2.1	6.4	5.28	3.68
30% ACRYLAMIDE (ml)	3.2	4	4	6.4	8
1.5M Tris – 8.8pH(ml)	4	4	4	4	4
10% SDS (µl)	160	160	160	160	160
10% APS (µl)	160	160	160	160	160
TEMED (µl)	16	14	8	8	8

k) Destaining solution (Storage at RT)

Same as staining solution but does not contain Coomassie Brilliant Blue R250.

l) 30% Acrylamide

(29.2% Acrylamide, 0.8 % N'N'-bis-methylene-acrylamide)

m) 10X Protease inhibitor cocktail *(sigma)* (Storage at -20°C)

n) 1M Imidazole

o) BSA standards

p) PVDF membrane for western blot

q) Luminiscence reagent (ECL-prime, GE-Healthcare)

3.1.6 Buffers

3.1.6.1 Buffers for molecular cloning

a) Tris-EDTA (TE) Buffer

Tris	10 mM
EDTA	1 mM

pH adjusted to 7.5, volume made up with Milli Q water and autoclaved.

b) Tris Acetate EDTA (TAE) running buffer (for agarose gel electrophoresis)

50X TAE (Stock concentration)

TRIS	2M
Acetic acid	114.2 mL (for 1Ltr)
EDTA	50 mM

pH adjusted to 7.5, autoclaved and stored at RT (room temperature).

Working concentration: 1X TAE (For 1 l, use 20 ml of 50X TAE and make up volume with Milli Q water).

c) Buffers for competent cell preparation

Transformation buffer 1:	
Potassium Acetate	30 mM
Manganous Chloride	50 mM
Potassium Chloride	100 mM
Calcium chloride	10 mM
Glycerol	15% (v/v)
Transformation buffer 2:	
MOPS	10 mM
Potassium Chloride	10 mM
Calcium chloride	75 mM
Glycerol	15% (v/v)

3.1.6.2 Buffers for protein purification and estimation

3.1.6.2.1 SDS-PAGE Running Buffer (no need to adjust pH)

Components	**Final concentration**
Tris Base	25 mM
Glycine	192 mM
SDS	1%

3.1.6.2.2 Ni -IDA Column Purification buffers

Buffers for HtrA3/HtrA2 and its variants

a) 1M Dibasic or disodium phosphate (Na_2HPO_4)

b) 1M Monobasic or monosodium phosphate (NaH_2PO_4)

c) 10X Phosphate buffer, pH 8

Stock concentration: 10X (200 mM NaH_2PO_4/Na_2HPO_4, 2 M NaCl, pH 8).

To 200 ml of 1M dibasic solution, 1M monobasic solution was added and pH was adjusted to 8. Later 2M NaCl was added, volume was made to 1l using Milli Q. Buffer was filtered using 0.2 µm filter and autoclaved.

Working concentration: 1X (20 mM NaH_2PO_4/Na_2HPO_4, 200 mM NaCl, pH 8).

d) Lysis buffer

Components	**Final concentration**
10X Phosphate buffer	1X
Imidazole	10 mM
Triton X 100	0.1%
10X Protease inhibitor	1X
Glycerol	10%

pH adjusted depending on the pI of the protein.

e) Binding/Washing Buffer

Components	**Final concentration**
10X Phosphate buffer	1X
1M Imidazole	10 mM
Glycerol	2%

For 1L, 20 ml of glycerol and 10 ml of 1M Imidazole was added to 100 ml of 10X Phosphate buffer and volume was adjusted to 1l using Milli Q. pH was adjusted as required. Buffer was filtered using 0.2 µm filter and autoclaved.

f) Elution Buffer

Components	Final concentration
10X Phosphate buffer	1X
Imidazole	1M
Glycerol	2%

For 1l, 20 ml of glycerol and 68 g of Imidazole was added to 100 ml of 10X Phosphate buffer and volume was adjusted to 1l using Milli Q. pH was adjusted as required. Buffer was filtered using 0.2 µm filter and autoclaved.

Gradient buffers for elution of protein of interest from column were further prepared as follows:

Imidazole concentration	Volume of Binding buffer (ml)	Volume of Elution buffer (ml)
10 mM (100 ml)	100	-
20 mM (100 ml)	99	1
50 mM (100 ml)	96	4
100 mM (30 ml)	27.3	2.7
250 mM (30 ml)	22.8	7.2

Note: All the buffers used were pre chilled at 4^0C.

g) Stripping buffer

Components	Final concentration
Sodium phosphate	20 mM
NaCl	100 mM
EDTA	50 mM

Volume was adjusted to the required amount with Milli Q water.

h) $NiSO_4$ charge solution (5X)

500 mM nickel sulphate hexahdrate. Volume was adjusted to the required amount with Milli Q water.

Buffers for DUSP9 and its variants

a) 1X Tris buffer

(20 mM Tris, 200 mM NaCl, pH 8.5)

b) Lysis buffer

Components	**Final concentration**
Tris buffer	1X
Imidazole	10 mM
Triton X 100	0.1%
10X Protease inhibitor	1X
Glycerol	10%
BME	10 mM

pH adjusted depending on the pI of the protein.

c) Binding/Washing Buffer

Components	**Final concentration**
Tris buffer	1X
1M Imidazole	10 mM
Glycerol	10 %
BME	10 mM

For 1L, 100 ml of glycerol and 10 ml of 1M Imidazole was added to 500 ml of 1X Tris buffer and volume was adjusted to 1L using1X Tris buffer. pH was adjusted as required. Buffer was

filtered using 0.2 µm filter and autoclaved.

f) Elution Buffer

Components	**Final concentration**
Tris buffer	1X
Imidazole	1M
Glycerol	10%
BME	10 mM

For 1L, 100 ml of glycerol and 68 g of Imidazole was added to 500 ml of 1X Tris buffer and volume was adjusted to 1l using 1X Tris buffer. pH was adjusted as required. Buffer was filtered using 0.2 µm filter and autoclaved.

Gradient buffers for elution of protein of interest from column were further prepared as follows:

Imidazole concentration	**Volume of Binding buffer (ml)**	**Volume of Elution buffer (ml)**
10 mM (100 ml)	100	-
20 mM (100 ml)	99	1
50 mM (100 ml)	96	4
100 mM (30 ml)	27.3	2.7
250 mM (30 ml)	22.8	7.2

Note: All the buffers used were pre chilled at 4^0C.

3.1.6.2.3 Amylose Column Purification buffers

a) 0.5 M Stock Maltose

b) Lysis buffer

1X Phosphate buffer with 10% glycerol, 0.1% Triton and 1X Protease inhibitor cocktail. pH was

adjusted depending on the pI of the protein.

c) Binding/Washing Buffer

1X phosphate buffer with 2% Glycerol. pH was adjusted as required. Buffer was filtered using 0.2 µm filter and autoclaved.

d) Elution buffer

Components	**Final concentration**
10X Phosphate buffer	1X
0.5 M Maltose	10 mM
Glycerol	2 %

pH was adjusted as required. Buffer was filtered using 0.2 µm filter and autoclaved.

3.1.6.2.4 GST Column Purification buffers

a) 0.5 M Stock Glutathione

b) Lysis buffer

1X phosphate buffer with 10% glycerol, 0.1% Triton and 1X Protease inhibitor cocktail. pH was adjusted depending on the pI of the protein.

c) Binding/Washing Buffer

1X phosphate buffer with 2% Glycerol. pH was adjusted as required. Buffer was filtered using 0.2 µm filter and autoclaved.

d) Elution buffer

Components	**Final concentration**
10X Phosphate buffer	1X
0.5 M Glutathione	30 mM
Glycerol	2 %

pH was adjusted as required. Buffer was filtered using 0.2 µm filter and autoclaved.

3.1.6.3 Buffers for Western Blot

a) 1X Transfer Buffer

Components	**Final concentration**
Tris Base	25 mM
Glycine	192 mM
Methanol	20%

No need to adjust pH. Pre chilled at 4^0C

b) TBST (for washing) [Storage RT]

TRIS 25 mM, pH 7.5, NaCl 150 mM, Tween-20 0.1% (w/v)

3.1.7 Resins and reagents used for protein purification and size exclusion chromatography

Resin	**Purpose**	**Source**
Glutathione Sepharose 4B	Affinity chromatography of GST-tagged proteins	GE healthcare, Healthcare, Bjorkgatan, Uppsala, Sweden
Ni-IDA	Affinity chromatography of His_6-tagged proteins	Biotex, Houston, USA
Amylose	Affinity chromatography of MBP-tagged proteins	NEB
Superdex 200	Size exclusion chromatography	GE healthcare, Healthcare, Bjorkgatan, Uppsala, Sweden

Protein standards (*Sigma*) for size exclusion chromatography: Alcohol dehydrogenase (ADH), Bovine serum albumin (BSA) and Lysozyme.

3.1.8 Specialized instruments

Instrument	Purpose	Source
Jasco J815 CD spectrophotometer	Secondary and tertiary structure of proteins	JASCO, Easton, MD, USA
Fluorolog-3 spectrofluorimeter	Fluorescence emission studies, enzyme kinetics, quenching, FRET	HORIBA scientific, Edison, NJ, USA
Mithras multiwell plate reader	Enzyme kinetics	Berthold technologies, Wildbad, Germany
AKTA purifier	Gel filtration	GE Healthcare, Buckinghamshire, United Kingdom

3.1.9 Mammalian cell culture reagents

3.1.9.1 Antibodies

Primary antibodies [Storage -20/4°C];

Anti-HtrA2 1:1000 (Rabbit monoclonal, *Sigma*)

Anti-FLAG 1:8000 (mouse monoclonal, *Sigma*)

Anti-HtrA3 1:1000 (Goat polyclonal, *Santa Cruz Biotechnology*)

Anti-DUSP9 1:1000 (Mouse monoclonal, *Santa Cruz Biotechnology*)

Secondary antibodies [Storage -20/4°C];

Anti-mouse-HRP *(GE-healthcare)*

Anti-rabbit-*HRP (GE-healthcare)*

Anti-goat-*HRP (Santa Cruz Biotechnology*)

3.1.9.2 Buffers

a) Immunoprecipitation Buffer: [Storage 4°C]

Components	Final concentration
Tris (pH 7.5)	50 mM
NaCl	150 mM
NP-40	0.4-0.6%
Protease inhibitor	1X

b) NP-40 Lysis Buffer: For Mammalian cell lysis. [Storage 4°C]

Components	Final concentration
TRIS	50 mM (pH 7.5)
NaCl	150 mM
NP-40	1%
DTT	1mM
Protease inhibitor	1X
NaF	10 mM
Na_3VO_4	1mM
β-Glycerophosphate	10 mM

3.1.9.3 Culture media

Dulbecco's Modified Eagle Medium (DMEM)

Components	Final concentration
Powdered medium (DMEM)	1 pack as per manufacturer's Protocol
Sodium carbonate	3.5 gm

FBS (GIBCO Invitrogen)	10%
Antibiotic mixture (100IU of penicillin per ml, and 100µg of streptomycin per ml and amphotericin)	1X

pH was adjusted to 7.2 using 1 N HCl and the volume was made up to 1l was filtered by a sterile filter assembly and stored at 4°C. Foetal bovine serum (FBS) and antibiotic mixture were later added to prepare the complete medium.

3.1.9.4 Other reagents

a) 1X sterile PBS [Storage 4°C].

b) 1X sterile trypsin made in 1X PBS [Storage 4°C].

c) DMSO (Sigma) [Storage RT].

d) Protein-G sepharose beads/ anti-FLAG-M2 Agarose beads (GE Healthcare).

e) 8M urea [denaturant].

3.1.9.5 Kits used

a) Lipofectamine 2000/3000 (Invitrogen)

3.1.9.6 Softwares used

a) ClusPro 2.0 server (https://cluspro.bu.edu/home.php)

b) PyMOL (The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC)

c) GROMACS 2018

d) XmGrace ((ftp://plasma-gate.weizmann.ac.il/pub/grace/)

e) Bioluminate (Bioluminate, Schrödinger, LLC, New York, NY, 2018)

f) SiteMap (Sitemap, Schrödinger, LLC, New York, NY, 2018)

g) Glide (Glide, Schrödinger, LLC, New York, NY, 2018)

h) 3D builder (Schrödinger, LLC, New York, NY, 2018)

i) Maestro (Schrödinger, LLC, New York, NY, 2018)

k) PDBSum generate server, http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html

3.2 METHODS

3.2.1 Methods in molecular cloning

3.2.1.1 Estimation of DNA concentration

A micro-volume UV-Vis spectrophotometer (NanoDrop, Model-ND 1000) was used to measure the absorbance of the DNA samples at 260 nm and calculate the concentration using the Beer-Lambert Law.

$$A = \varepsilon cl$$

Where,

A represents absorbance (no units), Extinction coefficient or ε (L mol^{-1} cm^{-1}) represents the molar absorptivity, l is the path length of the sample in cm and c (mol L^{-1}) is the concentration.

The absorbance is directly proportional to the path length of the sample and the concentration of the sample.

The ratio of absorptions at 260 nm versus 280 nm was used to assess the purity of the sample with respect to protein contamination. Ideally, for nucleic acids, the 260:280 ratio should be around 1.6-1.9.

3.2.1.2 Preparation of primers

Primers were received as lyophilized pellets. These were re-suspended in 10 mM Tris, pH 7.5 or autoclaved Milli Q water, pH 7 to obtain stock concentration of 1µg/µl. The suspension was

mixed thoroughly and centrifuged at 10,000 rpm for 1 min before storing at -20° C. The stock was used to prepare a working concentration of 125 ng/µl. Detailed list of primers are provided in **section 3.1.2 (materials).**

3.2.1.3 Polymerase chain reaction (PCR) for gene amplification and site directed mutagenesis

For gene amplification from cDNA library or sub cloning, the gene of interest was PCR amplified using gene specific forward and reverse primers flanking the region of interest and containing appropriate restriction sites.

For site directed mutagenesis (SDM), whole plasmid with gene of interest was amplified using the forward and reverse primers carrying the desired mutation. The reverse primer was complementary sequence of the forward primer. The primers were designed with the help of 'oligoanalyser' tool keeping in mind parameters such as an optimum primer length of 18-30 nucleotides, 40-60% GC content and a Tm of 50-65 °C. pfu turbo polymerase (Stratgene kit) with proofreading activity was the used for all reactions as mentioned in the section on materials. The typical cycling steps optimized for the PCR are:

STEPS	TEMPERATURE	TIME
Initial denaturation	95°C	5 min
Denaturation	95°C	30 sec
Annealing	53°C	55 sec
Extension (elongation time depends on size of the DNA and processivity of polymerase)	72°C	10 min
No. of cycles	for gene specific amplification 25 cycles and for SDM 18 cycles were considered	
Final extension	72°C	15 min
Termination	4°C	∞

For SDM, PCR products were treated with DpnI endonuclease enzyme at 37 °C to degrade the methylated parental template. This digested PCR product was then transformed in E. coli DH5α competent cells (described subsequently). Colonies so obtained were screened to isolate the plasmid carrying the desired mutation using the automated DNA sequencing facility.

3.2.1.4 Agarose gel electrophoresis

Appropriate amount (0.8 – 1 %) of agarose was weighed and dissolved in 1X TAE. The solution was boiled to dissolve the agarose completely and then allowed to cool to ~50-60 ^{o}C. Ethidium bromide (0.5µg/ml) was added for fluorescent visualization of the DNA fragments under UV light. The agarose solution was poured in a gel casting tray with comb and allowed to polymerize for 20-30 mins. The DNA samples as well as the DNA ladder (for size reference) were loaded and resolved at 100 volts for 30-50 min. Lastly, the DNA bands were analyzed under UV 365 nm followed by documentation (UVP, Bioimaging Systems) and visualization (LaunchVision Works LS software) of gel.

3.2.1.5 Preparation of competent cells

Strains used: *E. coli* DH5α, BL21 (DE3) and RosettapLysS.

Procedure:

1. A Glycerol stock vial of *E. coli* bacteria strain (stored at -80°C) was thawed and streaked on to LB-agar plate without any antibiotic and incubated at 37ºC overnight.

Note: for pLysS strains, chloramphenicol was used for selection.

2. A single isolated colony was inoculated in 5ml of LB and incubated overnight at 37°C at 180 rpm.

3. From this culture, 2.5ml was transferred to a flask containing 100ml of LB and incubated at 37°C; 180 rpm till an O.D of 0.6 was achieved.

4. Following this, the flask was kept on ice for 15 min to arrest cell multiplication prior to centrifugation at 5000 rpm for 15 min.

5. Supernatant was discarded and the cell pellet was re-suspended gently in 20ml of Transformation Buffer 1.

6. The reconstituted pellet was centrifuged at 5000 rpm for 8 min and supernatant was discarded.

7. Finally, 2 ml of Transformation Buffer 2 was used to re-suspend the pellet. 100 µl of cellsuspension was aliquoted in 1.5 ml tubes and snap frozen in liquid nitrogen. These vials were subsequently stored at - 80°C.

3.2.1.6 Transformation

Procedure:

1. The competent E. *coli* cells retrieved from -80 °C were thawed on ice.

2. 50 ng of plasmid DNA was added to a vial of competent cells and kept on ice for 10-15 min.

3. After incubation, the cells were subjected to heat shock at 42°C for exactly 90 seconds in a water bath followed by incubation on ice for 5 min.

4. 750 µl of LB medium was added to the tube and the cells were incubated at 37 °C for 45 min to 1 hr with shaking at 180 rpm.

5. Cells were then centrifuged at 5000 rpm for 5 min; supernatant was discarded leaving ~100 µl for re-suspending the pellet.

6. Cells were spread on LB plate with appropriate antibiotic and incubated overnight (~16 hrs) at 37 °C.

3.2.1.7 Restriction digestion

A typical restriction digestion reaction was set up by adding the following components (for a final volume of 20 µl):

Components	Final concentration
DNA to be digested	2 µg
Enzyme 1	1 unit
Enzyme 2	1 unit
Appropriate buffer (10 X)	1 X
Autoclaved MQ water pH, 7.0	Make volume up to 20 µl

All reactions were incubated at 37°C for 1-4 hours. The digested products were then isolated by gel extraction (Gel Extraction Kit, Sigma).

3.2.1.8 Ligation

A typical ligation reaction was set up by adding the following components (for a final volume of 20 µl):

Components	Final concentration
Vector concentration	60-80 ng
Insert concentration	Calculated using the formula mentioned below
T4 DNA ligase	1 unit
Ligase buffer (10 X)	1 X
Autoclaved MQ water pH, 7.0	Make volume up to 20 µl

Vector to insert molar ratio was calculated using formula:

$$\frac{\text{Vector concentration (ng)}}{\text{Vector size (bp)}} = \frac{\text{Insert concentration (ng)}}{\text{Insert size (bp)}}$$

$$\text{Insert concetration (ng)} = \frac{\text{Vector concentration (ng) X Insert size (bp)}}{\text{Vector size (bp)}}$$

This equation represents 1:1 molar ratio of vector: insert concentration. For a molar ratio of 1:3, 3 times the insert concentration calculated using the formula above was used.

All reactions were incubated at 16°C for 12 hours or 22°C for 4 hours. The ligated products were then transformed in *E. coli* DH5α cells to obtain colonies.

Blunt end ligation

pJET cloning kit was used and ligation was carried out according to the manufacture's protocol. The vector and insert were used in 1:3 molar ratios and ligated at 22°C for 30 min.

Cohesive end ligation

The vector and insert were used in 1:3 molar ratios and ligated at 22°C for 30 min.

3.2.1.9 Plasmid isolation (Mini prep)

Bacterial colonies (with plasmid of interest transformed in *E. coli* DH5α cells) were inoculated in 10 ml of LB broth containing appropriate antibiotics and incubated at 37ºC, 180 rpm for 14-16 hr. The cultures were then centrifuged and the pellets were used for plasmid extraction following the kit manufacture Protocol.

Procedure:

a) Column and lysate preparation:

1. Column containing silica resin was regenerated by adding 500 µl of column preparation solution to the column and centrifuging it at 14,000rpm for 1min.

2. Culture grown overnight was centrifuged at 5000 rpm at 4°C for 15 min.

3. Supernatant was discarded and the pellet was re-dissolved using 250 µl P1 (resuspension buffer).

4. To this mixture 250 µl of P2 (lysis buffer) was added and mixed gently by inverting the tube

several times.

5. Afterwards, 350 µl of N3 (neutralization buffer) was added.

6. The mixture was centrifuged at 13000 rpm for 10 min.

b) Binding of plasmid DNA:

1. The supernatant was transferred to the regenerated column, which is a spin column containing silica resin that binds DNA.

2. After the supernatant was added, the spin column was centrifuged at low speed and the flow through was discarded.

3. The column was then washed with 750 µl PE buffer (wash buffer) containing alcohol. This was then centrifuged at 13000 rpm for 2 min.

4. An empty spin was given at same speed and time to remove leftover alcohol.

c) Elution of plasmid DNA:

1. Column was placed in a fresh 1.5 ml tube and 30 µl elution buffer, having Tris-Cl of pH 8.5 was added to the column.

2. Tube was kept on stand for 5 min for optimal binding.

3. The tube was centrifuged at 13,000 rpm for 2 min to elute the plasmid DNA bound to the silica membrane.

4. 2nd and 3rd elutes were collected using less quantity of buffer; say 25 µl and 20 µl to recover all residual DNA bound to the membrane.

d) Nanodrop:

The eluates were checked for the concentration and purity (260/280 ration) of DNA using the buffer as a blank.

3.2.1.10 Gel extraction

Gel extraction was performed using Sigma Genelute Gel Extraction kit as mentioned in the section on materials. DNA was extracted from an agarose gel by cutting out the required section and solubilizing it. The extracted DNA fragment was adsorbed onto a silica membrane containing column provided in the kit and finally eluted in Tris buffer following the kit manufacture Protocol.

Procedure

1. The DNA fragment was excised from the agarose gel and trimmed to remove excess.

2. The gel slice was weighed in an eppendorf. Three gel volumes of the gel solubilization solution was added to the gel slice and incubated at 50-60°C for10 min. The mixture was vortexed briefly every 2-3 minutes during incubation to help dissolve the gel.

4. Meanwhile, the binding column (GenElute binding column G) was prepared by adding 500 µl of the column preparation solution to the binding column and centrifuging for 1 minute. Flow through was discarded.

5. One gel volume of 100% isoproponol was mixed to the dissolved gel mixture until homogenous.

8. The solubilized gel solution mixture was then loaded into the binding column assembled in a 2 ml eppendorf and centrifuge for 1 minute. Flow through was discarded.

9. 700 µl of wash buffer (PE) was added to the binding column and centrifuged for 1 min.

10. The remaining steps for elution and estimation of DNA concentration was done exactly as described in section ***3.2.1.9.***

3.2.1.11 Cloning of HtrA3and its variants

cDNA of full length HtrA3 (1-453 amino acids) was obtained in pDONR221vector from DNASU plasmid repository (The Biodesign Institute/Arizona State University). The open

reading frame (ORF) was amplified using PCR and then subcloned in pJET 1.2 blunt ligation vector between NdeI and BamHI restriction sites. From this construct, all the domains and their combinations were subcloned in expression vectors (pET28a, pET20b, pMALc5E) by PCR between NdeI and BamHI restriction sites. The mutants were generated using appropriate primers through site directed mutagenesis. Sequences of all these variants were confirmed by automated DNA sequencing facility at ACTREC.

3.2.1.12 Cloning of DUSP9

DUSP9 ORF was obtained in pBLUESCRIPT vector from Origene. This was subsequently subcloned into expression vectors (pET28a, pGEX4T, pMALc5E) by PCR between NdeI and EcoRI restriction sites using PCR amplification.

3.2.2 Methods in protein expression, purification and estimation

3.2.2.1 Growing bacterial cultures

Strains used: *E. coli* BL21 (DE3) and RosettapLysS.

Procedure:

Growing starter cultures

A single, transformed, isolated colony from a freshly transformed plate of the respective expression host was inoculated in 10 ml LB medium containing appropriate selective antibiotics and grown overnight at 37 °C with constant shaking at 180-200 rpm.

Growing secondary cultures

1. The following day, 10 ml of the starter culture was added in a ratio of 1:100 to autoclaved 1l LB medium containing appropriate antibiotics for growth in bulk.

2. The culture was grown till the O.D600 reached 0.6.

3.2.2.2 Protein expression

For protein expression the gene of interest was cloned downstream of a promoter in pET28a/ pET20b/pMALc5e/pGEX4T expression vectors. The vectors were then introduced into either in E. coli BL21 (DE3) or Rosetta pLysS host expression strains. These strains lack the Lon/OmpT proteases and carry a chromosomal copy of the T7 RNA polymerase under the control of lacUV5 promoter (inducible by IPTG) and therefore can conveniently express genes driven by the T7 promoter. Additionally, Rosetta pLysS was used for expression of toxic genes. This strain particularly contains a chloramphenicol resistant pLysS plasmid which encodes for T7 lysozyme to prevent basal or leaky expression. The expression of T7 polymerase was induced by IPTG (isopropyl-β-D- thioglactopyranoside) to initiate heterologous protein synthesis. IPTG prevents the Lac repressor from binding to the Lac operator, thereby enabling transcription to occur.

Procedure:

1. Culture flasks having cells with O.D600 of 0.6 were induced with 0.5 mM IPTG and incubated at 18 °C for 16 hrs.

2. The following day, cells were harvested by centrifugation at 5,000 rpm for 15 min at 4 °C. The cells pellets were then stored at -80 °C until further use**.**

3.2.2.3 Protein purification

Affinity chromatography was used to purify the recombinant protein from tagged vectors. The procedure involved separating molecules based on the reversible interaction between target protein and the specific ligand attached to a chromatography matrix.

For certain proteins an additional step of size exclusion chromatography (separation of molecules on the basis of size) was employed to ensure enhanced purity of the affinity-purified protein samples.

3.2.2.3.1 Ni-IDA agarose affinity chromatography

His_6 tagged proteins were purified using Ni–IDA affinity chromatography, where the His_6-tag selectively binds to Ni-IDA (a tridentate chelating agent, IDA covalently coupled to agarose beads) on a chromatographic column. Histidine tags are small and are less disruptive than other bigger tags to the properties of the proteins they are attached with.

Procedure:

A. Regeneration and equilibration of Ni-IDA resin

1. 3 ml of Ni–IDA agarose beads (Biotex) were taken in a 10 cm econo column (Millipore).

2. For fresh beads, 4-5 washes with autoclaved distilled water (DW) were given. For beads stored in 20% ethanol, the ethanol was first allowed to flow through and then the beads were thoroughly washed with DW.

3. The beads were stripped using stripping buffer (1 column volume i.e. ~10 ml) followed by one autoclaved distilled water wash.

4. One column volume 100 mM NiSO4 solution was then passed through the column to recharge the beads.

5. Finally, one or more autoclaved distilled water wash was given to remove unbound nickel and the column was equilibrated with pre-chilled binding buffer for at least half an hour.

B. Preparation of cell lysate

1. Cell pellet was resuspended in lysis buffer and mixed to obtain a homogenous slurry.

2. The viscous mixture was sonicated (to disrupt the cell membrane and remove its cellular contents) at 50 pulse rate with 50 % power for 1.45 minutes of duty cycle and centrifuged at 15000 rpm for 30 minutes (*Beckman coulter* JA-25.50 rotor) at 4^0C on a rocker with gentle agitation.

C. Affinity Purification

1. The supernatant was incubated with the equilibrated Ni-IDA beads and the column was kept on the rocker for binding for 45 minutes to 1hr at 4^0C.

2. The flow through was collected after incubation.

3. The His_6 tagged proteins were eluted with elution buffer containing different concentrations of imidazole (10-250 mM gradient). The proteins of interest were mainly collected in the 100 and 250 mM fractions.

3.2.2.3.2 Amylose affinity chromatography

All Maltose Binding Protein (MBP) tagged proteins were purified with amylose affinity purification method, since MBP has affinity for amylose. The bound protein is eluted using different concentrations of maltose since maltose has higher affinity for amylose therefore it replaces the MBP tagged protein. This displaced protein is collected in fractions.

Procedure:

A. Regeneration and equilibration of amylose resin

1. 3 ml of Amylose beads (Novagen) were taken in a 10 cm econo column (Millipore).

2. For fresh beads, 4-5 washes with autoclaved distilled water (DW) were given. For beads stored in 20% ethanol, the ethanol was first allowed to flow through and then the beads were thoroughly washed with DW.

3. One column volume of 0.1% SDS was then passed through the column twice.

5. Finally, 2-3 autoclaved distilled water wash was given and the column was equilibrated with pre-chilled binding buffer for at least half an hour.

B. Preparation of cell lysate

Procedure was same as described in section 3.2.2.3.1

C. Affinity Purification

1. The supernatant was incubated with the equilibrated amylose beads and the column was kept on the rocker for binding for 45 minutes to 1hr at 4^0C.

2. The flow through was collected after incubation and beads were washed with 3 column volumes of wash buffer.

3. Bound proteins were eluted using elution buffer containing 10 mM maltose.

3.2.2.3.3 GST affinity chromatography

All glutathione S-transferase (GST) tagged proteins were purified with this affinity purification method. GST-tagged proteins specifically bind to its tripeptide (Glu-Cys-Gly) substrate, Glutathione, immobilized through a sulfhydryl group to cross-linked beaded agarose. The fusion protein is eluted by passing excess reduced glutathione.

Procedure:

A. Regeneration and equilibration of Glutathione sepharose beads

1. The column was first washed with autoclaved D/W 3-4 columns.

2. It was then washed with two column volumes of 0.1 M Tris HCl and 0.5 M NaCl (pH 8.5).

3. Following this the column was washed with two column volumes of 0.1 M sodium acetate and 0.5 M NaCl (pH 4.5).

4. Steps 2 and 3 were repeated for 3-4 times.

5. Finally, the column was equilibrated with pre-chilled binding buffer for at least half an hour.

B. Preparation of cell lysate

Procedure was same as described in section 3.2.2.3.1

C. Affinity Purification

1. The supernatant was incubated with the equilibrated amylose beads and the column was kept on the rocker for binding for 45 minutes to 1hr at 4^0C.

2. The flow through was collected after incubation and beads were washed with 1 column volume of wash buffer.

3. Bound proteins were eluted using elution buffer containing 30 mM Glutathione.

3.2.2.4 Methods in protein estimation and visualization

3.2.2.4.1 Sodium Dodecyl Sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were electrophoretically separated based on their size. The proteins are denatured and imparted a uniform charge to mass ratio with the anionic detergent SDS and boiling. The protein samples were mixed with 5X SDS loading dye, boiled for 5 min and loaded onto the gel. The stacking gel concentrates the proteins prior to entering the resolving gel, where they are further resolved on the basis of their molecular weights. The gel was run at 100 V for about 60 min, stained with staining solution for 10-15 min, then destained overnight in destaining solution and finally preserved in 10% glacial acetic acid and documented.

3.2.2.4.2 Nanodrop

Proteins were checked for the concentration and purity by measuring their absorbance at 280 nm with a Nanodrop based on the Beer-Lambert law as described earlier in section ***3.2.1.1***. The buffer without protein was used as blank. Protein concentrations were generally represented in the following units:

1) Concentration in mg/ml = Absorbance $_{280}$/ ε x Molecular weight of protein (Daltons)

2) Concentration in µM = Absorbance $_{280}$/ ε x 1000000

3.2.2.4.3 Bradford assay for protein estimation

The assay is based on the principle that binding of proteins to acidic solution of coomassie brilliant blue G-250 shifts the absorbance maximum from 465 nm to 595 nm (brown to blue). This method spectroscopically analyses the presence of the basic amino acid residues (arginine,

lysine and histidine), which contribute to formation of the protein-dye complex [114].

Procedure:

Preparation of standard curve using microassay method

1. Standard solutions of BSA 1, 0.5, 0.25, and 0.125 mg/ml were prepared.

2. 5 µl from each standard was added to 200 µl of Bradford reagent (1:4 diluted, Bio-Rad) in a 96-well plate and readings were recorded with ELISA plate reader (Spectra Max 790) at 595 nm using SoftMaxPro 4.6 software. Experiment was performed in triplicates.

3. The standard curve was plotted with protein concentration on the X axis and absorbance reading on Y-axis

Determination of protein concentration

Concentration of protein was determined by extrapolating from the standard graph prepared.

3.2.3 Biophysical methods for protein characterization

3.2.3.1 Increasing protein concentration

Procedure:

1. The centricon with appropriate molecular weight cut-off was taken and washed thoroughly with water to remove the ethanol it is stored with.

2. It was then equilibrated with wash buffer by centrifuging it at 3500 rpm for 10 minutes at 4°C.

3. The protein of interest was collected in the centricon and centrifuged at 3500 rpm for 10 minutes at 4°C. This step was repeated till the volume was brought down to 2 ml. The flow through was discarded.

4. Finally, the protein sample was centrifuged for 10 minutes at 140000 rpm at 4°C to remove precipitates or aggregates and concentration was estimated using Nanodrop.

3.2.3.2 Size exclusion chromatography

Size exclusion chromatography or Gel filtration was used to determine the molecular weight of the proteins based on size and volume. The pores of column matrix act as molecular sieves where, small molecules can diffuse into the pores whereas large molecules migrate out faster since their size prevents them from entering the pores. This method is often used as an additional step in protein purification to improve the purity of affinity-purified proteins.

Procedure:

1. The chromatography column was first calibrated by running protein standards with known molecular weights in appropriate buffer to prepare the standard curve. The proteins used were Alcohol dehydrogenase (ADH), Bovine serum albumin (BSA), Lysozyme and MBP. All standards were purchased from Sigma except for MBP that was lab purified.

2. To find the molecular weight of the unknown protein, the column was first pre-equilibrated with appropriate buffer as described in section on materials (eg. 1X dialysis for HtrA3, 1X Tris buffer for DUSP9).

3. 1 ml of the concentrated protein (2-3mg) was injected on the column and eluted with the same buffer at a flow rate of 0.5 ml/min.

4. Elution of the protein was monitored with the help of the absorbance reading at wavelength 215 or 280 nm. The fractions were collected corresponding to the exact elution volume for protein. Elution volume (Ve) / void volume (V_0) of the protein was used to calculate its molecular weight from the prepared standard curve.

Data analysis:

1. Elution volume (Ve) / void volume (V_0) versus log of molecular weights of standards were plotted to generate the calibration curve. The graph was fitted to the linear equation given below and its slope was obtained.

$$y = mx + c$$

Where,

m is the gradient and c is the intercept on the y-axis.

2. Thereafter, molecular weights of unknown proteins were extrapolated from this curve. The molecular weight of the unknown protein was determined from plot of the log molecular weight of the standard proteins versus k_{av} ($k_{av}=V_e-V_0/V_c-V_0$), where V_e is elution volume of the protein, V_c is the volume of the column, and V_0 or void volume calculated by determining the elution volume of blue dextran (1 mg/ml).

All the standards and proteins of interest were run on the column three times and the average of the retention volumes were used for data analysis.

3.2.3.3 Circular Dichroism (CD)

The secondary structural properties (alpha helices, beta sheets, and turns) of HtrA3 and its variants were determined using CD spectroscopy [115, 116]. This method uses the property of chiral molecules to differentially absorb the right and left components of circularly polarized light. CD spectra in the far UV range (260-180 nm) was used to analyze the proteins, where, an absorption minima at 208 and 222 nm indicates α-helical structure, a minimum at 218 nm is typical of β-sheets and random coil regions are characterized by a low elipticity at 210 nm and negative band near 195 nm. CD can also determine the overall folding/unfolding and therefore the conformational changes in proteins as an effect of pH, denaturant or temperature. Hence, it is used for thermal stability studies, unfolding experiments and mutational analysis. The thermal stability of HtrA3 variants were also studied by this method.

Procedure:

1. 10 μM protein concentrations were taken in a quartz cuvette (1 mm path length).

2. Far-UV CD scans (260-190 nm) were recorded using a JASCO J 815 spectropolarimeter (JASCO). The parameters taken into consideration were:

scan rate-20 nm/s, accumulations-3, data pitch- 0.1, temperature -25 ºC.

3. For thermal denaturation studies, scans were collected as described with increasing temperature (22-100 ºC) with an increment of 2ºC/min. At each data point, the sample was equilibrated for 5 min. All experiments were performed in triplicates.

Data analysis:

Secondary structure determination

The CD spectra were plotted with ellipticity (θ) (Y-axis) as a function of wavelength (nm) (X-axis). For concentration independent plots, data was represented in the form of the Mean Residual Ellipticity (MRE or [θ]) with the unit $deg.cm^2.mol^{-1}$. Ellipticity was converted to mean residue ellipticity using the following formula:

$[\theta]MRE = (\theta * MRW)/ (10 * c * d)$

where,

MRW (Mean residue weight) = Molecular weight / (number of amino acids -1), 'c' represents concentration of protein (mg/ml) and d stands for the pathlength in cm.

Thermal denaturation

Ellipticity corresponding to 222 nm at different temperatures were used obtain the melting temperature (Tm). At the outset, the ellipticity of the completely folded (θ_f) and unfolded forms (θ_u) were estimated using nonlinear regression (GraphPad Prism). These values were then used to calculate fraction folded at any given temperature (α) with the following formula:

$$\propto = \frac{[F]}{[F] + [U]}$$

$$= \frac{\theta i - \theta u}{\theta f - \theta u}$$

Where,

[F] is the concentration of folded forms, [U] the concentration of unfolded forms, θ_i is the observed ellipticity at a given temperature. To calculate the Tm, the fraction folded at given temperature was further analysed using nonlinear regression (Igor Pro).

3.2.3.4 Fluorescence spectroscopy

Fluorescence emission spectroscopy was used to study the tertiary structural properties of HtrA3 and its variants by Fluorolog-3 spectrofluorometer (Horiba Scientific). Fluorescence is based on the principle that a molecule absorbs a lower wavelength photon, undergoes electronic excitation, and then emits longer wavelength. In proteins, aromatic amino acid residues (Phe, Tyr, Trp), especially tryptophan residue owing to high quantum yield, contributes towards the intrinsic fluorescence and provides information regarding the local environment of these aromatic residues [117]. This information in turn can be used to obtain the tertiary structural information of proteins and also be used studying dynamics, protein unfolding and biomolecular interactions. So this technique depends on the presence of aromatic amino acids, primarily tryptophan as Trp fluorescence has high quantum yield. Moreover, tryptophan is termed as the reporter molecule as it is solvent sensitive and provides information of the environment of the molecule in tertiary space. For an excitation wavelength of 280 nm, all the three amino acids are excited. To selectively excite Tryptophan (W) only, 295 nm wavelength is used. All experiments were performed in triplicates.

Procedure:

1. 2 μM protein concentrations were taken in a quartz cuvette (3 mm path length).

2. Fluorescence spectra were recorded Fluorolog-3 spectrofluorometer (Horiba Scientific). The

parameters taken into consideration were:

Excitation-295 nm, Emission between 310-400 nm, slit width- 5 nm, integration time-0.1s.

Data analysis:

The fluorescence spectra were plotted with fluorescence intensity (AU) (Y-axis) as a function of wavelength (nm) (X-axis). The emission maxima were obtained from the wavelengths that corresponded to the maximum intensity.

3.2.4 Biochemical methods for protein characterization

3.2.4.1 Gel based protease assays (Qualitative)

Beta casein (*Sigma*), the generic serine-protease substrate was used to qualitatively determine protease activity of HtrA3 as well as HtrA2 using SDS PAGE. A stock concentration of beta casein having concentration 1 mg/ml was made.

For a time based assays, for each 30 µl reaction mixture, the protein sample (of desired concentration in micrograms for eg: 2µg, 4µg, 6µg) in appropriate assay buffer (1X dialysis buffer) was taken and incubated with beta casein (6µg) for different time periods (0-2 hr) at 37ºC. For concentration based assays, the protein sample was incubated with increasing concentration of beta casein (2µg - 10µg) for 2 hr while enzyme concentration was kept constant. For temperature based assays, protease activity was monitored over 30-65°C.

The reaction was stopped by adding SDS gel loading dye. These samples were then loaded in 12% SDS polyacrylamide gel. After electrophoresis, the gel was stained using staining solution , destained and then visualized.

3.2.4.2 Enzyme kinetics (Quantitative)

For quantitative studies, FITC (fluorescein isothiocyanate) labeled β-casein (*Sigma*) was used as substrate.

Procedure:

1. For each 100 µl reaction mixture, 0.5 µM of enzyme was incubated with increasing concentration (0–20 µM) of FITC β-casein at 37 °C in assay buffer.

2. Proteolytic cleavage was assessed by monitoring increase in fluorescence intensity of unquenched FITC β-casein in a multi-well plate reader (Berthold Technologies) at 485 nm excitation and 535 nm emission wavelengths.

3. Assays were done in triplicates.

Note: FITC β-casein is photosensitive

Data analysis:

Initial velocities were calculated at each respective temperature using linear regression analysis. Graph was plotted with time (min) on X-axis and fluorescence intensity on Y-axis. Slope value for each substrate concentration was determined using linear regression analysis. Reaction rates V_0 (µM/min) of unquenched FITC at respective substrate concentration were determined by dividing the slope with 19422 (slope of free FITC standard curve). The steady-state kinetic parameters were obtained from the reaction rates by fitting data to Michaelis-Menten equation using nonlinear least squares subroutine in KaleidaGraph program (Synergy software).

$$\text{Velocity} = V\text{max}/(1 + \left(\frac{K_{0.5}}{[substrate]}\right)n)$$

Where,

'V_{max}' is the maximum velocity and $K_{0.5}$ is substrate concentration at half maximal velocity.

3.2.4.3 Pull down studies

Affinity pull down studies was used to understand protein-protein interaction. One of the interacting partners, the bait, is expressed as a fusion protein which binds to the chromatography resin. The prey protein is incubated with the bait in the next step. If the two proteins do indeed

interact in the cellular milieu, the prey binds to the immobilized bait. Extensive washes are given to remove unbound prey protein and non specific proteins. Finally, the bound samples were eluted and visualized on SDS-PAGE to verify the interaction.

Procedure:

1. Tagged bait protein was grown and lysed in an appropriate lysis buffer.

2. 10 µg of the lysate was centrifuged and incubated with 10 µl of affinity chromatography resin (Invitrogen) for 1hr at 4 °C.

3. Resin was then washed extensively with appropriate wash buffer.

4. After washes, 100 µg of prey protein was incubated with the bait the bait protein bound to the beads in a final volume of 200 µl in wash buffer.

5. After 3 hr incubation under agitation at 4 °C, beads were extensively washed for four times with wash buffer and boiled in 30 µl SDS loading buffer. Samples were then analyzed by 12% SDS-PAGE and coomassie blue staining.

3.2.4.4 N-terminal sequencing

The cleavage products of the enzyme-substrate reaction were analyzed through N-terminal sequencing to find specific sites of cleavage. The proteolytically degraded bands of the substrate were transferred electrophoretically onto a polyvinylidene difluoride (PVDF) membrane (pore size 0.45 µm, Millipore Corporation, Billerica, MA, US) using wet transfer apparatus in 1X transfer buffer at a constant voltage of 20V for 10 hrs. Transferred proteins were stained with 0.2% amido black in 50% methanol, followed by destaining in 50% methanol. The stained fragments were excised from the membrane and were air-dried. The N-terminal five amino acid residues of each fragment were identified using ABI 494 Protein sequencer, Tufts university core facility, Boston.

3.2.5 Computational methods

3.2.5.1 In silico preparation of models for mature HtrA3, HtrA3 N-SPD and its mutants

The structure of HtrA3 (PDB ID: 4RI0) was retrieved from Protein Data Bank [2]. Since PDZ domain was missing from chains B and C of the trimeric HtrA3 structure, these were modeled with the help of Prime (Prime, Schrödinger, LLC, New York, NY, 2018) tool using the PDZ domain from chain A as a template [118]. In addition, few loop regions, such as ^{130}LHQLS^{134} in chains A and C, as well as ^{163}HPLFGR^{168} and ^{277}AQREGRELGLRDS^{289} in all the chains were missing in the HtrA3 crystal structure. These missing regions were modelled and refined using loop filling and protein preparation programs in Prime (Prime, Schrödinger, LLC, New York, NY, 2018) [118, 119]. Loop filling was done on the basis of permissive dihedral angle values for different residues, followed by repetitive rounds of sample loop clustering, optimising the side-chain, and energy minimization of the loops [118]. For the preparation of HtrA3 N-SPD monomer (130 - 340 aa), chain B from the modelled structure was used after removing the hinge region and the PDZ domain. The modelled structure was further subjected to two subsequent mutageneses in all the three chains using PyMOL (The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC), where F142 and F255 were mutated to aspartic acid to understand their importance in trimerization [3]. A305 amino acid residue in all these modelled proteins (phenylalanine mutants, HtrA3 N-SPD, and mature HtrA3) was changed to serine to obtain the catalytically active form (S305 active site residue) prior to MD simulation analysis.

3.2.5.2 Peptide library designing and molecular docking study of HtrA3 and its binding partners

For molecular docking, receptor binding pocket needed to be identified. Hence, the generated HtrA3 trimeric model was subjected to SiteMap (SiteMap, Schrödinger, LLC, New York, NY,

2018) analysis to generate a list of four putative binding pockets. These pockets were scored based on their volume, hydrophobic and hydrophilic characters, exposure to solvent and degree to which ligand might donate or accept hydrogen bonds. The top ranked site based on the site score was selected for further docking study and named as the selective binding pocket (SBP_HtrA3). The generation of peptide library was accomplished using two parallel approaches. Firstly, four known HtrA3 substrates were collated from literature, namely β-casein, actin, α-tubulin and vimentin [88]. Domain signature peptides from each substrate were extracted using ScanProsite. α-tubulin (^{142}GGGTGSG^{148}), β-casein (^{109}CPRVYCENQPMLPIGLSDIPGEAMVKLYCPKC^{140}), vimentin (^{397}IATYRKLLE^{405}) and actin (^{54}YVGDEAQSKRG^{64} and ^{357}WISKPEYDE^{365}) were chosen as signature peptides and a peptide library was prepared using all possible heptameric peptide combinations [120]. Among these, ^{131}AMVKLYC^{137} which is a part of the β-casein peptide was chosen as the reference peptide on the basis of docking score obtained from docking of HtrA3 with the above-mentioned peptide library.

Secondly, we tried to form a mutation-based peptide library using affinity maturation and residue scanning tool in Bioluminate (Bioluminate, Schrödinger, LLC, New York, NY, 2018). The reference peptide, AMVKLYC (which was the top ranked one from the previous peptide library) was subjected to mutations where each position of the heptamer was replaced with all possible 20 amino acids. Top 1000 mutated peptides were obtained on the basis of binding affinity and stability with SBP_HtrA3 to prepare the second set of peptide library. An 11 mer peptide having the sequence LNTPLNNPKNN that does not match with the consensus SBP_HtrA3 binding peptide pattern was used as a negative control. All library peptides were built *in silico* using 3D builder (Schrödinger, LLC, New York, 2018). After that, they were pre-processed and energy

minimised using protein preparation wizard (Schrödinger, LLC, New York, 2018) and then docked with HtrA3. Energy minimised HtrA3 structure was prepared for docking and grid file was generated using the protein structure and SiteMap output (Sitemap, Schrödinger, LLC, New York, NY, 2018). Peptide docking tool from Bioluminate was used where the docked poses were scored on the basis of MM-GBSA (Molecular Mechanics using Born and surface area continuum) scoring pattern. This score was used to filter out energetically less favourable peptide poses and get a subset of best possible peptides for further studies [121].

3.2.5.3 MD simulation studies of bound and unbound and HtrA3-wildtype and its variants

Peptide bound and unbound HtrA3-wildtype, its phenylalanine mutants, unbound HtrA3 N-SPD and peptide-bound HtrA3 N-SPD were subjected to MD simulation using GROMACS 2018 where AMBER99sb-ILDNP force field was used to generate topology and parameter files [122, 123]. Each system was surrounded by a truncated octahedron box of TIP3P water molecules with the nearest distance from the protein to the box boundary being no more than 10 Å [124]. Since, net charges for all the systems were less than zero, adequate number of positive (Na^+) ions were added to each system for neutralization. Each system underwent one round of steepest-descent minimization and one round of conjugated gradient of 5000 steps [125]. Particle-mesh Ewald method (PME) was used for the calculation of electrostatic interactions and cut-off for Lennard-Jones interactions was set at 10 Å [126]. Throughout the simulation, to maintain constant temperature and pressure, V-rescale temperature coupling (a modified version of Berendsen thermostat) and Parrinello-Rahman pressure coupling were used [127, 128]. For attaining favorable orientation of water molecules and Na^+ around the systems, they were subjected to equilibration, performed under NVT (N=number of particles, V=system's volume, T= absolute Temperature) or isothermal-isochoric ensemble for 1 ns where the whole system

was heated at 300 K [129]. This was followed by NPT (N=number of particles, P=system's pressure, T= absolute Temperature) or isothermal-isobaric equilibration for 2 ns where pressure was kept constant at 1 atm with isothermall compressibility of solvent at 4.5e-5 atm^{-1} and constant temperature of 300 K [129]. During equilibration, LINCS (LINear Constraint Solver) constraint algorithm was used to apply position restraining force on all the heavy atom bonds present in the system [130]. Finally, 100 ns MD simulation (production run) were conducted for each system under NPT ensemble where coordinates were saved in every 1 ps of time interval. The resultant trajectories from the MD simulations were analyzed using energy, rms, rmsf, gyrate, sham, distance, bar and trjconv packages from GROMACS 2018 [122]. Clusters were generated at 100 ps intervals throughout the simulation. Using free energy landscape, the cluster representing the lowest energy conformers was extracted and analyzed on the basis of cross-relation plot between RMSD (Root Mean Squared Deviation) and radius of gyration [131]. Comparisons among bound and unbound wildtype HtrA3 and its phenylalanine mutants were done based on their overall calculated RMSD, residue-wise RMSF (Root Mean Squared Fluctuation) values, number of intramolecular hydrogen bonds and changes in the catalytic triad residues (H191, D227 and S305) distances [2]. For the comparison of peptide-bound and unbound HtrA3 N-SPD, overall RMSD as well as domain-wise RMSD were taken into consideration. All the graphs were plotted using XmGrace (ftp://plasma-gate.weizmann.ac.il/pub/grace/).

3.2.5.4 Protein-protein docking studies between HtrA2 and DUSP9

Docking studies were used to predict the binding interface and the residues participating in protein–protein interactions. Docking studies between the HtrA2 (receptor) –DUSP9 (ligand) was performed using the ClusPro 2.0, fully automated web server for the prediction of protein–

protein interactions. First, PIPER, a rigid body docking program was run, based on a novel Fast Fourier Transform (FFT) docking method with pair-wise potentials. Second, by using a clustering technique for the detection of near native conformations and by eliminating some of the non-native clusters, the 1000 best energy conformations were clustered, and the 30 largest clusters were retained for refinement. Third, by short Monte Carlo simulations, stability of these clusters were analyzed, and by the medium-range optimization method SDU (Semi-Definite programming based Underestimation), the structure refinement was done. The docked models based on several parameters such as electrostatic, hydrophobic, van der Waal-electrostatic and balanced interactions were generated. The structures based on the balanced type of interaction were studied, as it favors electrostatic, hydrophobic and electrostatic interactions which were preferred for docking. The interactions observed in these docked conformations were visually examined using the software PyMOL. The binding interface residues were evaluated using the PDBsum generate server with the default cut-offs.

3.2.6 Mammalian cell culture techniques

3.2.6.1Maintenance of cell lines

Cell lines were grown on culture dish (BD-Falcon) in DMEM supplemented with 10% FBS at 37ºC, 5% CO_2 and in a humid chamber.

3.2.6.2 Transient transfection

Cells were transfected by Lipofectamine-based technique according to manufacturer's protocol.

3.2.6.3 Cell lysis

Procedure:

1. 500 µL chilled NP40-lysis buffer was added to 90 mm plates containing 80-90% confluent cells.

2. Cells were then collected and incubated on ice for 30 min.

3. The suspension was centrifuged at 13000 rpm for 30 min at 4ºC and the supernatant was collected.

4. The protein concentration of this lysate was measured by Bradford protein estimation as described earlier.

3.2.6.4 Immunoprecipitation

Procedure:

1. 20 µL of Protein-G sepharose beads was washed twice with 1ml chilled buffer (NP-40 lysis buffer without NP-40) by gently inverting the tubes 4-5 times.

2. The beads were centrifuged at 1000rpm for 2 min at 4°C, wash buffer was discarded and beads were resuspended in 1 ml of fresh wash buffer.

3. 2-4µg of Ab/IgG isotype was added to 20µL of Protein-G sepharose beads and kept for binding for 2 h/overnight at 4°C on a rotator.

(IMPORTANT: Do not follow this step for M2-agarose bead and anti-HA-beads).

4. Following day, the cells were harvested, trypsinized and lysed by NP-40 Lysis Buffer.

5. Required amount of cell lysate was added to 20 µL of to previously prepared Ab/IgG bound Protein-A sepharose beads and the total volume was made upto 1mL with lysis buffer.This mixture was kept for binding for 2h at 4°C on a rotator.

6. Later, suspension was centrifuged at 1000 rpm for 2 min at 4°C and the supernatant was carefully separated.

7. The beads were washed (4 times) with chilled wash buffer by gently inverting 3-4 times and centrifugation at 1000 rpm for 2 min at 4°C.

8. Finally the beads were re-suspended with 20µL of wash buffer and 10µL of 3 X lamellae

buffer (Importantly without BME) and boiled at 100°C for 5 mins.

9. Then the IP-suspensions and 5-10% input (cell lysate) were loaded on a SDS-PAGE followed by western blot.

3.2.6.5 Western blotting

Western blotting was used to identify the protein of interest from a complex mixture of proteins by probing with specific antibodies. The technique is qualitative and semi-quantitative in nature. The proteins were first separated on SDS-PAGE and then blotted on a membrane like nitrocellulose, PVDF, etc before using two sets of antibodies (Ab), primary and secondary. The secondary Ab is conjugated to enzyme and chemiluminescent substrates (here ECL) to obtain a signal.

Procedure:

1. Cell lysates were run on SDS-PAGE in 1X SDS-running buffer (150 V, 300 mA and for 90 min) till the dye front exits the gel.

2. The samples from the gel were transferred onto a PVDF membrane in 1X transfer buffer at cold condition (120 V, 300 mA for 90 min in a cool environment.

3. Membrane was blocked with 3% BSA or 5%milk for 1h at RT.

4. Primary antibody (appropriately diluted in 1% BSA in TBST) was used for probing.

5. Membrane was washed with 1X TBST at least 4 times for 10 minutes each on rocker.

6. HRP conjugated secondary Ab was used for probing the membrane for 1h at RT.

7. Membrane was washed in TBST and signal detection was done by following manufacturer protocol of ECL-prime.

CHAPTER 4

Characterization of HtrA3 and its binding partner

4.1 Introduction

The available crystal structure of HtrA3 (3.27 Å) represents a trimeric pyramidal architecture similar to its homologs and is proposed to exhibit a similar mechanism of activation [2]. The structure highlights three monomeric chains where the N-terminal regions from each chain form the apex of the pyramid while the PDZ domains constitute the base of the pyramid. Moreover, the three phenylalanine residues (F140, F142 and F255) form a 'lock' through van der waal's interactions and might be crucial for stabilization of the trimer. The active-site pocket is buried 33 Å from the bottom of the pyramid. Another key feature includes a conserved hydrophobic ligand binding FIGI pocket (similar to GLGF motif of DegS) in the PDZ domain which also remains deeply buried and inaccessible to binding partners until conformational changes stimulates the groove to open up. The GLGF motif in HtrAs is known to be a part of a non-canonical substrate binding site which is allosterically regulated [4, 5]. These structural complexities hint towards a complex mode of regulation which might involve allostery as also observed in all its homologs. Though the crystal structure provides an overall idea of the structural organization of HtrA3, it fails at defining the regulatory and the specificity/substrate binding loops and consequently falls short of delineating the intrinsic conformational dynamics and structural reorganizations that mediate HtrA3 activation.

HtrA3 is prominently involved in cell signaling and apoptosis [7, 45]. It inhibits the TGF-β signaling pathway and has been reported to cleave several extra cellular matrix (ECM) proteins in vitro such as decorin, biglycan, actin, β-tubulin, vimentin and TCP1α chaperonin, suggesting a role in reorganization of extracellular environment. It has been proposed that HtrA3 brings about apoptosis by cleaving XIAP (X-linked inhibitor of apoptosis) to trigger the caspase pathway [44]. Moreover, its participation in non-classical cell death pathways has also been hypothesized

[78]. Since HtrA3 plays an important physiological role and has been implicated in various diseases such as pre-eclampsia and cancer [10, 94, 132], its mode of activation and regulation needs to be studied in detail to be able to modulate its functions with desired effects for therapeutic benefit. Understanding of its physiological roles is also limited since its intracellular substrates are very few in number. Therefore, with an aim at understanding the subtle structural reorganizations and intrinsic conformational dynamics that lead to HtrA3 activation, we dissected its various domains to decipher their roles individually as well as in combination in mediating HtrA3 specificity and functions. Additionally, we also explored the interaction of HtrA3 with its binding partner XIAP in order to better understand the role of this protein in mediating apoptosis.

4.2 Characterization of structural properties and conformational stability of HtrA3 and the role of different domains in protein function and specificity

4.2.1 Results

4.2.1.1 Generation of different HtrA3 domains and variants

Different variants and mutants were generated to understand the contribution of individual domains and residues in maintaining the structural integrity and functions of HtrA3 (**Fig. 4.1A**). On apoptotic stimulus, HtrA3 undergoes autocatalytic cleavage of a part of the N-terminal region and exits the mitochondria to reach the cytoplasm through 'maturation'. Wild type HtrA3 represents this 'mature' active form of the protein which is known to bring about apoptosis dependent on its protease activity. SPD-PDZ (HtrA3 SPD-PDZ) and PDZ (HtrA3 PDZ) domains were generated to understand the role of N-terminal region and SPD in structure, stability and protease activity. An N-SPD (HtrA3 N-SPD, PDZ domain deleted) variant was also used to

investigate the role of PDZ in regulating oligomerization, protease activity and stability of HtrA3.

Three N-terminal phenylalanine residues (F140, F142 and F255) from each monomer, which are also a part of the predicted homo-trimerization motifs in HtrA3, were observed to form strong intermolecular van der Waals interactions in the trimeric structure. This 'triple lock' has been proposed to stabilize the trimer (**Fig. 4.1 B**) as well as regulate the enzymatic activity. Residue F142 particularly, when mutated to aspartate was shown to make the protein monomeric in a previous study as well. Hence, these phenylalanine single mutants (HtrA3 F142D, HtrA3 F142A and HtrA3 F255D) were generated for our study to gain insight into the roles of these residues.

A few N-terminal deleted variants of the wild type protein (Δ134 HtrA3 and Δ143 HtrA3) were also constructed to understand the importance of specific N-terminal region residues. Active site mutant (HtrA3 S305A) was generated to be used as a negative control in protease assays and also for various biophysical studies since the wild type protein is prone to degradation.

A highly conserved 'G-Φ-G-Φ motif' (Φ denotes hydrophobic residues) in PDZ domains of HtrA proteins represent a canonical substrate binding site which might be regulated allosterically as observed in HtrA2 (YIGV groove). This sequence is represented by FIGI in HtrA3. Therefore, a FIGI groove mutant (FIAI), denoted here as HtrA3 G358A, was generated to understand the possible role of this groove in substrate binding in HtrA3.

All Primers used for generating the different constructs have been listed in **section 3.1.2 (Primers)** of **chapter 3 (Materials and Methods)**. The methodology for generating the constructs has further been explained in **section 3.2.1.11** of **chapter 3 (Materials and Methods)**.

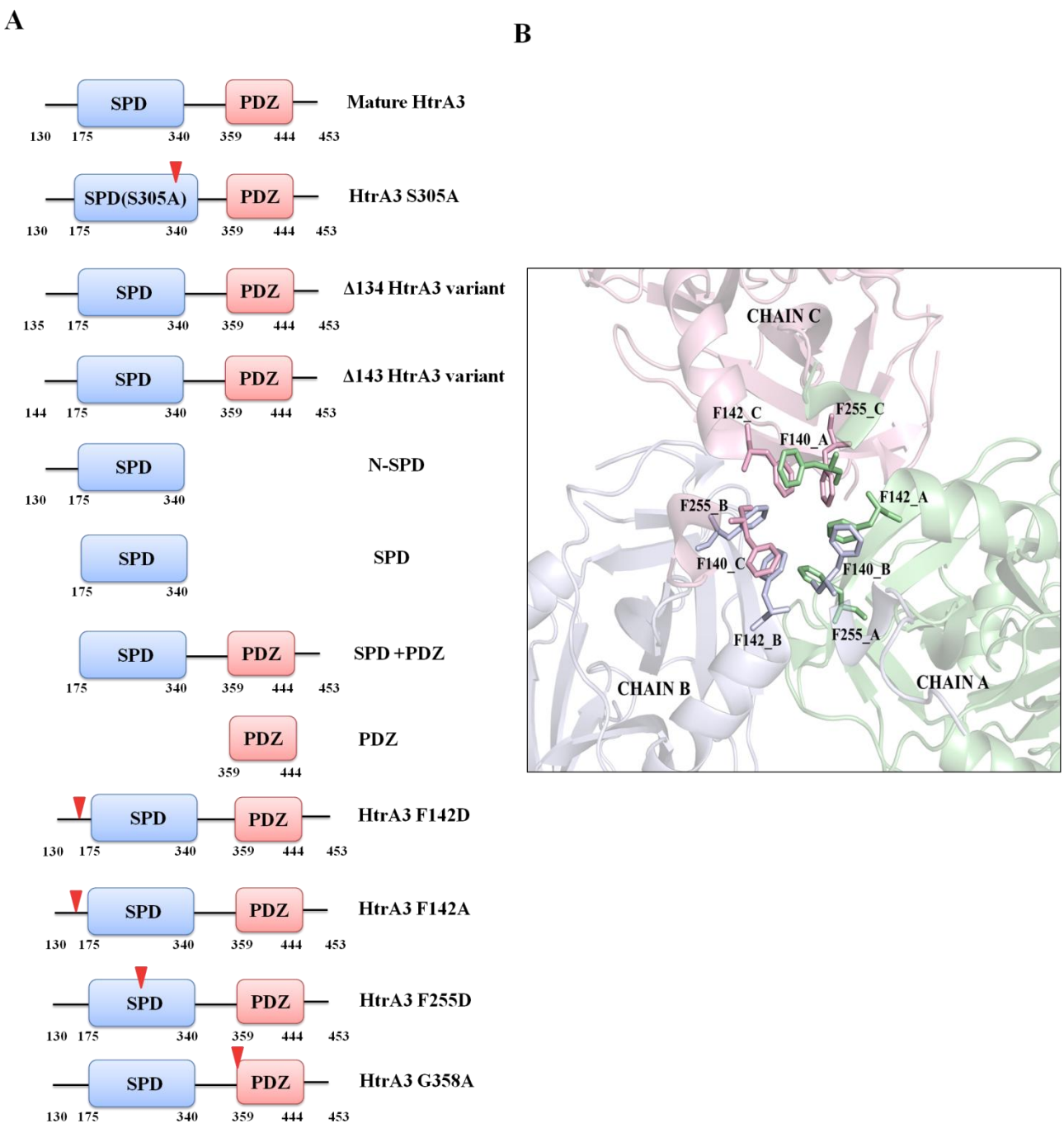


Figure 4.1 Representation of different domains and mutants of HtrA3. *.**A)** SPD indicates serine protease domain (175–340 residues); PDZ domain (359–444 residues); solid lines indicate N-terminal region (130–174 residues) and linker region (341–359 residues); triangles indicate positions of mutations on the respective domains or N-terminal region. **B)** Ribbon diagram of the crystal structure of HtrA3 (PDB accession number 4RIO). Amino acid substitutions in HtrA3 protease are shown in stick models. Ribbon diagram was generated using PyMOL.*

4.2.1.2 Secondary and tertiary structural properties of HtrA3 and its variants

Far UV CD and Fluorescence emission studies were performed to determine the secondary and tertiary structural properties of mature HtrA3 and its variants.

Secondary structure evaluation showed that Mature HtrA3, Δ134 HtrA3 variant, active-site mutant HtrA3 (S305A), monomeric mutant HtrA3 (F142D) as well as its domains SPD and PDZ showed both alpha-helical and beta-sheet characteristics as expected (**Fig. 4.3 A&B**). Compared to HtrA3 wild type, most mutants, HtrA3 F142D, HtrA3 SPD and HtrA3 PDZ showed a decrease in α-helical and an increase in β-sheet characteristics suggestive of significant changes in their secondary structural properties.

Far UV CD was also used for thermal denaturation studies to understand the stability of some of the variants. A temperature range of 20-80°C was chosen to monitor the ellipticity at 222 nm, which was then used to plot the folded-fraction of the respective proteins as a function of temperature to determine the T_m (melting temperature). HtrA3 S305A and HtrA3 PDZ recorded a T_m of 63°C and 46°C respectively (**Fig. 4.3 C**). This could be suggestive of the destabilizing effect the deletion of the N and SPD domains might have a on the protease.

HtrA3 protein has a single, fairly exposed tryptophan residue (W352) in the linker between SPD and PDZ (**Fig. 4.2**). Fluorescence studies were performed with an excitation of 295 nm and emission of 310-400 nm to delineate the tertiary structural properties. Tertiary structural studies with urea denatured HtrA3 S305A (emission maxima ~ 348 nm) and the native protein (emission maxima ~ 342 nm) showed a blue shift of the emission maxima by 6 nm for the native protein thereby demonstrating that the protein has a well folded tertiary structure (**Fig. 4.3 E**). Whereas, both mature HtrA3 (HtrA3 wt) and its active-site mutant (S305A), showed an emission maxima

of ~ 342 nm suggesting that the inactivating mutation did not alter the conformational properties of the protein (**Fig. 4.3 D**).

Far UV CD studies and T_m for other variants could not be performed because these proteins were extremely unstable beyond 50^oC or optimal concentration for CD spectroscopy could not be achieved.

Overall, these studies highlight that the HtrA3 variants mentioned in this section had well folded secondary and tertiary structures. Moreover, these observations underscored the importance of the N -terminal region, SPD and PDZ domains in maintaining protein stability.

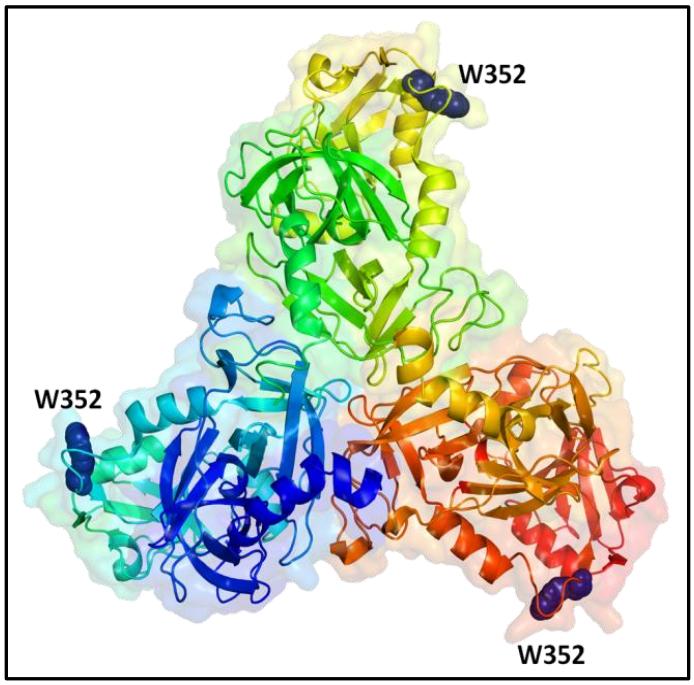
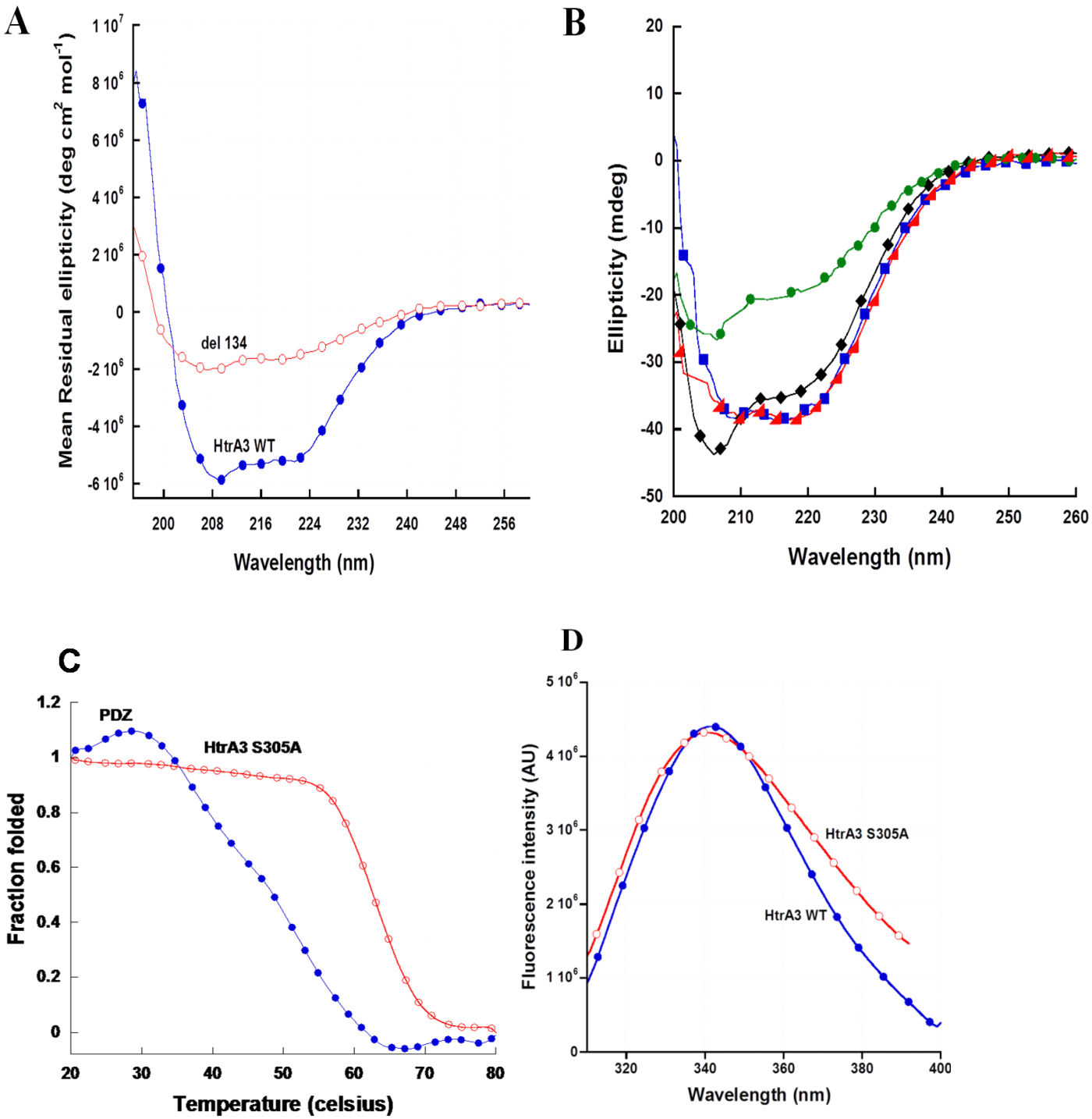


Figure 4.2. Residues W325 in HtrA3 trimer. *The tryptophan residues are present on the hinge between the SPD and PDZ domain on each of the three chains. The model was generated using PyMOL.*


A
Mean Residual ellipticity (deg cm^2 mol^{-1})
1 10^7
8 10^6
6 10^6
4 10^6
2 10^6
0
-2 10^6
-4 10^6
-6 10^6
del 134
HtrA3 WT
200 208 216 224 232 240 248 256
Wavelength (nm)
B
Ellipticity (mdeg)
20 10 0 -10 -20 -30 -40 -50
200 210 220 230 240 250 260
Wavelength (nm)
C
Fraction folded
1.2 1 0.8 0.6 0.4 0.2 0
PDZ
HtrA3 S305A
20 30 40 50 60 70 80
Temperature (celsius)
D
Fluorescence intensity (AU)
5 10^6
4 10^6
3 10^6
2 10^6
1 10^6
0
HtrA3 S305A
HtrA3 WT
320 340 360 380 400
Wavelength (nm)

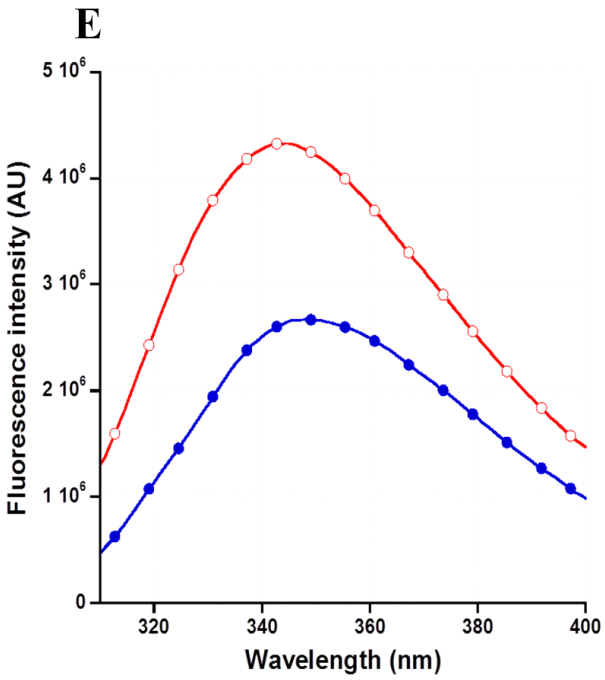


Figure 4.3 Secondary and tertiary structural properties of HtrA3 and its variants. A-B) *Far UV CD between 260-195 nm of (**A**) mature HtrA3 (WT in solid circles), Δ134 HtrA3 (del 134 in open circles) and (**B**) HtrA3 SPD (triangles), HtrA3 PDZ (circles), HtrA3 F142D (squares), HtrA3 S305A (diamond) at 25°C with 20 μM of each protein. **C**) Thermal denaturation curves for HtrA3 S305A (open circles) and PDZ (solid circles) within the temperature range of 25-80°C. The CD at 222nm was selected at each temperature and a plot of fraction folded as a function of temperature is plotted.* ***D-E)*** *Fluorescence emission spectra for HtrA3 and its active-site mutant with excitation at 295 nm and emission between 310-400 nm. Comparision of emission maxima for D) mature HtrA3 (wt, emission max. 342 nm) and HtrA3 S305A (emission max. 342 nm) and E) denatured HtrA3 S305A (solid circles, emission max. 348 nm) and the native protein (open circles, emission max. 342 nm) have been shown.*

4.2.1.3 Role of different HtrA3 domains in oligomerization

Size exclusion chromatography was performed to understand the oligomeric properties of HtrA3 and its variants in order to delineate the roles of each domain/their combination and critical residues in holding HtrA3 together in a trimeric ensemble. Our observations show that the wild type HtrA3, active site mutant (S305A) and PDZ form trimers (**Table 4.1**). In order to understand the structural basis of the PDZ trimer, we used protein-protein docking analysis using PIPER (Bioluminate, Schrödinger, LLC, New York, NY, 2018). Through this, a 10000

conformations of trimeric assemblies were generated which were classified into 53 clusters. Finally, the cluster containing the highest number of conformations was selected as the most favourable trimeric structures energetically. Further analysis of the trimeric ensemble highlighted the residues involved in intramolecular H-bonds, van der Waals and salt bridge interactions. These residues have been shown in following **Table 4.1 and Fig. 4.5**:

Table 4.1 Residues involved in the interacting interface of the PDZ trimer

N-terminal residues	S351, H352, K354, F356, R360 and M361
C-terminal residues	E450, G456, G458, R459, W460 and V461
Other residues	V391, P393 and Q398

Other variants such as HtrA3 N-SPD, HtrA3 SPD-PDZ, HtrA3 F142D and HtrA3 F255D formed monomers. HtrA3 F142A comprised a mix of trimers (~30%) and monomers (~70%) suggesting that the mutation resulted in weak interaction between the three monomeric chains. While HtrA3 variants with intact N-terminal regions were seen to form trimers as described earlier, HtrA3 N-SPD, contrary to the expected outcome was found to be a monomer. This also hinted towards the additional role of the PDZ domain in oligomerization. All the observations put together, besides the N-terminus and the PDZ domain, the 'triple lock' (F140, 142, 255) residues were also found to be important in maintaining protein stability and consequently the oligomeric ensemble of HtrA3. The 'triple lock' residues are held tightly together by van der Waals interactions, therefore, mutation of any of these residues lead to disruption of the interaction network followed by falling apart of the trimeric organization.

The molecular weights of HtrA3 SPD alone and Δ143 HtrA3 variant could not be calculated because these proteins aggregated at higher concentrations. Moreover, the ΔPDZ and Δ143

variants of HtrA3 were found to be very unstable as observed in the thermal denaturation studies, which reiterate the concerted role of the N-terminus and the PDZ in maintaining homotrimerization of HtrA3.

The predicted molecular weights and oligomerization status of the proteins are shown in **Table 4.2**. The representative elution profiles of some of the HtrA3 variants have been illustrated in **Fig. 4.4**.

Protein	**Theoretical molecular weight (kDa) of single chain**	**Calculated molecular weight (kDa)**	**Oligomeric status**
HtrA3 wild type (130-453)	36.2	110.2	Trimer
HtrA3 S305A	36.2	101.3	Trimer
HtrA3 PDZ	11.6	33.4	Trimer
HtrA3 F255D and HtrA3 F142D	36.2	42.3	Monomer
HtrA3 SPD-PDZ	31.0	23.3	Monomer

Table 4.2 Oligomeric properties of HtrA3 and its variants. *Molecular weights were calculated using superdex 200 gel fitration column*

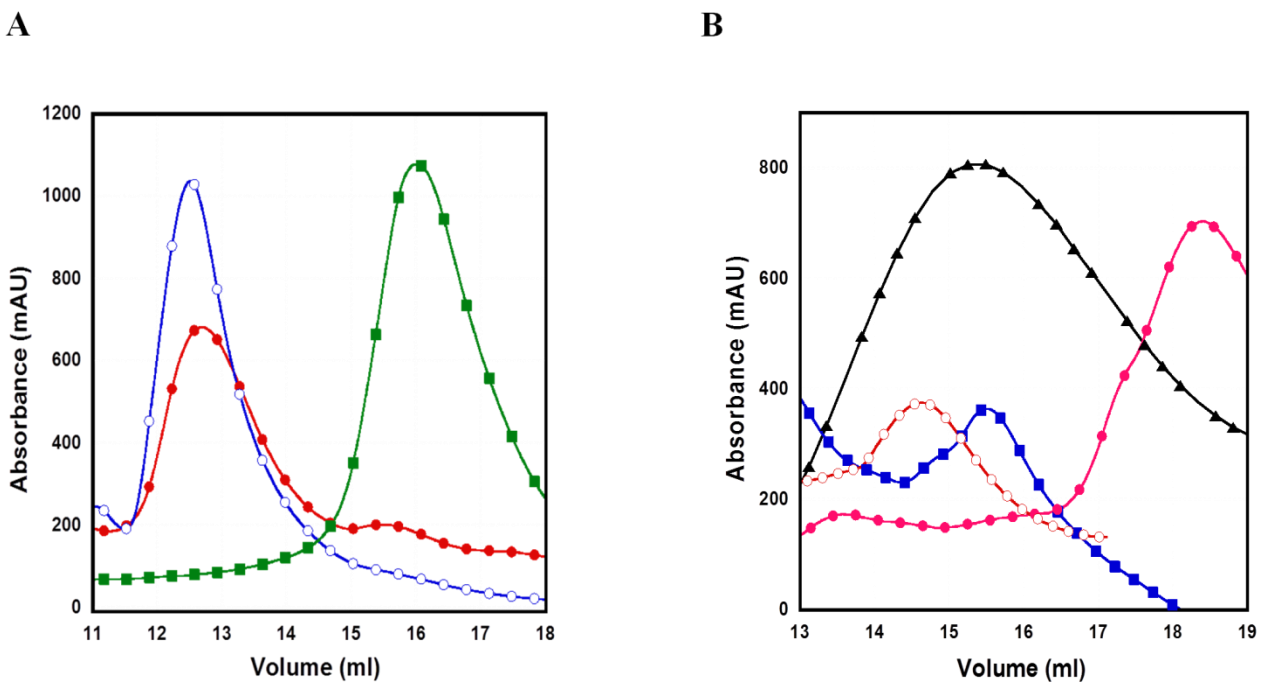


Figure 4.4 Elution profiles of HtrA3 and some of its variants using Superdex 200 gel-filtration column. A) *Plot of absorbance at 215 nm as function of elution volume for mature HtrA3 (solid circles), HtrA3 S305 (open circles) and HtrA3 PDZ (solid squares).* ***B)*** *Plot of absorbance at 215 nm as function of elution volume for HtrA3 F142D (solid triangles), HtrA3 F142A (open circles), HtrA3 F255D (solid squares) and HtrA3 N-SPD (solid circles).*

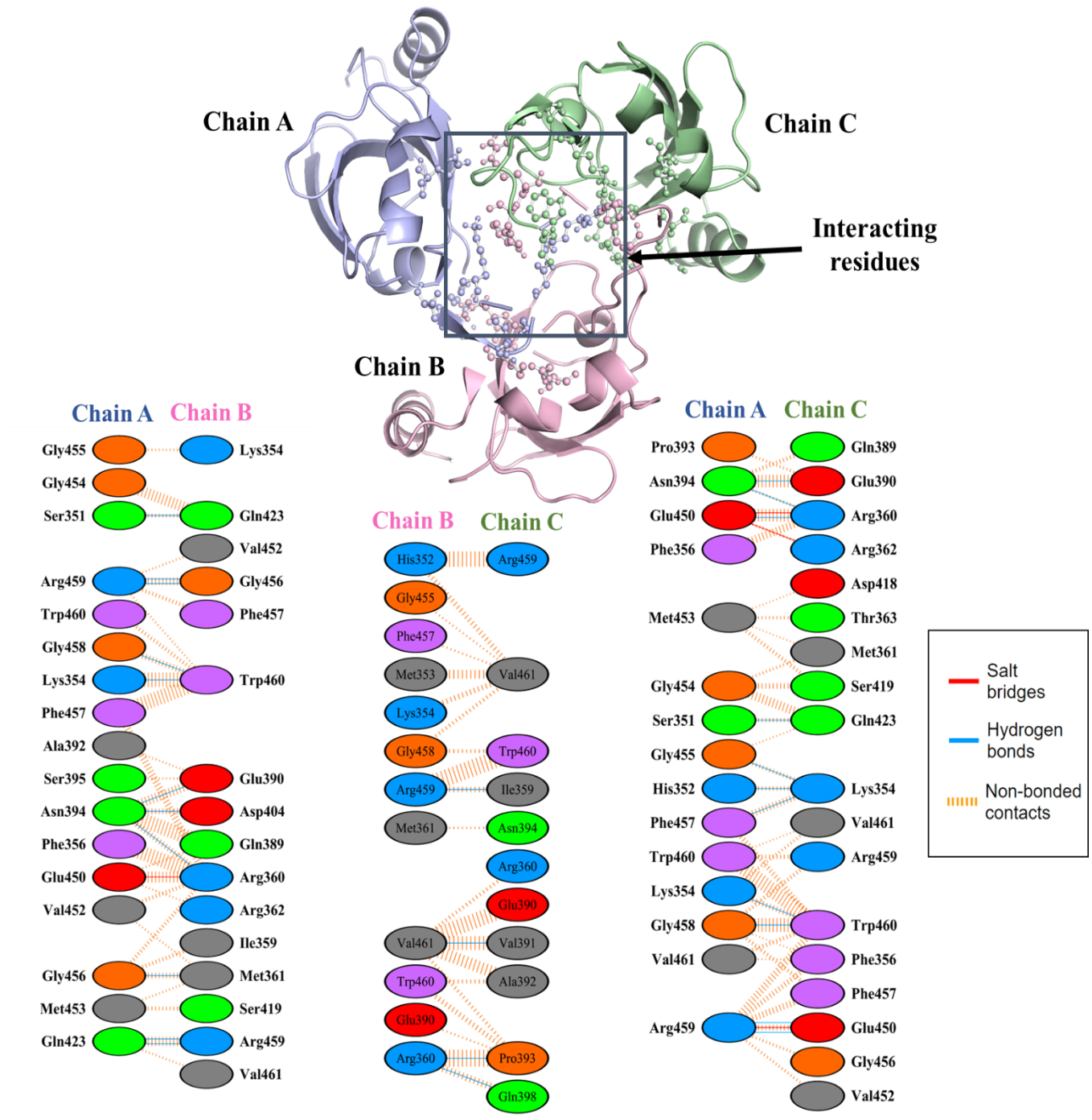


Figure 4.5 Interacting residues in the PDZ trimer. *Cartoon representation of HtrA3 PDZ trimer assembly along with the interacting residue details generated from PDBsum.*

4.2.1.4. Role of individual domains and critical residues in protease activity

Generic serine protease substrate β-casein was used to study the protease activity of HtrA3, its variants and mutants and determine the role of different domains and critical residues in enzyme regulation (**Fig. 4.6**). These observations along with our data on the oligomeric properties of HtrA3 and its variants were used to predict how protease activity is affected as a consequence of structural/conformational changes due to initial substrate binding.

The protease activity for all these variants and mutants was monitored over a time period of 2 hr (**Fig. 4.6**). It was observed that mature HtrA3, its variants (Δ134 HtrA3, Δ143 HtrA3) as well as HtrA3 N-SPD cleaved β-casein. While, mature HtrA3 and Δ134 HtrA3 had comparable activity, Δ143 HtrA3 and HtrA3 N-SPD exhibited activity much less than the wildtype. HtrA3 F142D, HtrA3 F142A, HtrA3 F255D, HtrA3 SPD, HtrA3 SPD-PDZ were completely inactive or showed minimal activity even at high substrate concentrations. These observations suggest that N-terminal region comprising the critical 'triple lock' residues as well as the PDZ domain are important in HtrA3 activity. The decrease in catalytic efficiency of HtrA3 NSPD may possibly arise due the lack of inter-domain PDZ-protease crosstalk for optimal HtrA3 activation or due to the absence of conformational changes mediated by the PDZ that might aid formation of a competent active-site. Δ143 HtrA3 variant, which exhibits a substantial decrease in activity, lacks the F140 and F142 residues that are a part of the 'triple lock' which further underscores the role of phenylalanine residues in maintaining structural integrity, and consequently optimal catalytic activation of HtrA3. Collectively, this emphasizes the role of the N-terminal region, critical N-terminal residues as well as the PDZ domain in substrate catalysis and reiterates the requirement of a complex yet precisely co-ordinated intra- and inter-molecular teamwork of different domains of the trimeric protease to perpetuate its activity.

HtrA3 G358A, the FIGI groove mutant, surprisingly cleaved beta casein (**Fig. 4.7)** highlighting that the FIGI groove may not be required for substrate binding in HtrA3 unlike its homologs. This also suggests the existence of an alternate non-canonical substrate binding pocket as will be discussed in detail in the upcoming **section 4.3**.

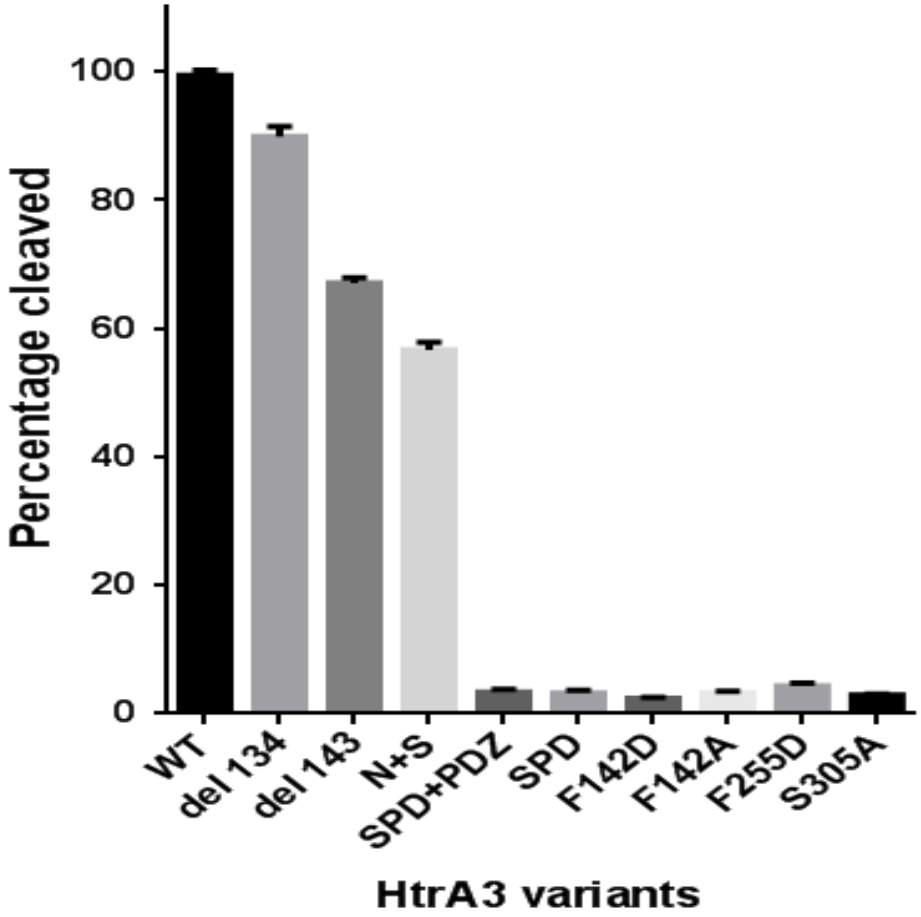


Figure 4.6 Proteolytic activities of HtrA3 and its variants with β-casein as a substrate. *Plot representing percentage of β-casein cleaved by different constructs of HtrA3. Respective enzymes were incubated with 6 μg of β-casein at 37°C for 2 hrs. Reaction samples were resolved by SDS-PAGE and visualized with coomassie brilliant blue staining. Intensity of the substrate remaining after 2 hrs was semi-quantified using Image J. (The error bars are the representations of SE, n=3). Values were normalized with respect to wild type where 100% cleavage was observed. Enzyme and β-casein only used as controls for each assay were also incubated for 2 hr at 37℃.*

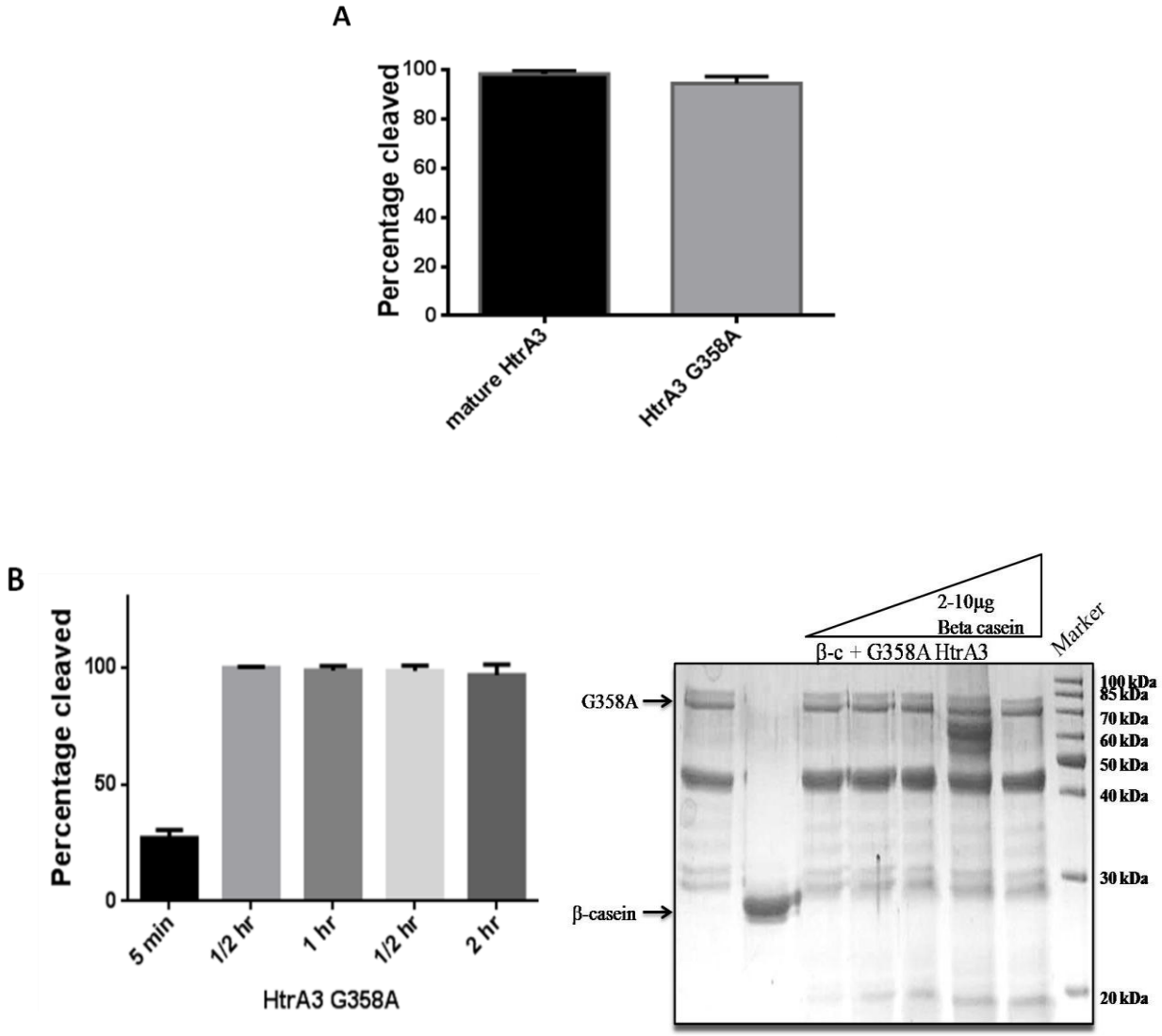


Figure 4.7 Proteolytic activities of HtrA3 WT and HtrA3 G358A. A-B) *Plot representing percentage of β-casein cleaved by HtrA3 WT and HtrA3 G358A. Respective enzymes were incubated with 6 µg of β-casein at 37°C for 2 hrs. Reaction samples were resolved by SDS-PAGE and visualized with coomassie brilliant blue staining. Intensity of the substrate remaining after 2 hrs was semi-quantified using Image J. (The error bars are the representations of SE, n=3). Values were normalized with respect to wild type where 100% cleavage was observed.* ***C)*** *Representative gel based β-casein (β-c/β-cas) cleavage assay over a period of 2 hr at 37℃ with HtrA3 G358A.*

4.2.1.5. Comparative analysis of the structural stability of wildtype HtrA3 and its mutants using MD simulation studies

To understand the roles of the phenylalanine 'lock' residues F142 and F255 in HtrA3 oligomerization, HtrA3 mutants F142D and F255D were studied using MD simulations. Models

for these mutants were prepared using the available HtrA3 crystal structure (PDB ID: 4RI0) as template (methodology described in **section 3.2.5** of **chapter 3**). Structural stability of these two mutants were compared with that of mature HtrA3 by generating RMSD plots of all the atoms in each system over a period of 100 ns (**Fig. 4.8A**)

The plot demonstrated that HtrA3 wildtype and its mutant systems converged at approximately 18 ns. However, beyond which the RMSD values for the wildtype slowly decreased to a stable level of trajectories as compared to the mutants. HtrA3 F255D exhibited the highest deviation from the starting structure, followed by HtrA3 F142D as the trajectories reached near 100 ns (**Fig. 4.8A**). The mean values of RMSD for mature HtrA3, HtrA3 F142D and HtrA3 F255D were 3.2 Å, 3.7 Å and 4.8 Å respectively, indicating greater conformational changes in the mutant proteins, especially F255D, which exhibited a cumulative deviation of 1.6 Å from the wildtype. Compactness of these structures was elucidated by the radius of gyration values plotted against their respective RMSD values in the free energy landscape (FEL) diagrams (**Fig. 4.9**). FELs depicted that the lowest energy conformers for mature HtrA3, HtrA3 F142D and HtrA3 F255D have radius of gyration values of 29.2 Å, 31.1 Å and 32.8 Å (**Fig. 4.9**), respectively. Lower RMSD and radius of gyration values of the wildtype as compared to the mutants further emphasized that mature HtrA3 has a more condensed structure, and intrinsic perturbations in HtrA3 F142D and HtrA3 F255D mutants might be responsible for lesser stable structures.

As the convergence of the three systems occurred at 18 ns, the RMSF plots for their residues were generated taking trajectories between 18-100 ns into consideration. For simplified representation of RMSF data, residue-wise RMSF values for one chain (in this case, chain A) have been plotted. Critical analysis of the plot demonstrated higher fluctuations in various loop regions of the mutant proteins (**Fig. 4.8B**). The highest fluctuation was observed in the linker

region (residues 341-359) of HtrA3 F255D mutant which deviated from mature HtrA3 by ~2 Å (**Fig. 4.8B**). Time-based plot of the secondary structure evolution further implicated increase in the number of disordered regions (predominantly coils and beta turns) in HtrA3 F255D and HtrA3 F142D with respect to the wildtype in Supplementary Fig. S4. In addition, when the number of intramolecular hydrogen bonds (H-bonds) was calculated using filtering parameters such as distance (cut-off value of 3.5 Å) and angle (cut-off value of 35°), a decrease in values was observed. The mean H-bond count decreased from 542 (in mature HtrA3) to 527 (in HtrA3 F142D) and 521 (in HtrA3 F255D) respectively (**Fig. 4.8C**).

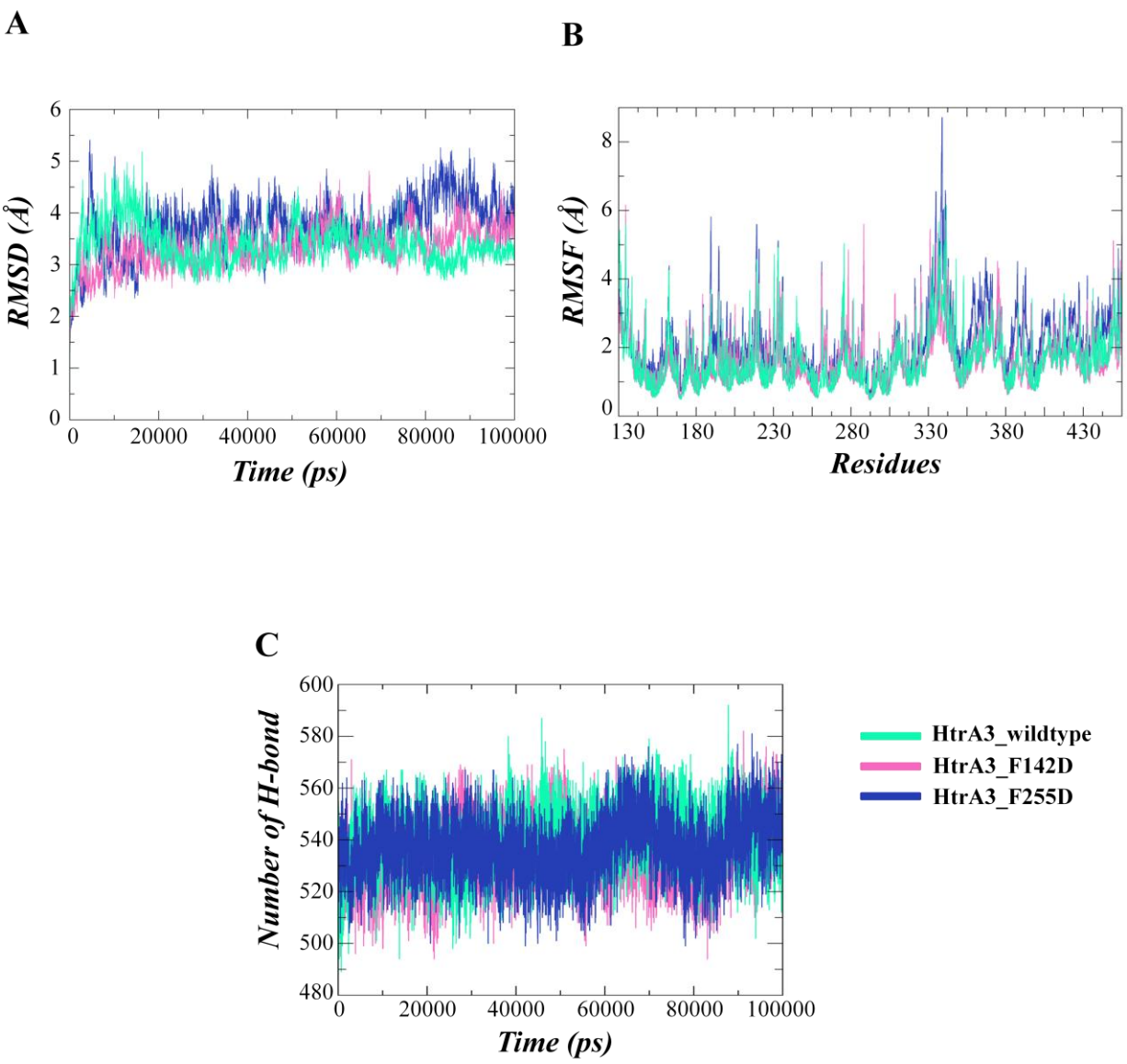


Figure 4.8 In silico comparison of the stability of HtrA3 wildtype and mutants. *Structures of HtrA3 wildtype (green), HtrA3 F142D (pink) and HtrA3 F255D (blue) were analyzed using* ***A****) Root mean squared deviation (RMSD) plot where the trajectories throughout 100 ns MDS run represent structural perturbations;* ***B****) Root mean squared fluctuation (RMSF) plot indicating structural fluctuation of each residue;* ***C****) Hydrogen-bond plot where the changes in the total number of hydrogen bonds throughout the 100 ns MDS run are represented.*

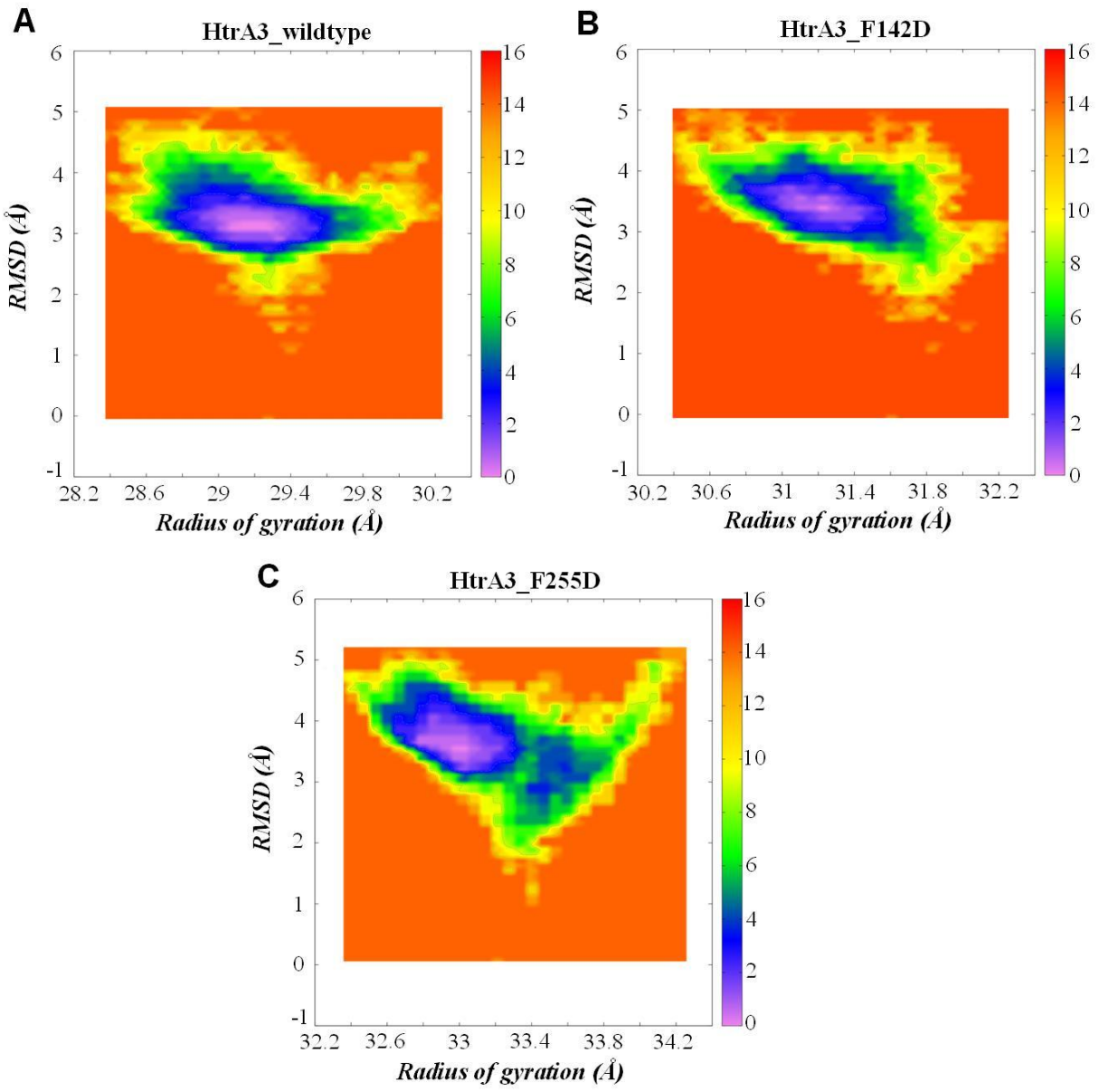


Figure 4.9 *Free energy landscape of HtrA3 wildtype (mature), HtrA3 F142D and HtrA3 F255D representing a cross-relation plot between RMSD and Radius of gyration of different conformations (generated during MDS run) where each conformation is coloured in rainbow spectrum from violet (the lowest energy) to red (the highest energy) on the basis of its Gibb's free energy value.*

4.2.1.6. Cα atomic distance analysis of the catalytic triad residues

Distances among Cα atoms of H191, D227 and S305, which encompass the HtrA3 catalytic pocket were calculated at 2 fs intervals and plotted against 100 ns of MD simulation run for all

the three systems to understand the effect of mutations in HtrA3 catalytic activity. For a proper catalytic pocket formation in serine proteases, histidine usually moves closer to aspartate while serine moves away from the histidine residue. This results in an open conformation that is sufficient for accommodating substrate molecules as well as maintaining an optimum distance between histidine and serine to facilitate the formation of a metastable tetrahedral intermediate (active form), resulting from the histidine-induced deprotonation of serine residue. However, in case of HtrA3 F142D and HtrA3 F255D, the average distance between H191 and D227 in all the three chains increased within a range of 0.6-1.1 Å when compared to the mature HtrA3 (**Fig. 4.10 A, C and E**). Similarly, a drastic decrease within range of 1.2-3.8 Å was observed when the average distance between H191 and S305 was compared with mature HtrA3 (**Fig. 4.10 B, D and F**). These movements resulted in a closed active site pocket where the relative orientation of the catalytic triad is not conducive of maintaining optimal HtrA3 activity. This analysis further explains the inactivity of F142D and F255D variants of HtrA3 where the catalytic activity is substantially reduced due to disordered catalytic pocket formation.

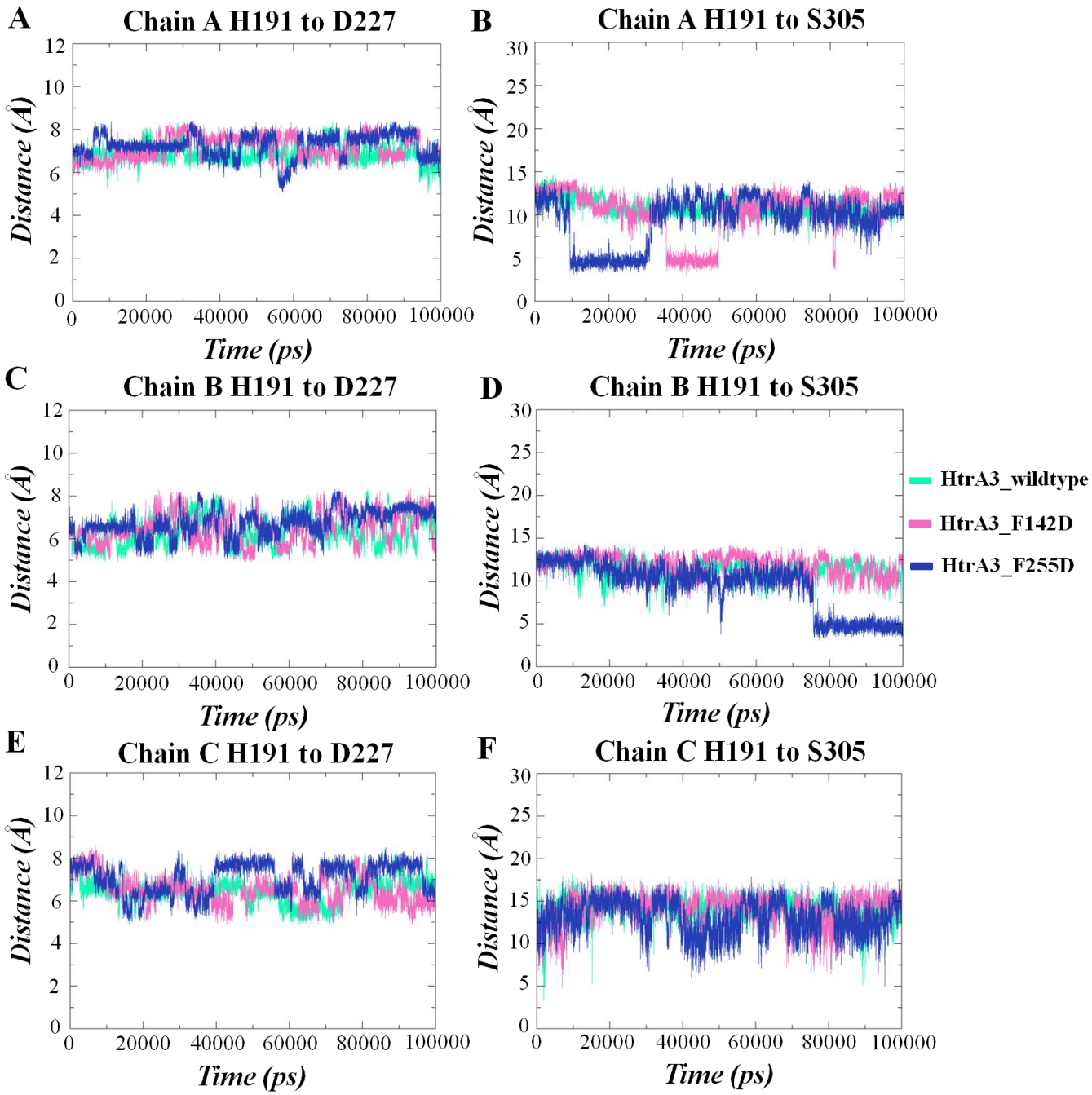


Figure 4.10 Analysis of distances between the catalytic triad residues for HtrA3 and its mutants. *Distance analysis plots showing 100 ns trajectories representing the changes in the distance between H191and D227 in* ***A)*** *chain A;* ***C)*** *chain B; and* ***E****) chain C respectively as well as distance between H191 and S305 in* ***B****) chain A;* ***D****) chain B and* ***F****) chain C respectively. Colour representation for HtrA3 and its mutants are: HtrA3 wildtype in green, HtrA3 F142D in pink and HtrA3 F255D in blue.*

4.2.1.7. Role of temperature in protease activity

Protease assays of HtrA3 were performed as a function of temperature to delineate whether temperature affects the activity of this protease. HtrA3 WT was incubated with 6 µg of β-casein in the temperature range of 30ºC to 60ºC for 10 mins and 30 mins and the cleavage products were analyzed using SDS-PAGE gels. Beta casein was completely cleaved at all temperatures except at 30ºC when incubated for 10 mins (**Fig. 4.11 A and B**) which indicates that with rise in temperature above 30ºC the necessary conformational changes required for efficient proteolytic activity might occur. However we also observed significant degradation of HtrA3 enzyme at temperatures beyond 50ºC, which might be due to protease destabilization (**Fig. 4.11 C**).

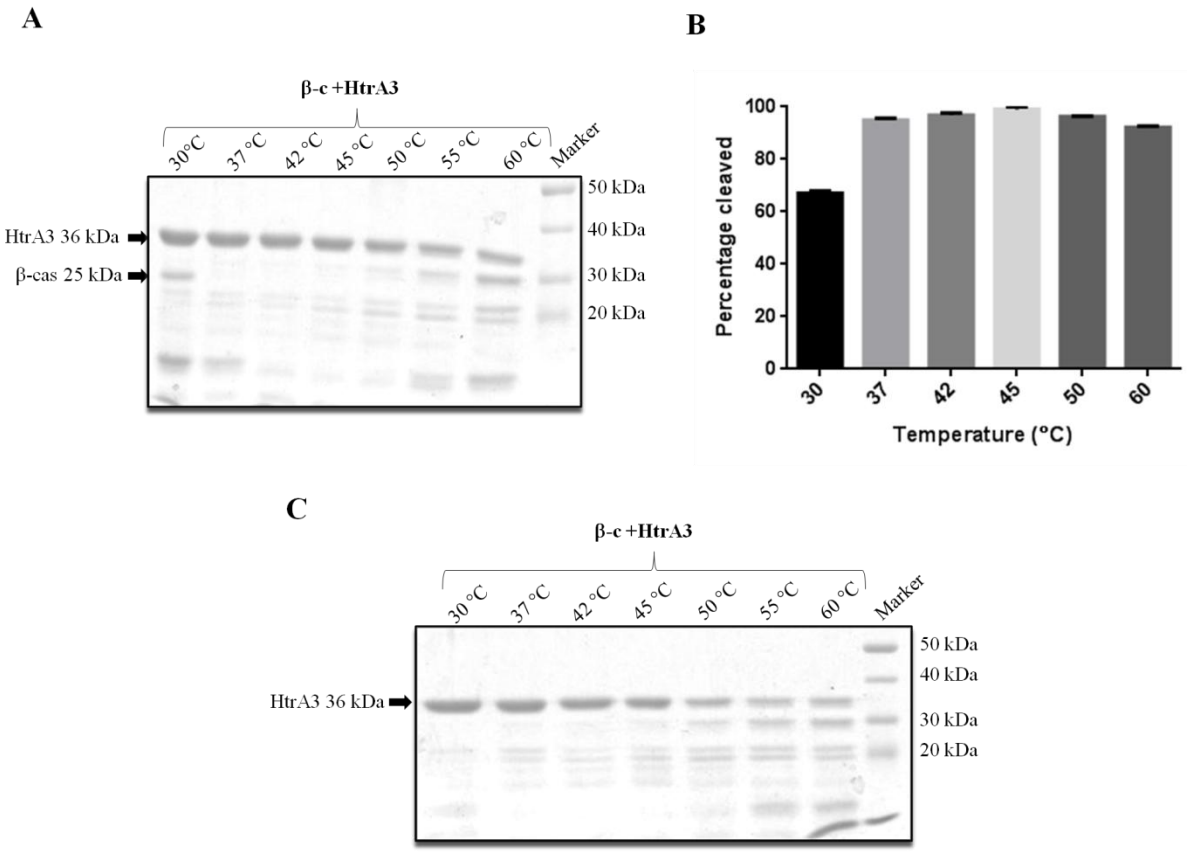


Figure 4.11 Proteolytic activity of mature HtrA3 as a function of temperature (30℃-60℃) using β-casein as substrate. (A and C) *Gel based β-casein (β-c/β-cas) cleavage assays over 30℃-60℃, where reactions were incubated for 10 and 30 mins respectively.* ***B)*** *Plot of*

proteolytic activities of HtrA3 over 30℃-60℃ have been semi-quantified using Image J. (The error bars are the representations of SE, n=3).

4.3 Role of substrate binding and conformational dynamics in regulating mode of HtrA3 action

4.3.1 Results

4.3.1.1 An allosteric mechanism regulates substrate binding and catalysis in mature HtrA3

FITC-labeled generic serine protease substrate, β-casein, was used to study the steady-state kinetic parameters to determine whether HtrA3 exhibits allostery. A wide range of substrate concentration (0-10 μM) was used to measure the initial rates of substrate cleavage which was then plotted as a function of the corresponding substrate concentrations. This plot of reaction velocities were fitted with modified Michaelis-Menten equation incorporating Hill's coefficient suggesting cooperative substrate binding behaviour (**Fig. 4.12**). The steady state parameters were as follows:

Table 4.3 Steady state kinetic parameters for mature HtrA3 with FITC beta-casein

Protein	$K_{0.5}$ **(μM)**	**Hill constant**	**Maximum velocity** V_{max} **(M.s^{-1})**	k_{cat}**(s^{-1})**	**Catalytic efficiency** $k_{cat}/K_{0.5}$ **(M^{-1}s^{-1})**
Mature HtrA3	1.31±0.06	2.98±0.50	1.95±0.20 x 10^{-9}	0.56±0.07 x 10^{-3}	0.42±0.03 x 10^{3}

Comparing the enzymatic parameters of HtrA3 and its homolog HtrA2, we observed that V_{max} for HtrA2 exceeded that of HtrA3 by ~ 2 folds. Moreover, the catalytic efficiency and substrate turnover rate of HtrA2 was also more than that of HtrA3. However, Hill constant and $K_{0.5}$ value for both HtrA2 (HtrA3 2.8±0.2 and HtrA2 $K_{0.5}$ 2.3±0.1 μM) and HtrA3 were not significantly different suggesting similar cooperativity and substrate binding affinity. Overall, our observations prove the existence of an allosteric activation of HtrA3 at the molecular level.

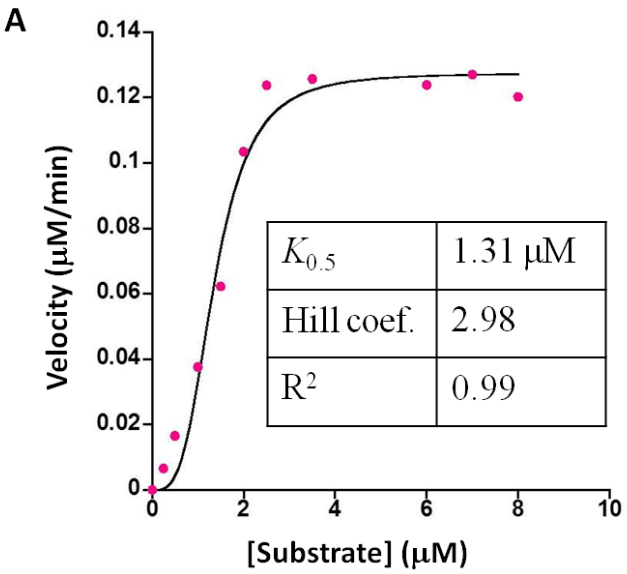


Figure 4.12 Steady-state kinetics of HtrA3 wildtype*. (A) Plot representing the steady-state kinetics of β-casein cleavage by mature HtrA3.*

4.3.1.2 Elucidating the mechanism of allosteric activation in HtrA3

Identification of Selective Binding Pocket and important interacting residues using docking study

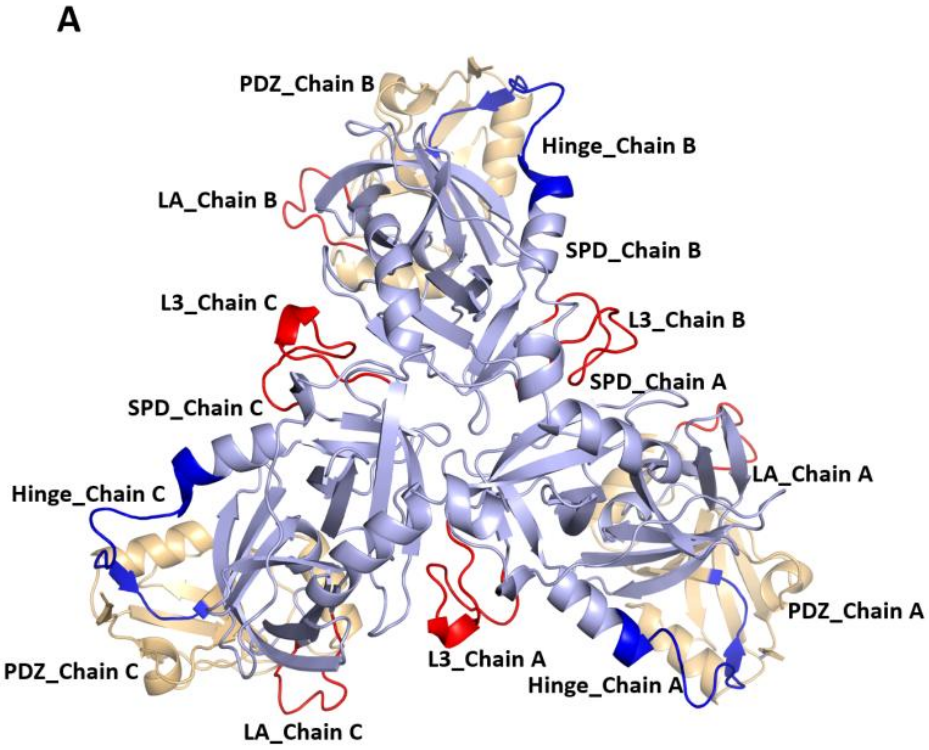
The modeled HtrA3 protein was subjected to SiteMap analysis and four putative binding sites were identified (**Table 4.4**). Among them Site 1 (**Fig 4.13 B**), designated as SBP_HtrA3, was chosen as the best site according to the score. Scoring was done based on the following parameters:

a) Number of available hydrogen donor and acceptor groups available for interaction

b) Size of the pocket depending on its capability to accommodate the heptameric peptides generated from the peptide library analysis

Site 1 was then chosen for further docking and MD simulations. Docking was performed with chain A of HtrA3 trimer model and heptamers from the library were generated as described in **section 3.2.5** in **chapter 3**. Using this method, the best ranked β-casein peptide ^{131}**AMVKLYC**^{137},

was taken as a positive control. A second library was generated by mutating each residue of ^{131}**AMVKLYC**^{137} with all the 20 amino acids one at a time and the best ranked peptides were chosen which have been enlisted in **Table 4.5**. An 11 mer peptide (LNTPLNNPKNN), which did not match the consensus pattern, was used as the negative control. The residues in SBP_HtrA3 which were most involved in the protein-peptide interactions have been enlisted in (**Table 4.5)**. P200 and R202 from the LB loop of the SPD and Q403, D404, G405 from the PDZ were the most common residues (**Table 4.5**, **Fig. 4.14**).



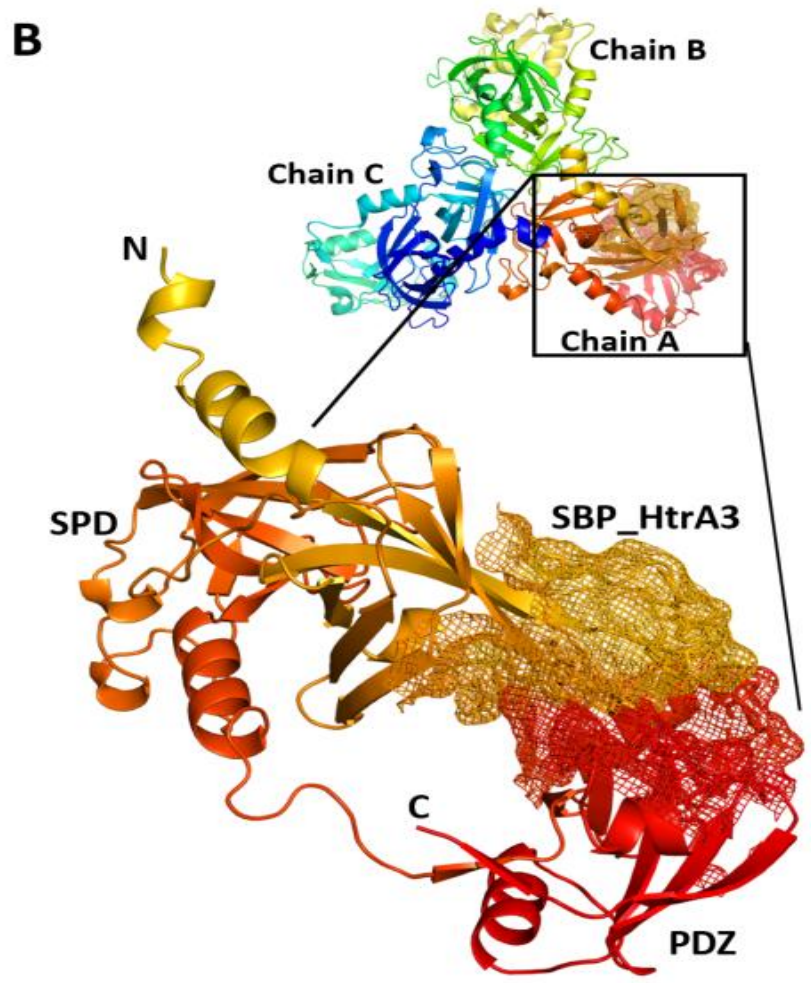


Figure 4.13 Loop filled model of HtrA3 trimer showing selective binding pocket. A) *Ribbon representation of energy minimised loop-filled model of HtrA3 trimer showing domain organisation for all the three chains where hinge regions are indicated in blue colour. Missing loop regions namely LA (163-168) and L3 (277-289) are shown in red colour.* ***B)*** *Selective binding pocket of HtrA3 (SBP_HtrA3) is represented as mesh in chain A of HtrA3 trimer which indicates the accessible area for peptide binding.*

Table 4.4 SiteMap analysis of putative binding sites in HtrA3

SiteMap number	Residues present in the site	Site score*
Site 1	R162, H163, P164, L165, F166, R168, N196, S197, A198, A199, P200, G201, R202, S375, N376, D378, F379, P380, Q389, E390, V391, A392, P393, Q403, D404, G405, D406, I407, R437, R438	1.211
Site 2	L172, S173, S174, G175, N189, H191, V192, V193, S194, S195, N196, Q203, Q204, L205, A217, T218, I219, T321, L322, K323	0.879
Site 3	S247, A248, A277, Q278, R279, D280, D281, R282, L284, D288,	0.806

	D290, D292, Y293, D335, R336, R339	
Site 4	D223, K225, S226, D290, M291, R339, F340, D343, F344,	0.705

**Site score is given on the basis of total volume, solvent accessible surface area (SASA), hydrophobic and hydrophilic nature of the predicted sites*

Table 4.5 Interacting residues in SBP_HtrA3 identified by peptide docking of HtrA3

Chosen peptides for analysis	**Interacting residues**		**Docking Score (Kcal/mole)**
	H-bonds	**Vdw interactions**	
Designed (GMMMRFC)	*Q403*, *D404*, D408, R438	F166, *P200*, P380, D390, *D404*, *G405*	-10.814
Designed (GMTLHHQ)	R162, *R202*, D378, I402, *Q403*, *D404*	*R202*, S375, *G405*	-9.928
Designed (GMQASHC)	R162, *R202*, D378, *Q403*, *D404*	*R202*, N376, *Q403*	-9.796
Designed (GMMAFHV)	*R202*, N376, D378, *D404*	*R202*, I402, *Q403*, G439	-9.753
Designed (GMTLMHC)	D378, *Q403*, *D404*,	*P200*, F379, *D404*	-9.596
β-casein (AMVKLYC)	*Q403*, *D404*, D441	F166, A199, *P200*, L372, S375, R438,	-9.472
Actin (GDEAQSK)	R168, E371, D378, D406	F166, R168, *P200*, S375, N376, *D404*	-8.420
Vimentin (TYRKLLE)	*R202*, *D404*, *G405*	S375, *Q403*, *D404*, *G405*	-8.397
C-terminal β-casein (GPFPIIV)	V391, *D404*	P164, *P200*, N376, *D404*	-7.797
α-tubulin (GGGTGSG)	*D404*, *G405*	F379, *D404*	-6.742
LNTPLNNPKNN	Did not dock with HtrA3		

Common residues in SBP_HtrA3 that interact with most of the peptides are given in italics

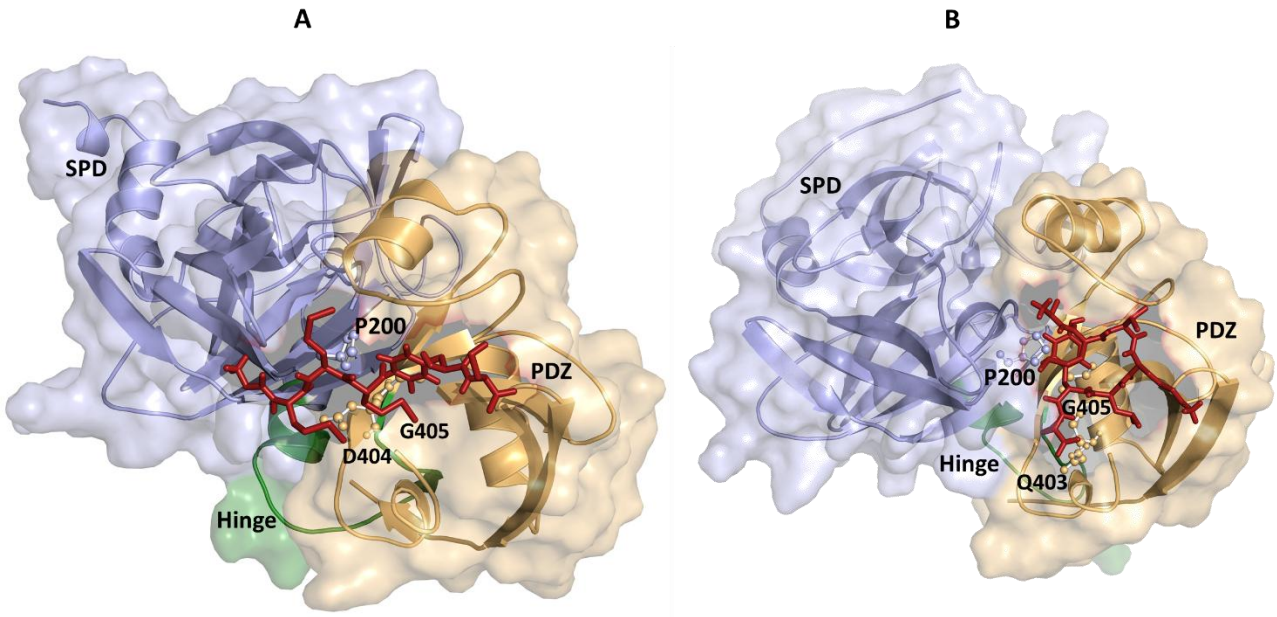


Figure 4.14 Surface representations of peptide-docked HtrA3. *Lowest energy conformer of representative Chain A of (a) HtrA3-GMMMRFC and (b) HtrA3-AMVKLYC complexes. GMMMRFC peptide was obtained using affinity maturation residue scanning analysis. AMVKLYC peptide is a part of beta casein (a known substrate of HtrA3 from literature), which was extracted from the conservative motif region of beta casein protein. Both the peptides are represented by stick (dark red) whose common interacting residues in SBP_HtrA3 are labelled and shown in ball and stick conformation. These residues are highlighted with golden yellow and light purple in PDZ and SPD domains respectively. SPD denotes 'serine protease domain' in both the figures.*

MD simulation analysis shows conformational reorientation of the catalytic pocket of HtrA3 in the presence of peptides

RMSD (Root Mean Squared Deviation) and RMSF (Root Mean Squared Fluctuation) plots of HtrA3-GMMMRFC and HtrA3-AMVKLYC highlighted largest fluctuations in L3 and L1 loop regions (260-305) (**Fig. 4.15 A, B, C and F**) ranging from 2.2 Å-2.5 Å and 1.8 Å in L2 loop when compared to unbound HtrA3. This might suggest that loops L1, L2 and L3 may together regulate allostery by relaying information from the PDZ domain, towards the active site pocket. Hinge region in both the peptide bound complexes showed a deviation of about 1.2 Å from that of unbound HtrA3 (**Fig. 4.15 A, B and C**). Comparative analysis with unbound HtrA3 also

showed that there were no significant overall SPD domain movements (175-340), however PDZ domain showed movements ranging from 0.5 Å to 1.9 Å which could be due to the movements seen in PDZ helices, namely H6 and H7 (**Fig. 4.15 A, B, C, D and E**). Critical assessment of the bound complexes demonstrated deviation in flexible regions that contain the catalytic triad, namely, LB (188-202) and LC loops (223-227) harbouring H191 and D227 respectively (**Fig. 4.15 D and E**).

Both the peptide-bound complexes were subjected to atomic distance analysis to understand the conformational changes in the catalytic triad. The analysis showed that distances between nitrogen (Nδ1) atom of H191 and oxygen (Oγ) atom of S305 decreased by 1.8 Å in HtrA3-GMMMRFC complex, and 1.1 Å in HtrA3-AMVKLYC complex when compared to unbound HtrA3 (**Fig. 4.16 A**). However, the distance between nitrogen (Nε2) atom of H191 and oxygen (Oδ2) atom of D227 increased by 2 Å for HtrA3-GMMMRFC, and 1.7 Å for HtrA3-AMVKLYC with reference to the unbound HtrA3. Apart from this, overall angular distance between nitrogen (Nε2) atom of H191, oxygen (Oγ) atom of S305 and oxygen (Oδ2) atom of D227 increased by 11.3 deg in HtrA3-GMMMRFC complex and 8.3 deg in HtrA3-AMVKLYC complex with respect to unbound HtrA3 (**Fig. 4.16 B**). Moreover, conformational changes were also observed in the relative orientation of oxyanion-hole residues, namely Y302, G303 and N304. The aromatic ring of Y302 exhibited an anti-clockwise flip moving towards the imidazole ring of H191 (**Fig. 4.16 C and D**). This further disrupted the interaction between G260 and D296, resulting in an inward L1 loop movement towards the orthosteric pocket leading to its subsequent stabilization. MD simulation analysis further demonstrated that LC loop that comprises D227 residue of the catalytic triad moved in close proximity to the active site pocket upon peptide binding by disrupting its Van der Waals interaction with β11 strands. All these

movements synergistically shift the LB loop away from the proximal region of SPD thereby bringing the catalytic triad residues within optimal distance to facilitate the assembly of a well-formed catalytic pocket. Overall, post peptide binding in SBP_HtrA3, the reorientation of the catalytic triad residues along with formation of a functional oxyanion hole at the active site makes it compact and stable for further substrate binding and catalysis.

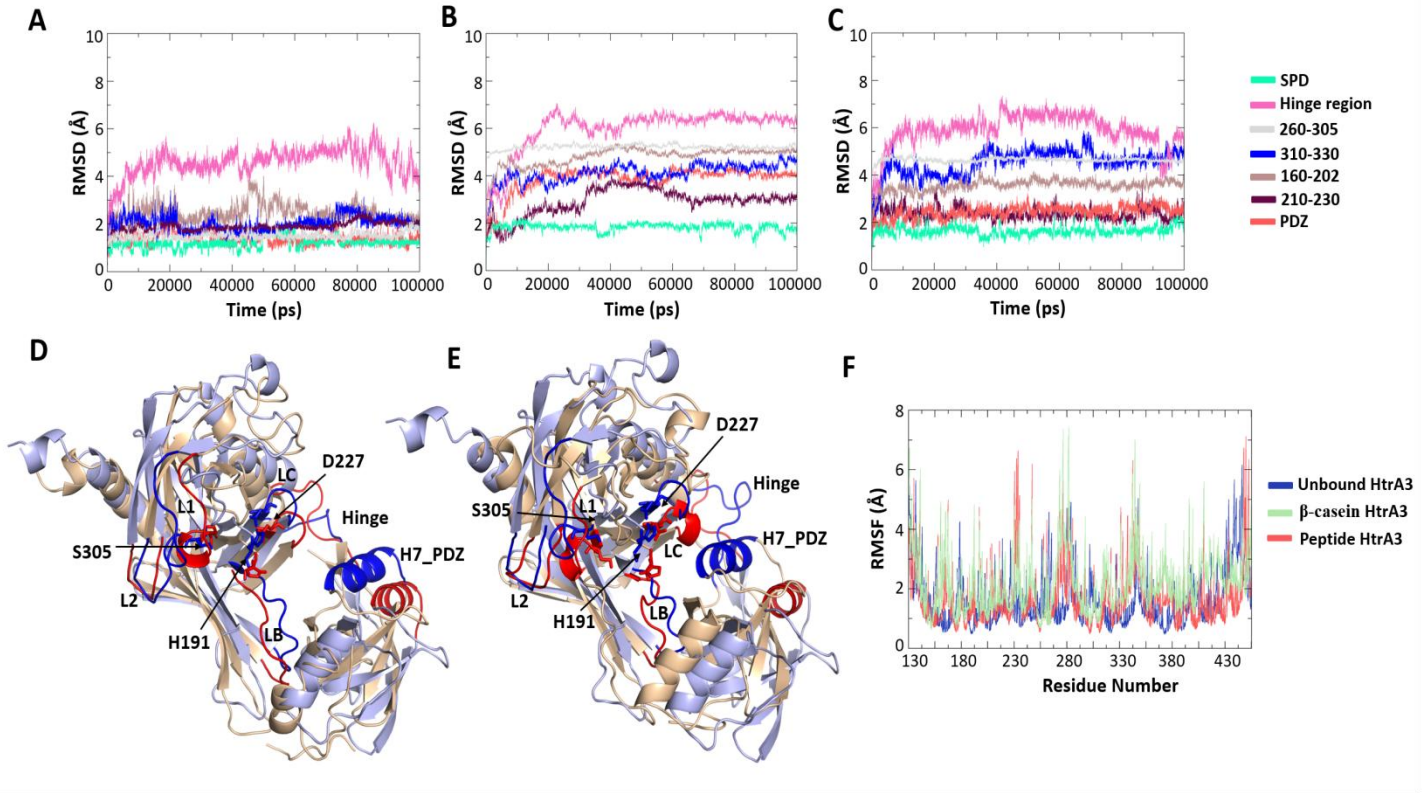


Figure 4.15 RMSD and RMSF plots showing structural fluctuations in the loop regions upon peptide binding at SBP_HtrA3. *RMSD plots of the 100 ns MD simulation trajectory for **A**) unbound HtrA3 (negative control), **B**) HtrA3-AMVKLYC (positive control) and **C**) HtrA3-GMMMRFC) are shown. **D**) Structural alignment of chain A of HtrA3-GMMMRFC (light blue) and unbound HtrA3 (golden yellow) showing the fluctuations in the loops and helices where peptide-bound loops and helices are denoted by red and for unbound they are represented in blue. **E**) Structural alignment of chain A of HtrA3-AMVKLYC (light blue) and unbound HtrA3 (golden yellow) showing the fluctuations in the loops and helices where peptide-bound loops and helices are denoted by red and for unbound they are represented in blue. **F**) Graphical representation of RMSF trajectory for 100 ns in unbound HtrA3 (blue), HtrA3-AMVKLYC (green) and HtrA3-GMMMRFC (red).*

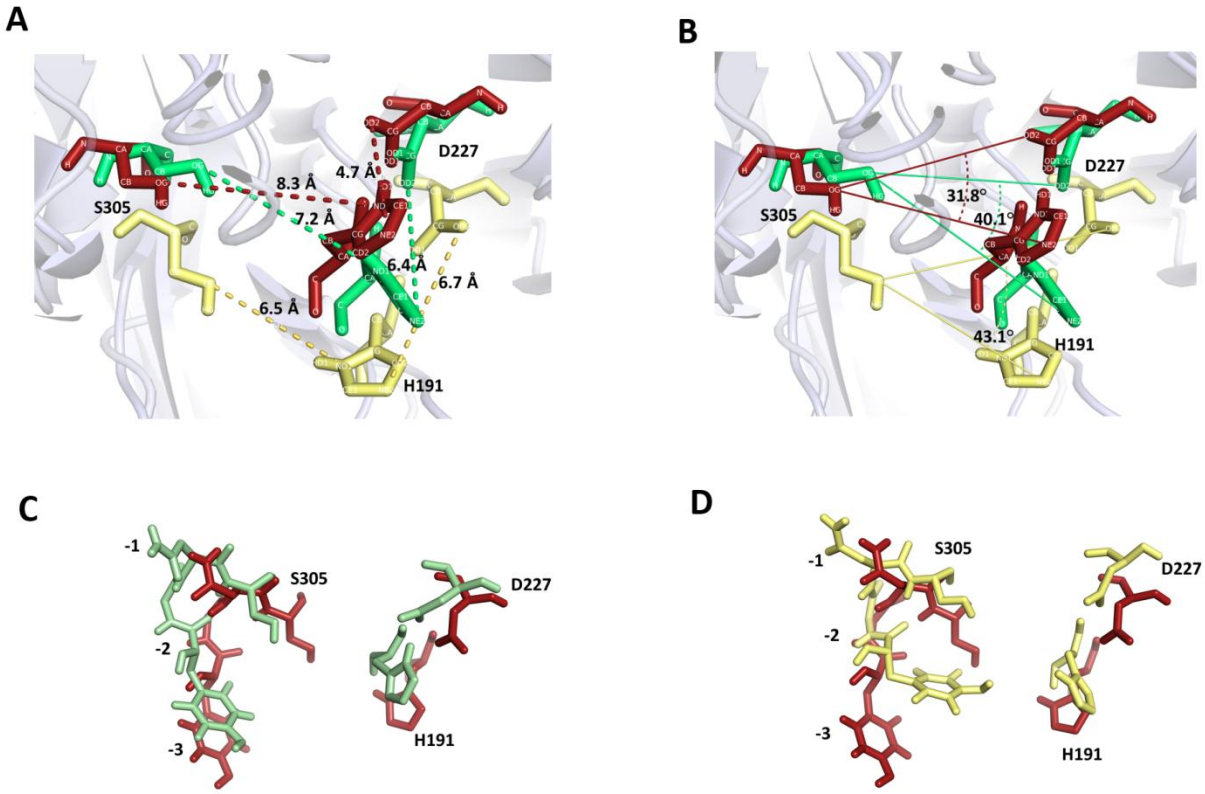


Figure 4.16 Analysis of the interatomic distance and conformational changes at catalytic triad and oxyanion hole residues*. Stick representation showing structural alignment catalytic triad residues in unbound HtrA3 (dark red) with HtrA3-GMMMRFC (yellow) and HtrA3-AMVKLYC (green).* ***A****) Interatomic distances between nitrogen (Nδ1) atom of H191 and oxygen (Oγ) atom of S305 as well as nitrogen (Nε2) atom of H191 and oxygen (Oδ2) atom of D227 are shown.* ***B****) Representation of interatomic angular deviation among nitrogen (Nε2) of H191, oxygen (Oγ) atom of S305 and oxygen (Oδ2) atom of D227.* ***C-D****) Superimposition of the oxyanion hole and catalytic residues showing inward flip of Y302 at -3 position towards the imidazole ring of H191 for both HtrA3-GMMMRFC (yellow) and HtrA3-AMVKLYC (green), when compared to unbound HtrA3 (dark red). The positions of the oxyanion residues namely N304, G303 and Y302 are denoted by -1, -2 and-3 respectively, keeping S305 at 0th position.*

Table 4.6 Interatomic distances and angular deviations of the catalytic triad in peptide-bound and -unbound form of HtrA3

Docked complex	**Nδ1 (H191) - Oγ (S305)**		**Nε2 (H191) – Oδ2 (D227)**		**Nε2(H191) – Oγ(S305) – Oδ1 (D227)**	
	Bound	Unbound	Bound	Unbound	Bound	Unbound
HtrA3-GMMMRFC	6.5 Å	8.3 Å	6.7 Å	4.7 Å	43.1°	31.8°
HtrA3-AMVKLYC	7.2 Å	8.3 Å	6.4 Å	4.7 Å	40.1°	31.8°

4.3.1.3 Allosteric signal propagation and substrate binding in HtrA3 does not recruit the conserved GLGF motif

It was demonstrated earlier by our group that the FIGI groove mutant (FIAI or HtrA3 G358A), cleaved beta casein [14] indicating that this groove might not be required for initial substrate binding unlike other HtrAs. Additionally we wanted to understand if this groove played any role in allosteric signal propagation in HtrA3. The steady-state kinetic parameters were calculated as shown in **Fig. 4.17** and **Table 4.7**. HtrA3 G358A showed cooperativity and substrate binding affinity (Hill constant of 2.93±0.006 and 1.07±0.32 μM respectively) comparable with that of mature HtrA3.

Overall, our findings highlight that the FIGI groove may not be required for initial substrate binding and allostery in HtrA3. We further believe that the novel alternate pocket (SBP_HtrA3), as shown by our *in silico* studies, may be the site for substrate binding.

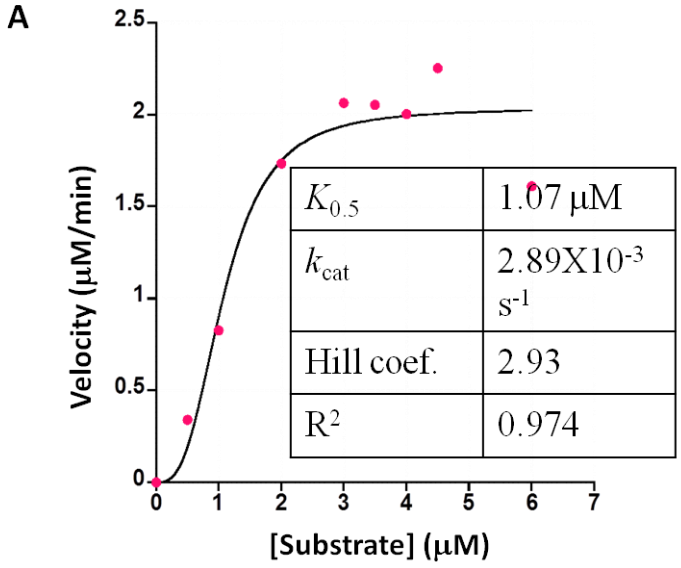


Figure 4.17 (A) Plot representing the steady-state kinetics of β-casein cleavage by HtrA3G358A.

Table 4.7 Steady state kinetic parameters for HtrA3 G358A with FITC beta-casein.

Protein	$K_{0.5}$ (μM)	Hill constant	Maximum velocity V_{max} ($M.s^{-1}$)	k_{cat}(s^{-1})	Catalytic efficiency $k_{cat}/K_{0.5}$ ($M^{-1}s^{-1}$)
G358A (HtrA3)	1.07±0.32	2.94±0.006	3.39±1.1 x 10^{-8}	2.89±0.01 x 10^{-3}	2.68±0.08 x 10^{3}

4.3.1.4. Optimal catalytic activation of HtrA3 is impaired on removal of PDZ domain

Kinetic studies were performed with HtrA3 N-SPD, to understand if the C-terminal PDZ domain participates in allosteric activation of HtrA3 (**Table 4.8**). We had already observed that removal of the PDZ domain leads to a decrease in substrate turnover from our gel based protease assays. The plot of the reaction velocity against the corresponding substrate concentration was fitted with modified Michaelis–Menten equation that accommodates Hill's cooperativity parameters (**Fig. 4.18 A**). On comparison of the steady state kinetic parameters with that of the wild type, we observed that the V_{max}, k_{cat} and $k_{cat}/K_{0.5}$ of this variant were significantly less compared to the wild type mature HtrA3 by ~6.7 x 10^3, 3.6 x 10^3 and 2.8 x 10^3 fold respectively (**Table 4.8**). The drastic decrease in maximum velocity, substrate turnover rates and enzyme catalysis hints toward the presence of a malformed oxyanion hole. The K_m was slightly less than the wild type by 1.2 fold suggesting that deletion of PDZ domain possibly provides greater accessibility to the active-site leading to a the slight increase in substrate affinity for this variant. Additionally, we reported that the Hill coefficient of 1.53±0.03 of this variant is less than that of the mature HtrA3.

To investigate if there is any malformation of the oxyanion hole upon β-casein binding, HtrA3 N-SPD was docked with the β-casein peptide (AMVKLYC) at the allosteric pocket identified previously by our *in-silico* studies. We observed a low docking score of -5.194 kcal/mole (**Fig. 4.18 B**). Moreover, the comparative molecular dynamics (MD) simulation study of the unbound

and bound HtrA3 N-SPD showed no significant reorientation of the oxyanion hole residues to form a pocket conductive for substrate cleavage (**Fig. 4.18 C**). These observations provide a solid basis for the decrease in catalytic efficiency for this HtrA3 variant. Furthermore, exhibition of residual allostery in this ΔPDZ variant might be due to the binding of the β-casein peptide to an altered allosteric pocket (encompassing the residues in the SPD domain alone), since the residues of the PDZ domain which are a part of the pocket are missing in this variant. (**Fig. 4.18 B**).

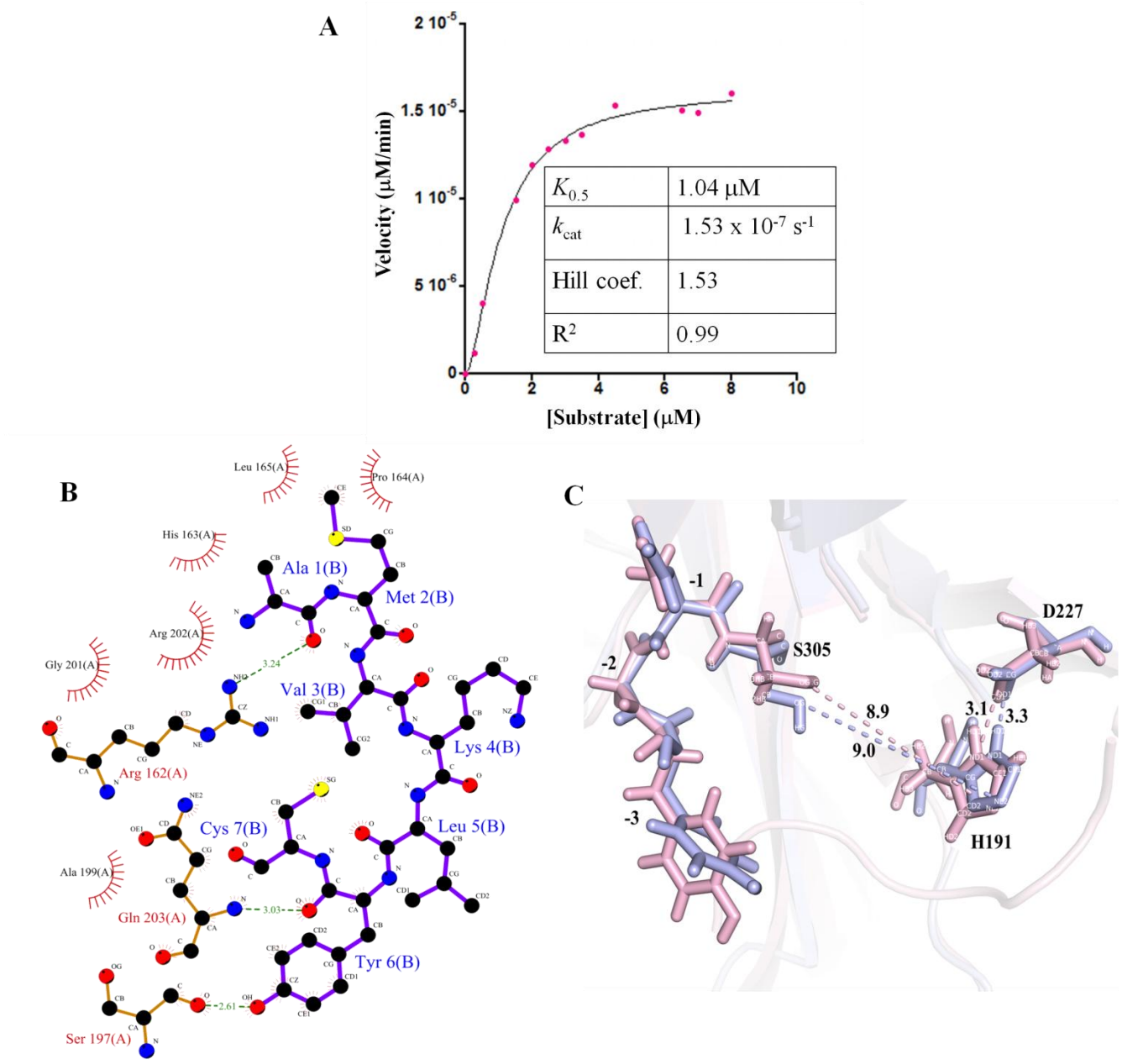


Figure 4.18 Mechanism of catalytic activation in HtrA3 N-SPD A) *Plot representing the activity and steady-state kinetics of β-casein cleavage by HtrA3 N-SPD. For clarity, SEM values have been mentioned in Table 2.* ***B)*** *Ligplot showing the interacting residues of HtrA3 N-SPD (Chain A) with β-casein (AMVKLYC) peptide (Chain B). Green dotted lines indicate H-bond interactions with distance labelled; red semi-circular dash lines indicate van der Waal's interactions.* ***C)*** *Stick diagram showing the alignment of unbound HtrA3 N-SPD (light blue) and peptide-bound N-SPD (pink), representing the oxyanion hole residues along with catalytic residues H191, D227 and S305. Oxyanion residues are marked as -1, -2 and -3 starting from the S305 residue which is considered as 0. Distances among the catalytic residues are represented by dash lines (light blue for unbound and pink for bound HtrA3 N-SPD).*

Table 4.8 Steady state kinetic parameters for HtrA3 N-SPD with FITC beta-casein as the substrate*. This data is the average of three independent experiments.*

Protein	$K_{0.5}$ **(μM)**	**Hill constant**	**Maximum velocity** V_{max} **(M.s^{-1})**	k_{cat}**(s^{-1})**	**Catalytic efficiency** $k_{cat}/K_{0.5}$ **(M^{-1}s^{-1})**
HtrA3 N-SPD	1.04±0.06	1.53±0.03	2.90±0.26 x 10^{-13}	1.53±0.14 x 10^{-7}	0.15±0.02

4.4 Characterization of HtrA3-ligand interaction

4.4.1 Results

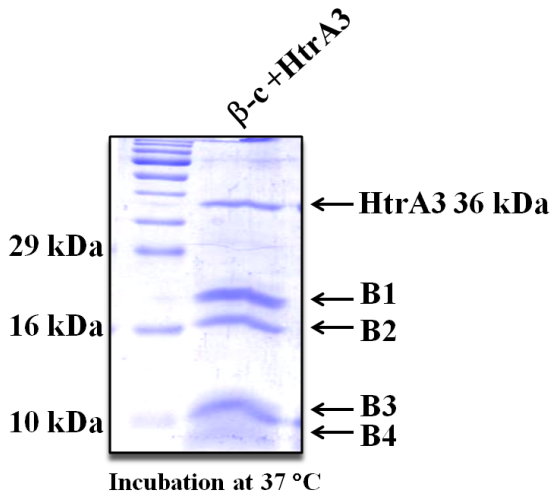
4.4.1.1 Substrate specificity of HtrA3 using beta casein as substrate

To determine the cleavage specificity of HtrA3 proteolytic studies were performed using β-casein as the substrate. The proteolytic products were run on SDS-PAGE gel and four cleaved fragments (between ~5-24 kDa) were separated (**Fig. 4.19**) which were further analysed using N-terminal sequencing by Edman degradation. The four cleavage sites are as follows:

sites	positions
1	15A-16R
2	71Q-72S
3	138Q-139S
4	170V-171M

The specificity profile of the four sites has been highlighted in **Table 4.9** and the comparative substrate specificities with other HtrA homologs have been detailed in **Table 4.10**.

A



1. B1, fragment RELEE, 23.5 kDa
2. B2, fragment SLYVP, 17 kDa
3. B3, fragment SLTLT, 9.6 kDa
4. B4, fragment MFPPQ, 5.8 kDa

B

MKVLILACLV ALALA RELEE LNVPGEIVES LSSSEESITR INKKIEKFQS EEQQQTEDEL QDKIHPFAQT

Q SLVYPFPGP IPNSLPQNIP PLTQTPVVVP PFLQPEVMGV SKVKEAMAPK HKEMPFPKYP

VEPFTESQ SL TLTDVENLHL PLPLLQSWMH QPHQPLPPTV MFPPQSVLSL SQSKVLPVPQ

KAVPYPQRDM PIQAFLLYQE PVLGPVRGPF PIIV

Figure 4.19 Substrate specificity of HtrA3. ***A)*** *Protease assay of HtrA3 with substrate β-casein peptide fragments B1-B4 were excised and sent for N-terminal sequencing.* ***B)*** *Identified cleavage sites are highlighted in red, green, pink and blue respectively. The four cleavage sites are also indicated by black arrows.*

Table 4.9 Substrate specificity of HtrA3. *Sites1-4 of cleavage that were identified. Prime (P) and nonprime (P') residues at these sites have been identified through N-terminal sequencing.*

	P4	P3	P2	P1		P1'	P2'	P3'	P4'	P1' residue
Site 1	**L**	**A**	**L**	**A**	**-**	**R**	**E**	**L**	**E**	**Arg 16**
Site 2	**A**	**Q**	**T**	**Q**	**-**	**S**	**L**	**V**	**Y**	**Ser 72**

Site 3	**T**	**E**	**S**	**Q**	**-**	**S**	**L**	**T**	**L**	**Ser 139**
Site 4	**P**	**P**	**T**	**V**	**-**	**M**	**F**	**P**	**P**	**Met 171**

Table 4.10 Substrate specificity of HtrA family members

Position	**HtrA1**	**HtrA2**	**HtrA3**	**HtrA4**
P1	A,M,I,V,L	M,I,V	A,Q,Q,V	A,M,V
P2	L,T	L,R	L,T,S,T	L,V,T
P3	Q,L	K,R,Y	A,Q,E,P	E,A,P
P4	A,P	L,P,I	L,A,T,P	L,P
P1’	R,S,T	S,A	R,S,S,M	R,G,M
P2’	L,P,W	Y,F	E,L,L,F	E,V,F
P3’	V,M,D	S,P,Y	L,V,T,P	S,P,L
P4’	F	F,Y,S	E,Y,L,P	E,K,P

Our observations suggest that HtrA3, similar to the other human HtrA homologs, has a strong preference for aliphatic residues at P1 (A, V, and Q) position. At the P2 position, aliphatic (L) and polar (S, T) residues are preferred, again similar to HtrA1 and HtrA2 (L). P3 and P4 positions are occupied by aliphatic (L, A), polar residues (T, Q), non polar (P) and acidic (E) residues, which also matches with that of HtrA1 & 2 (A, L, P, Q). For the P1' positions, we observed that HtrA3 preferred aliphatic (M) and polar amino acids (S, R), similar to other family members. While, P2' position was occupied by acidic (E), aromatic (F) and aliphatic (L) residues, P3' position had aliphatic (L, V), polar (T) as well as non polar (P) residues. At P4', HtrA3 showed a preference for aliphatic (L), aromatic (Y), non polar (P) and acidic (E) residues. Comparatively, HtrA1 and HtrA2 showed preference for similar residues at P1-P4 and P1'- P4' sites suggesting that they have similar (but not identical) substrate specificities

4.4.1.2 XIAP is a binding partner and substrate of HtrA3

Pull down studies with HtrA3 and XIAP

Pull down studies were performed with GST fused recombinant XIAP as bait and HtrA3 S305A (active site mutant) as prey as described in **section 3.2.4.3** of **chapter 3**. No pull down was detectable using GST control alone **(Fig. 4.20)**.

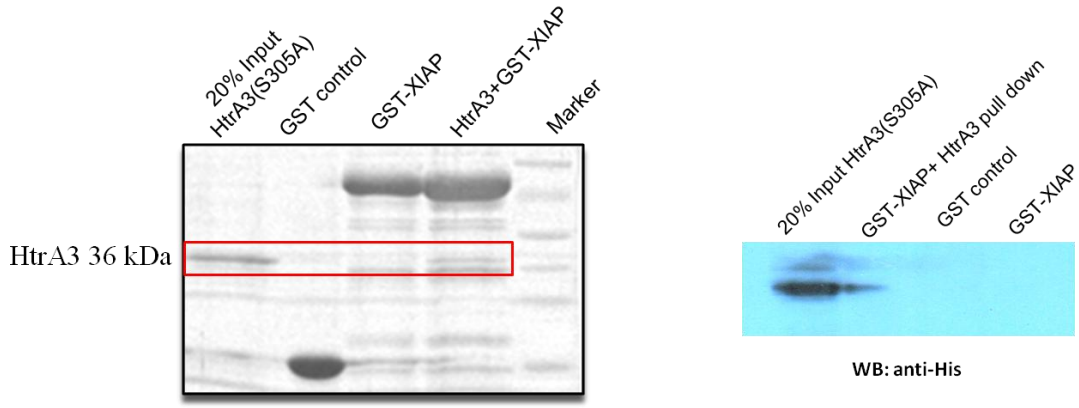


Figure 4.20 Pull down studies with HtrA3 and XIAP. *Pull down was performed using GST tagged XIAP as bait and His tagged HtrA3 as prey proteins. The blot was probed with anti-His to detect HtrA3.*

These studies support the hypothesis that HtrA3 regulates the apoptotic pathway through XIAP protein.

Proteolytic assays with XIAP as a substrate of HtrA3

To address if HtrA3 regulates the turnover of XIAP protein in the apoptotic pathway through binding (as seen in pull down studies) and cleaving, protease activity was studied using XIAP as a substrate (**Fig. 4.21**). Our data reports complete hydrolysis of XIAP within 10 mins of incubation at 37°C.

Overall, these results suggest that HtrA3 wild-type might regulate the apoptotic pathway by binding and cleaving XIAP.

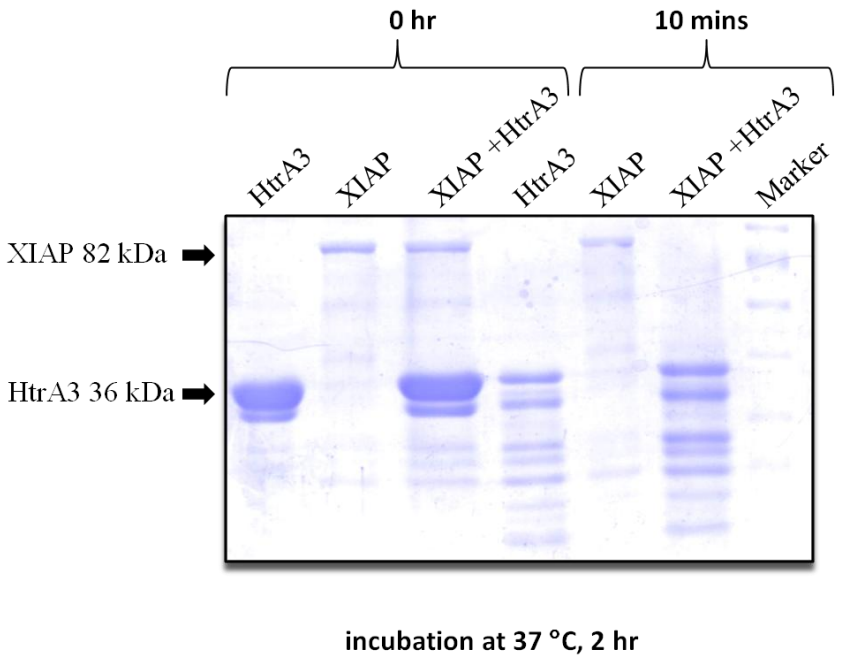


Figure 4.21 Proteolytic cleavage of XIAP by HtrA3 wild type. *HtrA3 was incubated with XIAP at 37°C for 10 min at 37℃. Reaction samples were resolved by SDS-PAGE and visualized with coomassie brilliant blue staining.*

4.5 Summary

Our study focuses on delineating the mode of activation of HtrA3 as well as understanding the regulatory switch that drives HtrA3 activation**.** *For the first time,* we have underlined the roles of both the N-terminus and the PDZ in holding the trimeric structure in a stable oligomeric ensemble and proper active-site formation. This study also reports that allostery modulates HtrA3 enzyme activity. Allosteric modulators have the potential to develop into important drug targets owing to specificity, diversity and less toxicity. Therefore, our findings, besides providing new insights into the structure-function coordination of HtrA3, also highlight an excellent approach to regulate HtrA3 activity efficiently in strategies devised against diseases it is associated with.

4.6 Discussion

The aim of this study was to gain an in-depth knowledge as well as to re-examine existing hypotheses on structure-function relationship, mode of activation and specificity of HtrA3. We delineated the role of different domains and their combinations, critical residues, and oligomerization in modulating HtrA3 activity and specificity with an aim at developing a working model for the activation mechanism of HtrA3.

We demonstrated that an allosteric mode of regulation exists for HtrA3 through a novel non-canonical substrate binding pocket (SBP_HtrA3). Moreover, this pocket does not include any major residue from the conserved binding groove *aka* GLGF motif (here FIGI) in PDZ domain, which is unique in this protease. Peptide binding to the allosteric site is accompanied by conformational changes in the regulatory loops LD, L1 and LA that eventually lead to a catalytically favourable active site formation. Interestingly, the ΔPDZ variant (HtrA3 N-SPD), that also demonstrates allosteric propagation with positive Hill coefficient has very less activity due to the formation of a malformed oxyanion hole. Therefore our findings suggest that the PDZ is essential for catalytic activation of HtrA3.

Our studies with the phenylalanine single mutants (HtrA3 F142D, HtrA3 F142A and HtrA3 F255D) showed that these mutations render the proteins inactive and monomeric. Δ143 HtrA3 (lacking F140) deletion variant shows destabilisation of the protein in both thermal denaturation and size exclusion chromatography studies, which suggests that the F140 is also critical for oligomerization. *In silico* MD simulation studies demonstrated higher RMSD and radius of gyration (Rg) values, along with loss of intramolecular hydrogen bonds for HtrA3 F142D and HtrA3 F255D mutants, when compared to mature HtrA3. Apart from structural instability, malformation of the catalytic pocket was also observed in the mutant systems where the catalytic triad residues (H191, D227 and S305) show movements, which are not conducive towards

formation of proper catalytically active conformation. These results from the molecular dynamic studies further substantiate the fact that 'triple-lock' residues such as F142 and F255 are critical in maintaining structural integrity and catalytic activity of HtrA3 protein. We also showed that while mature HtrA3 (wild type), HtrA3 S305A and HtrA3 PDZ tend to form trimers, HtrA3 SPD-PDZ and HtrA3 N-SPD are completely monomeric. These observations unambiguously demonstrate for the first time that the N-terminal region, PDZ domain and the 'triple lock' residues all contribute in a coordinated manner towards homotrimerization, stability and hence formation of a functional trimeric ensemble of the protease.

Protease assays with HtrA3 mutants and variants showed that proteins lacking the N-terminal region (HtrA3 SPD-PDZ) or both PDZ and N-terminal region (HtrA3 SPD) or any of the critical phenylalanine residues (HtrA3 F142D, HtrA3 F142A and HtrA3 F255D) were completely inactive. Constructs without certain regions of the N-terminus or the PDZ domain, exhibited a drastic loss in their enzymatic activity (Δ143 HtrA3 and HtrA3 N-SPD respectively). These observations unambiguously suggest that N-terminal region plays an important role in regulating its activity.

Examination of cleavage specificity using β-casein as substrate, revealed four that HtrA3 showed preference for similar residues at P1-P4 and P1'- P4' sites when compared to HtrA1 and HtrA2.

Temperature based activity studies suggested that HtrA3 activity increases beyond 30°C and is accompanied by destabilization of the protein post 50°C.

Based on our observations we provided a comprehensive idea of the mode of HtrA3 activation and have underlined the roles of both the N-terminus and the PDZ in holding the structure in a stable oligomeric ensemble and proper active-site formation. Moreover, this fundamental study

opens up possibilities of designing appropriate peptidomimetics or small molecule analogs to favourably manipulate HtrA3 functions for disease intervention.

HtrA3 was also shown to bind and cleave XIAP. This sheds new light on HtrA3's apoptotic functions and might help in better understanding of the pathway involved.

4.7 Significance of the study

1. Overall, this study proves for the first time that HtrA3 shows allosteric behaviour on substrate binding like its homologs, HtrA2 and HtrA4 which might be required for the specific cellular functions it is involved in.

2. For the first time, we have underlined the roles of both the N-terminus and the PDZ in holding the trimeric structure in a stable oligomeric ensemble and proper active-site formation and provided a comprehensive illustration of the mode of HtrA3 activation. This opens up possibilities of designing appropriate peptidomimetics or small molecule analogs to favourably manipulate HtrA3 functions for disease intervention.

CHAPTER 5

Identification and characterization of a novel binding partner of HtrA2

5.1 Introduction

HtrA2 is involved in several critical biological functions such as protein quality control, unfolded protein response (UPR), cell growth, apoptosis, arthritis, cancers and neurodegenerative disorders [1, 100]. Under normal physiological conditions, it acts as regulator of mitochondrial homeostasis where its protective function switches on its pro-apoptotic properties in response to stress-inducing agents [26, 33]. HtrA2 mediates apoptosis through classical pathways via caspase activation by displacing XIAP from caspases [26]. HtrA2 also mediates non classical cell death [24].

There is also evidence that interaction with IAPs is not the only mechanism involved in HtrA2 induced cell death, as the propteolytic activity of HtrA2, is sufficient to promote caspase cell death in absence of the caspases. HtrA2, via its protease activity degrades several cellular substrates such as Pea15, FLIP, Hax1 and cytoskeletal proteins such as Actin, Vimentin, α- and β-tubulin [51]. Despite substantial evidence of involvement of HtrA2 in triggering apoptosis, little is known about its mode of regulation, substrate recognition and specificity.

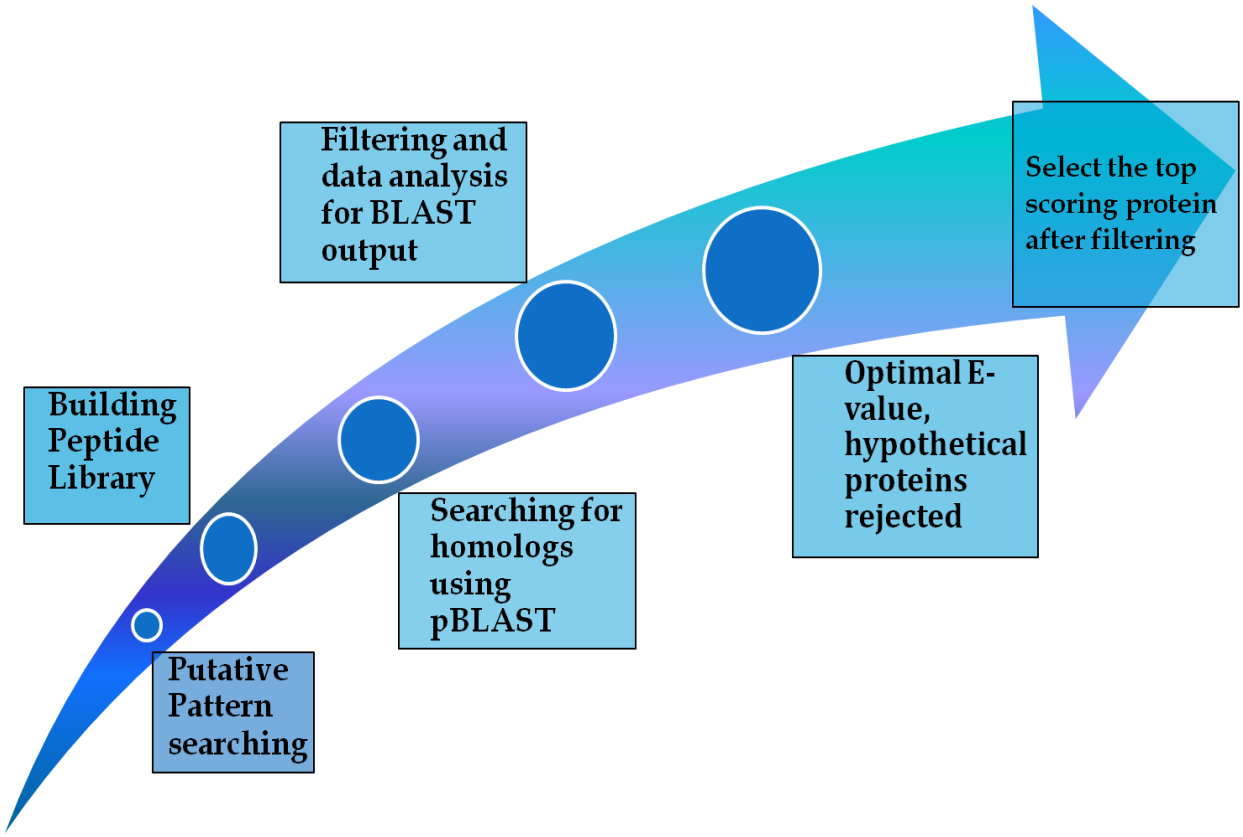
Although substrate unbound form of HtrA2 provided a broad overview of its structural organization, it could not decipher the mechanism of action or its regulation which is a prerequisite for understanding its role in various biological pathways and diseases [3]. It was reported that several repertoire of proteins (for e.g. GRIM 19, WARTS kinase and Presenilin) bind to the C-terminal PDZ domain of HtrA2 and stimulate the protease activity to promote cell death [133, 134]. Interestingly, in addition to the regulatory role played by the PDZ domain, HtrA2 manifests an additional mode of regulation via its N-terminus by binding to and cleaving IAPs [5]. This phenomenon apparently demonstrates existence of multiple or complex mechanisms of HtrA2 activation and regulation involving PDZ as well as other regions of the

protein. A clear understanding of the molecular mechanism of HtrA2 interaction with its binding partners or substrates will define ways of regulating its functions. Therefore, we performed a comprehensive *in silico* and biochemical study to identify and characterize a novel binding partner for HtrA2. These studies will help to further delineate the biological roles of HtrA2.

5.2 Results and discussion

5.2.1 In silico approach to find a putative binding partner/substrate of HtrA3

An *in silico* approach was used to predict the interacting/binding partners of HtrA2 using putative pattern search. After generating putative binding peptides from this pattern, these peptides were used for searching the homologues in human proteome. An overview of the work is as follows:



5.2.1.1 Designing a putative peptide pattern and building a library

Putative binding pattern was extracted from the HtrA2-PDZ phage-displayed peptide library using the consensus sequence of the aligned peptides based on literature review.

$$_{-6}[G/V]-_{-5}[\phi]-_{-4}[X]-_{-3\ to\ -1}[\phi\text{-}\phi\text{-}\phi]-_{-0}[V]-_{-1}[V/G]-_{-2}[G/W]-_{-3}[L]-_{-4}[G]$$

X= any of the 20 amino acid.

ϕ = any of the hydrophobic amino acid.

Figure 5.1 The 11-mer putative binding pattern. *X represents any of the 20 amino acids while ϕ stands for hydrophobic animo acid residues.*

The 11-mer pattern comprised of G/V at the -6th position, any hydrophobic amino acid at the -5th position, any amino acid at the -4th position, three consecutive hydrophobic amino acids in the -3rd to -1st position, V at 0th position, V/G at 1st position, G/W at 2nd position, L at 3rd position and G at 4th position (**Fig. 5.1**). Using this pattern, with the help of an in-house PERL script, random combinations of sequences were generated which were designated as peptides and accommodated into two sets of peptide libraries. The first set consisted of 24192000 peptides with a total size of 612 Mb, and the second one had 6912000 peptides with a total size of 170 Mb.

5.2.1.2 Identifying homologs using pBLAST

All these peptides were compared against protein database using standalone BLAST to retrieve the homolog protein sequences. A total of 34514 positive hits were found, out of which 28100 had e-value of 3 or more, 5567 had e-value between 2 and 3, 827 with e-value more than 0.3 and 26 had e-value less than 0.3.

5.2.1.3 Filtering proteins obtained by pBLAST

Using subcellular localization prediction software CELLO2GO, cellular locations of these 26 proteins were identified and only the cytoplasmic and mitochondrial proteins were shortlisted. Among the shortlisted proteins, hypothetical and already known binding partners of HtrA2 were excluded which resulted in a total of three proteins. Out of three, DUSP9 showed the highest docking score when bound to HtrA2 (**Fig. 5.2**) and hence, with further experimental validation, it was confirmed as HtrA2's novel binding partner.

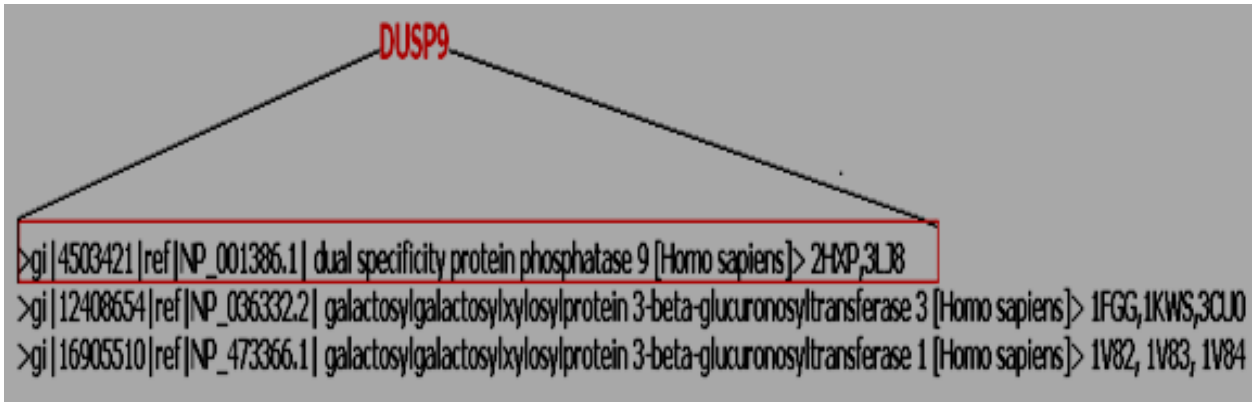


Figure 5.2 Homology search using pBLAST. *DUSP9 was identified as the top scoring protein among the final three selected based on filtering parameters.*

5.2.2 Validating DUSP9 as a binding partner of HtrA2

5.2.2.1 Cloning of DUSP9 in bacterial expression vectors

For the ease of purification, DUSP9 was subcloned in vectors with MBP, GST and His_{6x} tag as shown in **Fig. 5.3**.

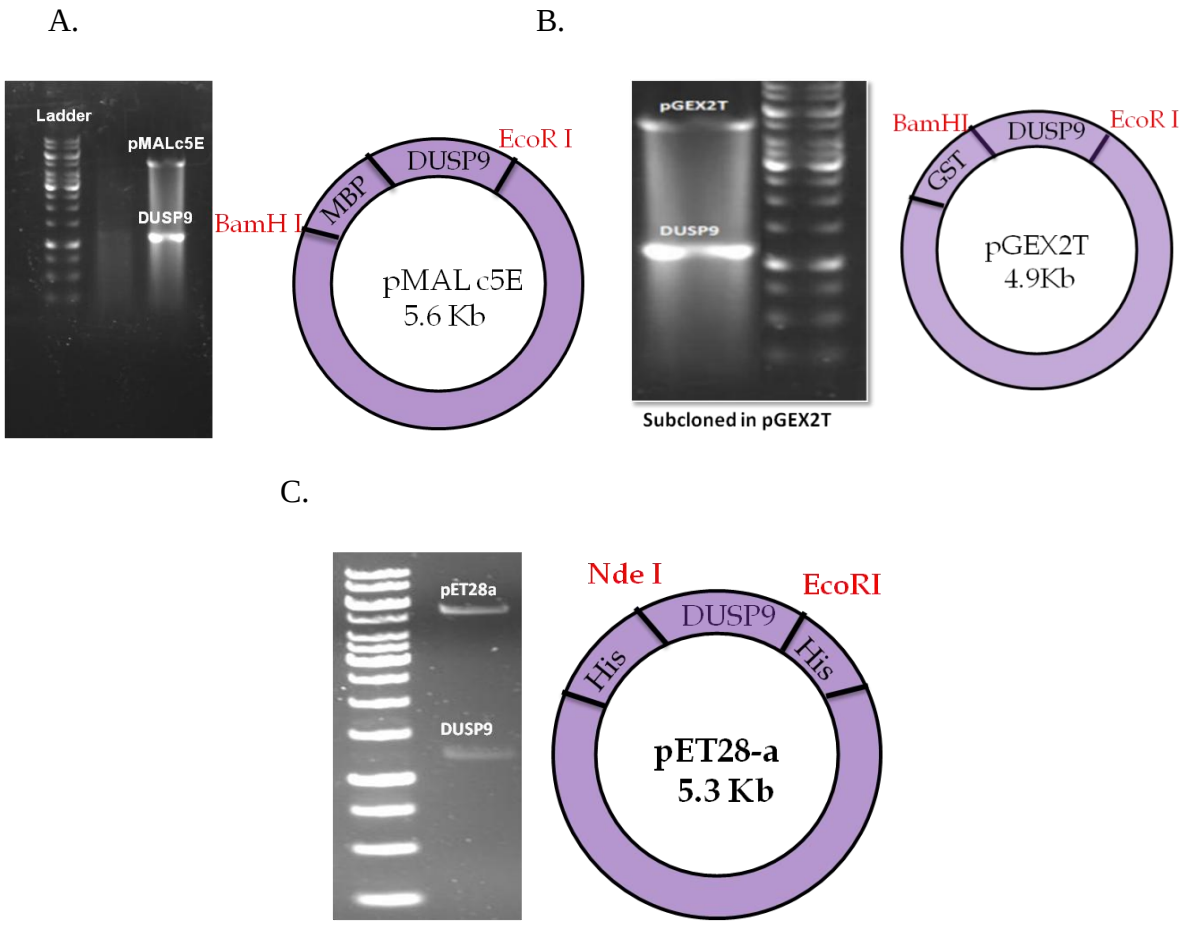


Figure 5.3 Different DUSP9constructs. *Expression vectors with different tags were used to clone DUSP9 to facilitate easy purification: a) MBP-DUSP9, b) GST-DUSP9 and c) His_{6x}-DUSP9.*

5.2.2.2 Purification of HtrA2 for pull down studies

Recombinant His_{6x}-HtrA2 was expressed in pET20b and purified using affinity chromatographic purification approach to obtain >90 % pure recombinant proteins (**Fig. 5.4)**. The plasmid pET28a was transformed in BL21 (DE3) pLysY cells for protein expression. The expression of polyhistidine tagged HtrA2 (pI – 8.03) was checked by inducing the cells at IPTG concentration of 0.5 mM. The supernatant was then bound to Ni-beads and samples were observed on an SDS PAGE gel which showed a prominent band at the expected size.

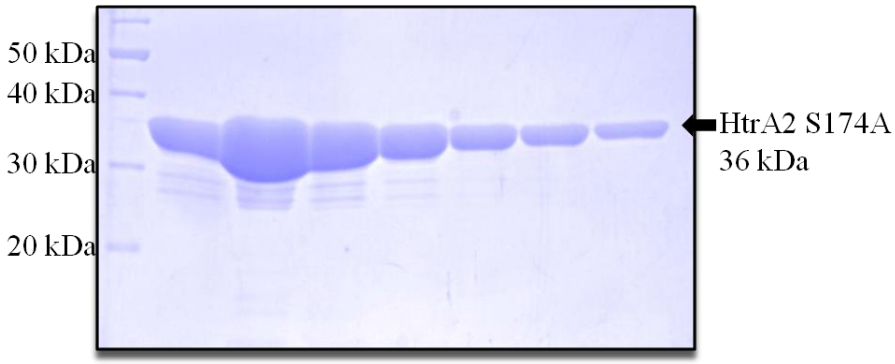


Figure 5.4 Active-site mutant of HtrA2 purified for pull down studies. *His_{6x}-HtrA2* was *expressed in BL21 (DE3) pLysY cells and purified using affinity chromatography.*

5.2.2.3 DUSP9 is a binding partner of HtrA2

Pull down studies were performed with cell lysate expressing GST fused recombinant DUSP9 as bait and purified HtrA2 S174A (active site mutant) as prey as described in **section 3.2.4.3** of **chapter 3**. No pull down was detectable using GST control alone **(Fig. 5.5)**.

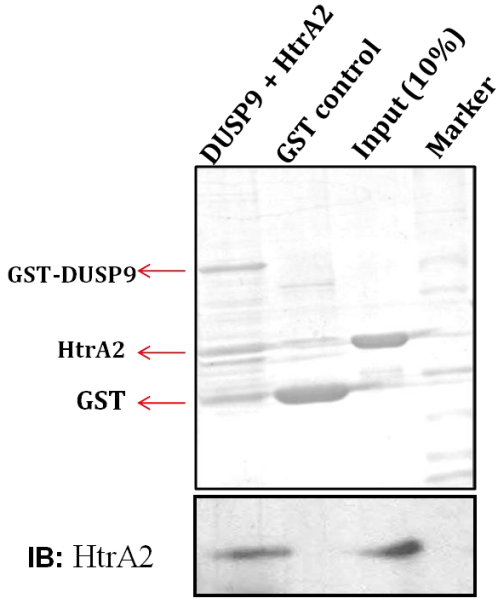


Figure 5.5 Pull down studies with HtrA2 and DUSP9. *Cell lysate expressing GST tagged DUSP9 was used as the bait and purified HtrA2 was used as the prey protein for pull down. GST alone was used as control.*

The pull down study suggests that DUSP9 is a binding partner of HtrA2.

5.2.2.4 Co-IP studies with HtrA2 and DUSP9

To see if these findings were indicative of a functional relation between HtrA2 and DUSP9, we investigated a physical interaction between these two proteins. Therefore, a carboxy-terminally FLAG tagged version of DUSP9 was expressed in HEK-293T cells, in order to recover tagged DUSP9 along with interacting proteins. We observed a direct interaction of overexpressed FLAG-tagged DUSP9 protein with endogenous HtrA2 protein in co-immunoprecipitation experiments using HEK293T cells **(Fig. 5.6).** Collectively, we established that DUSP9 interacts with HtrA2.

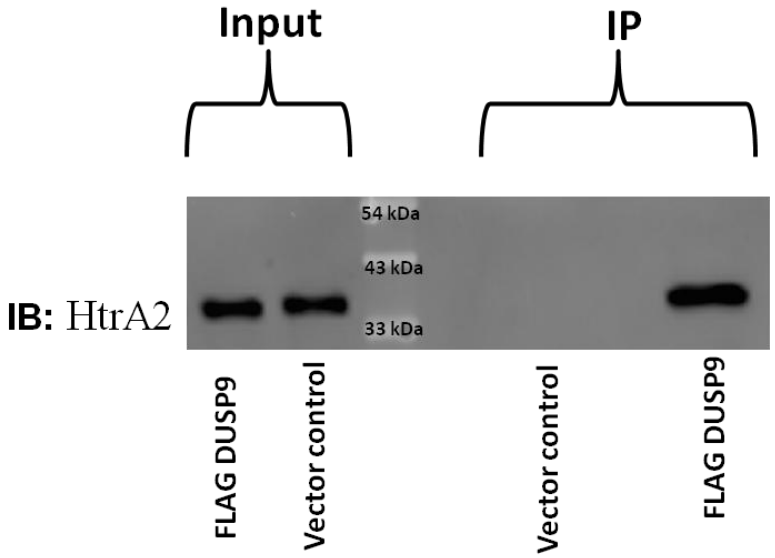


Figure 5.6 HtrA2 interacts with DUSP9. *HEK293T cells were transfected with FLAG-tagged DUSP9 while untransfected cells were used as controls. Lysates were subjected to Western blot analysis directly (inputs) or after incubation with anti-FLAG coupled agarose. Immunoblot were probed with anti-HtrA2 as indicated.*

5.2.2.5 DUSP9 is a substrate of HtrA2

To address if DUSP9 is a substrate of HtrA2 as well, proteolytic assays were performed (**Fig. 5.5**). Our data reports complete hydrolysis of DUSP9 within 4 hr of incubation at 37°C.

Overall, these results suggest that DUSP9 is a binding partner and substrate of HtrA3

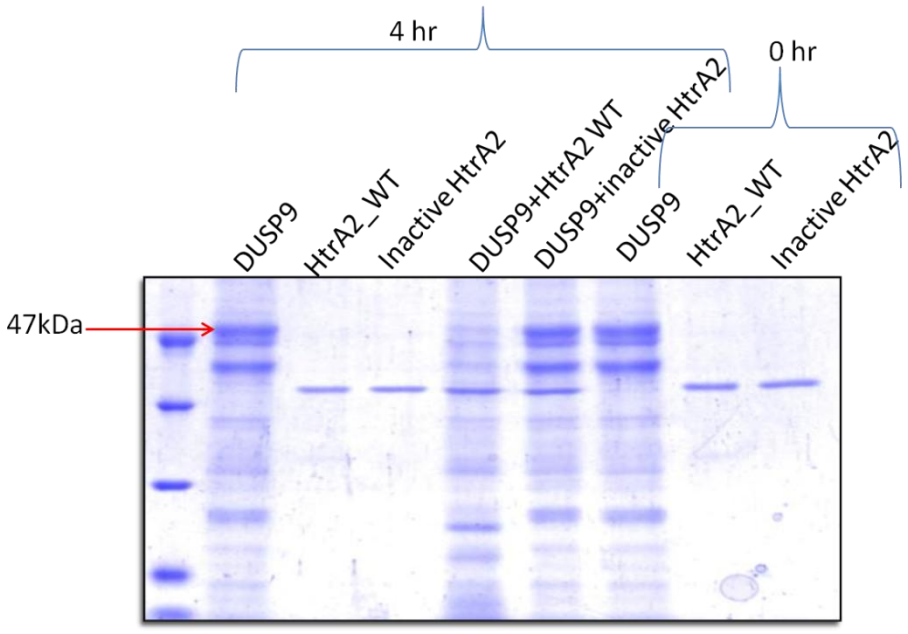


Figure 5.7 Proteolytic cleavage of DUSP9 by HtrA2. *HtrA2 was incubated with DUSP9 for 4 hr and the products of the proteolytic cleavage were visualized using SDS PAGE. HtrA2, its active-site mutant (S174A) and DUSP9 alone were used as controls.*

5.2.3 Docking studies of HtrA2 and DUSP9

Since the full length crystal structure of the protein was not available, homology modeling of DUSP9 was used to generate a full length model. This model was then assessed using ProSA and SAVES server.

Using an in silico tool/server Cluspro 2.0, a blind docking approach was used to dock HtrA2 and DUSP9 to identify the critical interface residues (**Fig. 5.8 and table 5.1**).

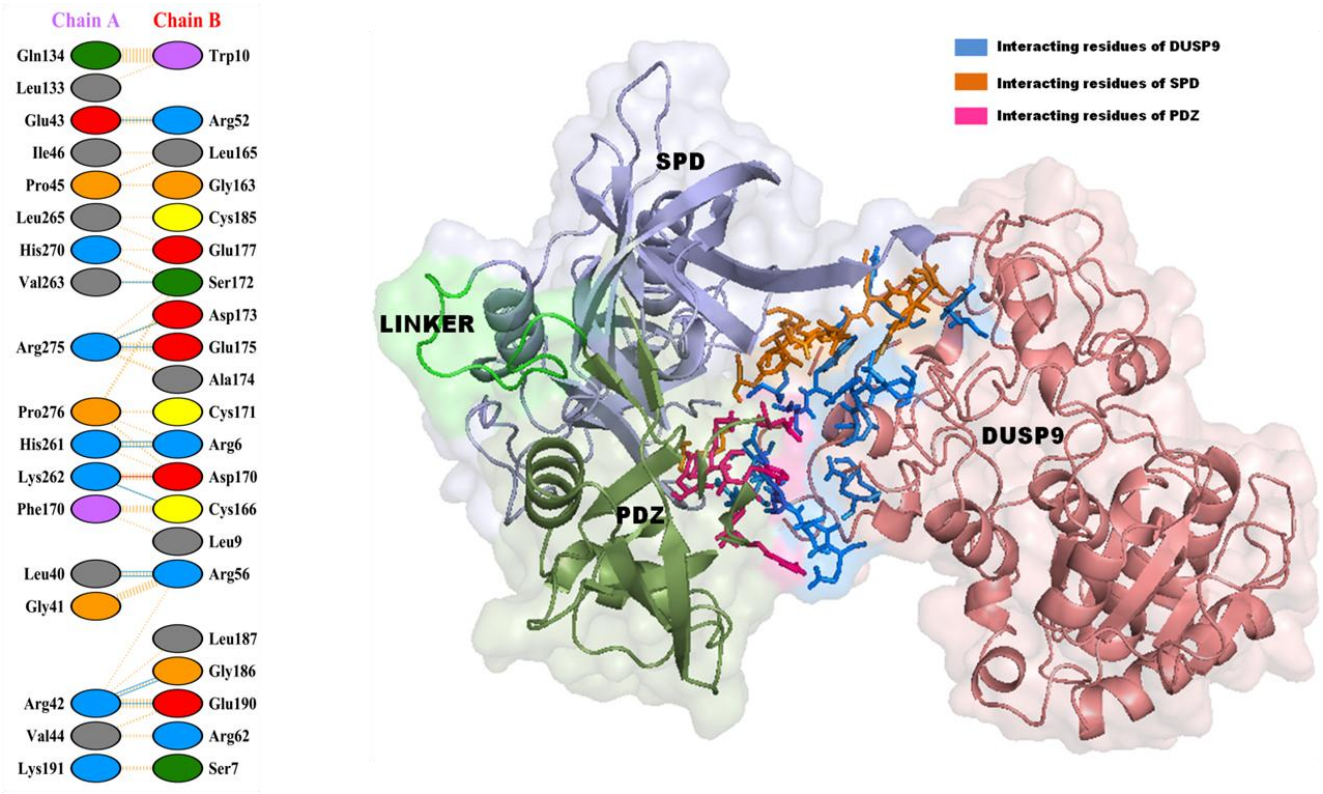


Figure 5.8 Docked structure of HtrA2 and DUSP. *Chain A represents HtrA2 while chain B represents DUSP9.The residues important for interaction from each chain were identified based on the number and strength of contacts made.*

HtrA2	DUSP9
• L40, G41, R42, E43	• D170, S172, D173, A174, E175
• H261, K262, V263	• R6, S7, W10
• R275, P276	• R52, R56
• L133, Q134	• G186, L187, E190

Table 5.1 Interacting residues of the HtrA2-DUSP9 binding interface. *Four patches of residues were identified on both the proteins based on the docked structure.*

5.3 Summary

We identified DUSP9 as a putative binding partner of HtrA2 through *in silico* studies and validated the interaction using *in vitro* studies. The two proteins were docked and the residues participating in the interaction were identified. The critical residues in the interaction will be identified using mutational analysis.

5.4 Significance of the study

This study was used to develop an accurate and reliable *in silico* method to predict putative binding partners/substrates of proteins having PDZ domains. Most importantly, we validated DUSP9 as a novel binding partner cum substrate of HtrA2. Given the biological and pathological significance of HtrA2, identifying novel natural substrates of this protein will help us understand the diverse functions HtrA2 plays in a cellular background. Furthermore, it might also help in devising strategies for modulating HtrA2 functions for therapeutic benefits. DUSP9 primarily regulates the ERK signaling pathway which in turns influences growth and apoptosis [107, 108, 111]. Several studies have shown that DUSP9 modulates the insulin signaling pathway and is also implicated in cancers and placental development and it is interesting to note that there is a huge functional overlap in DUSP9 and HtrA proteins suggesting that the intricacies of this interaction needs to be elucidated to understand its biological significance.