

**Structural and Functional Characterization of Proapoptotic
Proteins: Human Papillomavirus E2 and Serine Protease HtrA2**

By

Nitu Singh

[LIFE09200804010]

Tata Memorial Centre

Mumbai

*A thesis submitted to the
Board of Studies in Life Sciences
In partial fulfilment of requirements
for the Degree of*

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE

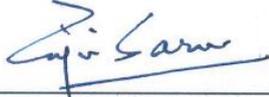


June, 2015

HOMI BHABHA NATIONAL INSTITUTE

Recommendations of the Viva Voce Board

As members of the Viva Voce Board, we certify that we have read the dissertation prepared by Nitu Singh entitled “**Structural and functional characterization of proapoptotic proteins: Human papillomavirus E2 and serine protease HtrA2**” and recommend that it may be accepted as fulfilling the dissertation requirements for the Degree of Doctor of Philosophy.



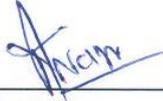
Chairperson: Dr. Rajiv Sarin

Date: 25.6.15



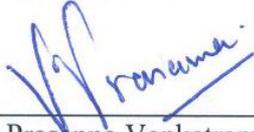
Guide/Convenor: Dr. Kakoli Bose

Date: 25/6/15



Member 1: Dr. Ashok Varma

Date: 25/06/15



Member 2: Dr. Prasanna Venkatraman

Date: 25/6/15



Invitee: Dr. Smita Mahale

Date: 25/6/15



External Examiner: Dr. S. Mazumdar

Date: 25/6/2015

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

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

Date: 25/6/15

Place: Navi-Mumbai



Kakoli Bose
(Guide)

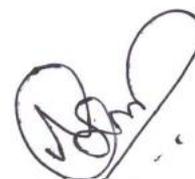
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfilment 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.

Date: 25/06/2015

Navi Mumbai



(Nitu Singh)

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.

Date: 25/06/2015

Navi-Mumbai

A handwritten signature in black ink, appearing to be 'Nitu Singh', written in a cursive style.

(Nitu Singh)

List of publications arising from the thesis

Journal

1. Equilibrium dissociation and unfolding of human papillomavirus E2 transactivation domain. **Singh N.**, Kanthaje S, Bose K. *Biochemical and Biophysical Research Communications*, **2015**. <http://dx.doi.org/10.1016/j.bbrc.2015.05.057>
2. Dual regulatory switch confers tighter control on HtrA2 proteolytic activity. **Singh N**, D'Souza A, Cholleti A, Sastry GM, Bose K. *FEBS Journal*, **2014**, 281, 2456–2470.
3. Allosteric regulation of serine protease HtrA2 through novel non-canonical substrate binding pocket. Bejugam PR, Kuppili RR, **Singh N**, Gadewal N, Chaganti LK, Sastry GM, Bose K. *PLoS One*, **2013**, Vol. 8, Issue-2, e55416.
4. The structural basis of mode of activation and functional diversity: a case study with HtrA family of serine proteases. **Singh N**, Kuppili RR, Bose K. *Archives of Biochemistry and Biophysics*, **2011**, Vol. 516, 85–96.
5. Molecular evidence of death effector domain chain assembly and its role in caspase-8 activation. **Singh N**, Hassan A, Bose K. *FASEB*, **2015** (Under Review)
6. Unraveling caspase activation mechanism by a non-death-fold proapoptotic inducer: human papillomavirus E2 protein. **Singh N**, Bose K. (manuscript prepared)

Book Chapters

1. Apoptosis: Pathways, Molecules and Beyond. **Singh N.** and Bose K. *Springer*. ISBN 978-3-319-19496-7
2. Cathepsins and HtrAs - multitasking proteases in programmed cell death. Chaganti L, **Singh N.** and Bose K. Springer Ed. ISBN 978-3-319-19496-7

Conference proceedings

1. Insights into the novel adaptor-Independent apoptosis mediated by human papillomavirus E2 protein. **Singh N**, Bose K. *Biophysical journal*, **2012**, 102(3), 629.

Conferences

1. Oral presentation entitled “Dual regulatory switch confers tighter control on HtrA2 proteolytic activity” at international symposium on “**Conceptual Advances in Cellular Homeostasis Regulated by Proteases & Chaperones: The Present, the Future & Impact on Human Diseases**” on Dec 3-6, 2013 at ACTREC, Navi Mumbai, India.
2. Poster presentation entitled “Unraveling the mechanism of procaspase-8 activation by a novel adaptor independent pathway” at **Biophysics Paschim Fourth Meeting (BPP IV)** on March 2, 2013 at NCL, Pune, India.
3. Presented a poster entitled “Structural and functional determinants of XIAP mediated activation of pro-apoptotic serine protease HtrA2/Omi” at **37th Annual meeting of Indian Biophysical Society (IBS)** and “**National Symposium on frontiers of Biophysics, Biotechnology and Bioinformatics**” held during Jan 13-16, 2013 at University of Mumbai, India.
4. Poster presentation entitled ‘Effect of mutations on dimer interface of human papillomavirus E2 transactivation domain’ at **International Symposium on Protein Folding and Dynamics**, October 15-17th 2012 in NCBS, Bangalore, India.
5. Presented a poster entitled ‘Insights into the novel adaptor-independent apoptosis mediated by human papillomavirus E2 protein’ in **Biophysical Society 56th Annual Meeting**, February 25 – 29th 2012, San Diego, California, USA.
6. Presented a poster entitled “Dissecting the novel adaptor independent apoptotic mechanism mediated by human papillomaviral E2 protein” at **Indo-US Symposium and Workshop, Modern Trends in Macromolecular Structure**, Feb. 2011, Indian Institute of Technology, Mumbai, India.



Nitu Singh

Date: 25/06/2015

ACKNOWLEDGEMENT

First and foremost, I offer my sincere gratitude to my research supervisor Dr. Kakoli Bose for her deft guidance and being such a wonderful and motivating guide. One simply could not wish for a better or friendlier supervisor. Her stimulating suggestions and encouragement accompanied by the freedom of thought that she granted, has not only helped me to discover my potential but also contributed new dimension to my scientific career. Ma'am, I thank you for everything and will always be proud to have been mentored by you.

I am grateful to Dr. S.V. Chiplunkar (Director- ACTREC), Dr. R. Sarin (Former Director- ACTREC), and Dr. S. Zingde (Former Deputy Director-ACTREC) for providing research facility and excellent infrastructure. I sincerely thank the honourable members of my doctoral committee, Dr. M.V. Hosur, Dr. R. Sarin, Dr. A.K. Varma, Dr. V. Prasanna, and Dr. Smita Mahale. Their critical evaluation and suggestions have contributed significantly to the work.

A special thanks to Dr. Madhavi Sastry and Dr. Sanjib Senapati for their help with in silico studies. I am extremely thankful to the common instrument facility, BTIS, library, IT, administration and accounts department of ACTREC for their constant help and support.

I am fortunate to have wonderful lab members, Mr. Padale ji, Snehal, Lalith, Raja, Saujanya, Ajay, Raghu and our excellent trainees and research fellows who helped me in innumerable ways during the entire period. I thank them all from bottom of my heart for their unconditional love and support. Special thanks are due to my batchmates-2008 and student's community of ACTREC for their cooperation, help and pleasure of their company.

I owe a lot more than I can express in words to my parents and family members who have been my major guiding force. Words are not enough to thank Mummy, Papa and Ritu for making me what I am. It is difficult to express my feelings to my better-half, Vikrant, for being my source of perennial support and loving me for who I am. Thank you for coming into my life. Nothing is impossible with You by my side!

And last but not the least, I thank God for everything.

CONTENTS

	Page No.
Synopsis	1
List of Abbreviations	14
List of Figures	16
List of Tables	19
Chapter-1: Introduction	21
Chapter-2: Review of literature	26
2.1 Human papillomavirus	27
2.1.1 Functions of papillomavirus proteins	29
2.1.2 HPV life cycle and pathogenesis	29
2.1.3 Current cervical cancer therapies	31
2.2 Human papillomavirus E2 protein	33
2.2.1 HPV18 E2-induced apoptosis	34
2.2.2 Structure and function of typical death-fold domain	37
2.2.3 Structure and function of HPV E2 protein	43
2.3 HtrA2 – An overview	46
2.3.1 Caspase-dependent mechanism	47
2.3.2 Caspase-independent apoptosis	48
2.3.3 Alterations in cancer	50
2.3.4 Structural details	52
2.3.5 Mechanism of activation	53
<i>Aims and Objectives</i>	55
Chapter-3: Methods and Materials	57
3.1 Cloning, expression and purification of recombinant proteins	58
3.2 Animal cell culture	
3.2.1 Routine maintenance of cell lines	71
3.2.2 Transient transfection	73
3.2.3 Live cell confocal imaging	74
3.2.4 Preparation of whole cell lysate	74
3.2.5 Protein estimation using Peterson method	75
3.2.6 SDS-PAGE	75
3.2.7 Western blotting	76
3.2.8 Cell death assays	77
3.3 Protein-protein interaction analyses	
3.3.1 Pull-down assay	79
3.3.2 Co-immunoprecipitation	80

3.3.3	Isothermal titration calorimetry	80
3.3.4	Surface plasmon resonance	81
3.4	Biophysical characterization	
3.4.1	Circular dichroism spectroscopy	83
3.4.2	Fluorescence spectroscopy	83
3.5	Protein oligomerization and size characterization	
3.5.1	Size-exclusion chromatography	90
3.5.2	Glutraldehyde cross-linking	91
3.5.3	Dynamic light scattering	91
3.6	Serine protease enzymatic assay	92
3.7	Molecular modeling, docking and simulation	93
3.7.1	Types of interactions	94
3.7.2	General steps for molecular modeling, simulation and docking	96
Chapter-4: Human papillomavirus E2 protein		98
4.1	<i>Understanding protein-protein interactions involved in non-death-fold E2-induced apoptosis</i>	
4.1.1	Experimental procedures	102
4.1.2	Results	
4.1.2.1	Mapping minimal binding region	106
4.1.2.2	<i>In silico</i> prediction of E2-procaspase-8 binding interface	109
4.1.2.3	Identification of procaspase specific binding site on E2	111
4.1.2.4	Identification of E2 specific binding site on procaspase-8	117
4.1.2.5	Monitoring the effect of interface E2 mutant on cell death	119
4.1.2.6	E2-FADD interaction analysis	120
4.1.3	Discussion	122
4.2	<i>Molecular details of DED assembly and its role in death-fold mediated caspase-8 activation</i>	
4.2.1	Introduction	126
4.2.2	Experimental procedures	128
4.2.3	Results	
4.2.3.1	Structural basis of DED-DED interaction	134
4.2.3.2	Molecular evidence of procaspase-8 DED assembly	136
4.2.3.3	Molecular basis of FADD DED self-association	144
4.2.3.4	Mechanism of FADD – procaspase-8 assembly	150
4.2.4	Discussion	156

4.3	<i>Biochemical and biophysical characterization of E2 transactivation domain</i>	
4.3.1	Introduction	162
4.3.2	Experimental procedures	164
4.3.2	Results and Discussion	
4.3.2.1	<i>In silico</i> analysis of E2 TAD dimerization	168
4.3.2.2	<i>In vitro</i> characterization of wild-type and mutant E2 TAD	171
4.3.2.3	Unfolding studies	174
4.2.3.4	E2 TAD homo vs. E2 TAD – DED heterodimerization	179

Chapter-5: Human serine protease HtrA2

5.1	Introduction	182
5.2	Experimental procedures	184
5.3	Results	
5.3.1	Elucidating the role of IAPs in activation of HtrA2	
5.3.1.1	Determination of the proteolytic activity of HtrA2 in presence of IAPs	189
5.3.1.2	Disruption of the XIAP–HtrA2 interaction prevents protease activation	190
5.3.1.3	Interaction analysis	193
5.3.2	Determining specificity and catalytic efficiency of HtrA2	
5.3.2.1	Steady-state kinetic parameters for HtrA2	196
5.3.2.3	Comparison of the catalytic efficiency upon N- and C-termini induced activation	198
5.3.3	Delineating the global conformational changes following IAP binding	
5.3.2.1	Monitoring conformational changes using <i>in silico</i> approach	202
5.3.2.2	Structural evaluation of peptide bound and unbound HtrA2	203
5.4	Discussion	206

Chapter-6: Concluding remarks	211
--------------------------------------	-----

Chapter-7: Future perspective	217
--------------------------------------	-----

References	220
------------	-----

Publication Reprints

Synopsis



Homi Bhabha National Institute

Ph. D. PROGRAMME

- | | |
|--|--|
| 1. Name of the Student: | Nitu Singh |
| 2. Name of the Constituent Institution: | Tata Memorial Centre, Advanced Centre for
Treatment Research and Education in Cancer |
| 3. Enrolment No. : | LIFE09200804010 |
| 4. Title of the Thesis: | “Structural and functional characterization of
proapoptotic proteins: Human papillomavirus E2
and serine protease HtrA2” |
| 5. Board of study: | Life Sciences |

Introduction:

In maintaining homeostasis, multicellular organisms tightly couple the rate of cell proliferation and death. Disruption of this fine balance leads to several life-threatening diseases such as cancer and neurodegenerative disorders. Cancer, which is characterized by a breakdown in cellular apoptotic machinery, might occur due to alteration of structural and functional properties of critical proteins in the apoptotic pathway which is closely related to the genesis and progression of this deadly disease.

Tremendous progress in apoptotic research has occurred to understand the apoptotic pathways with an aim at targeting them for therapeutic intervention. Recently, human papillomavirus (HPV) early protein E2 and human serine protease HtrA2 have been reported to trigger novel alternate pathways of apoptosis (1,2). Therefore, characterizing the structural and functional properties of these proapoptotic proteins will help understand the apoptotic pathway better and might open new avenues in cancer research.

Part-I) Structural and functional characterization of human papillomavirus E2 protein

High risk HPV-18 E2 protein has been found to induce apoptosis via direct interaction with procaspase-8 through a unique pathway that bypasses homotypic interactions between death domains of proteins in death inducing signaling complex (DISC) (1,3). Our aim was to delineate the mechanism of interaction between HPV18 E2 and procaspase-8 as well as to understand how it aids in E2-induced apoptosis.

Objectives:

1. To identify minimal binding regions and critical residues involved in the E2-procaspase-8 interaction and understanding mechanism of caspase activation.
2. To design interface mutants and study the effect on procaspase activation and apoptosis.

Methodology:

The gene of interest was cloned in bacterial or mammalian expression vectors and site-directed mutagenesis was performed to introduce desired mutations. Recombinant proteins were expressed and purified using different chromatography techniques including affinity and size-exclusion, and their percentage purity was determined using SDS-PAGE. Protein-protein interaction analysis was carried out using pull-down and co-immunoprecipitation (Co-IP) assays. Biophysical probes, such as circular dichroism, fluorescence spectroscopy, dynamic light scattering and mass spectrometry were used to characterize mutant and wild-type proteins. Structural analysis was also carried out using computational approach such as molecular modeling, docking and simulation.

Results:

1. Mapping the minimal binding region and critical residues involved in the interaction

From *in-vitro* studies we identified that HPV18 E2 transactivation domain (TAD) strongly associates with death effector domain (DED)-B of procaspase-8. In order to map the minimal binding region we generated deletion constructs comprising different combination of six α -helices of DED-B and performed pull-down assay. From this analysis we determined that $\alpha 2$ and/ or $\alpha 5$ helices of DED-B are important in mediating E2-procaspase-8 interaction.

1.1 *In silico* prediction of E2-procaspase-8 binding interface

The MC159 viral FLIP (vFLIP) crystal structure (PDB entry 2BBR) provides an excellent template to study the structure function relationship of tandem DED containing proteins (4). The procaspase-8 tandem DED is 48% similar in sequence to the MC159 vFLIP. Therefore, by homology modeling we generated the structure of caspase-8 DED using *Modeller 9v7*. HPV18 E2 TAD structure was also modeled for the first 65 missing residues in crystal structure (PDB ID. 2QQH) using HPV16 E2 TAD (1DTO) as the template. From protein-protein docking we determined that the interface between DED and E2 TAD complex

notably involved residues from $\alpha 2$ and/ or $\alpha 5$ of DED-B, which very well corroborated with our *in vitro* deletion mutagenesis data.

1.2 Identification of critical residues in E2 for procaspase-8 binding

To identify the critical residues, we introduced series of mutations in E2 TAD and performed pull-down assay. It was found that R41, W42 and F48 are the key residues involved in E2-procaspase-8 interaction. Co-IP confirmed that E2 wild-type associates with procaspase-8 but the triple mutant does not. In addition, docking of mutant E2 with DED-B showed significant reduction in the docking score.

1.3 Identification of critical residue in procaspase-8 for E2 binding

Taking clues from deletion constructs, we targeted residues of $\alpha 2$ and/ or $\alpha 5$ helices of DED-B. Series of mutations were carried out individually and in combination on these two helices to identify the key residues. Using single and combinatorial mutants we performed pull-down and Co-IP assay which identified that F122, L123 of $\alpha 2$ helix and D158, Q166 ($\alpha 5$ helix) are important in mediating E2-procaspase-8 interaction.

To further comprehend the interaction data, co-localization studies were carried out with ectopic expression of GFP-E2 and mcherry-procaspase-8 wild-type and their corresponding mutants. It was observed that E2 interaction with procaspase-8 results in the formation of cytoplasmic speckles, however, the mutants did not co-localize. Interestingly, DED-B interacts with E2, while DED-AB does not show any interaction. It has been reported that DED-AB over expression leads to formation of filamentous structure called as ‘death effector filaments’ (5). We observed that DED-AB filaments did not interact with E2 suggesting that in process filament formation, the E2 binding site plausibly is unavailable.

2. Delineating the mechanism of procaspase-8 activation

In recent years, new insights into procaspase-8 activation, especially at the level of DED-containing proteins are emerging (6). DISC assembly and caspase activation occurs mainly at

the level of DED-chain formation (7). Therefore, it was intriguing to understand the mechanism of DED filament formation and its involvement in procaspase-8 activation via E2 mediated novel adaptor-independent pathway and classical FADD-dependent signaling.

2.1 Understanding the mechanism of DED-chain assembly - a prerequisite for procaspase-8 activation

To investigate the mechanism of DED-chain formation, structure-guided mutational analysis was carried out. From this study, we identified that the hydrophobic patch comprising residues F122, L123 termed as 'FL motif' on DED-B and L42, F45 residues i.e. 'LXXF-motif' (where X represents any amino acid) from DED-A hetero-oligomerizes to form DED-chain. Substitution of either of this motif resulted into complete abrogation of filament formation.

2.2 Comparison of novel adaptor-independent pathway with the classical death receptor-mediated caspase activation

To delineate the mechanism of classical death-receptor pathway, co-localization was carried out in HEK 293 cells, using ectopic expression of FADD and procaspase-8. Upon co-expression, we observed that cytoplasmic diffused localization of procaspase-8 is transformed into co-localized short filaments. When we mutated the 'FL-' or 'LXXF-motif' of procaspase-8, it was observed that FL-motif mutant localized with FADD, whereas the LXXF-motif mutant did not. Thus, DED-A is involved in FADD-procaspase-8 interaction as opposed to DED-B observed in case of E2-mediated pathway. These observations were further tested with minimal binding domains and confirmed using Co-IP and *in silico* analyses. Overall, procaspase-8 preferentially engages FADD using $\alpha 1/\alpha 4$ surface and E2 via $\alpha 2/\alpha 5$ surface. Based upon these observations we propose that procaspase-8 is recruited to FADD via DED-A and E2 through its DED-B, subsequently initiating procaspase-8 chain

formation to allow proximity induced cleavage to release active caspase-8 and hence promote apoptosis.

2.3 E2, a non- death-fold protein with death-fold characteristics

It was intriguing as to how E2 adopts a death-fold behavior and behave as a potent binding partner of procaspase-8. In order to understand this, we characterized E2 TAD wild-type, single and double mutants corresponding to residues R41, W42. Secondary and tertiary structure analyses suggested that all the proteins have similar conformation. Size exclusion chromatography showed presence of small population of dimer in case of wild-type and single mutants whereas the double mutant was predominantly obtained as monomer. This was further confirmed using glutaraldehyde cross-linking, dynamic light scattering, equilibrium unfolding and *in silico* studies.

From these studies we infer that E2 TAD can homo-dimerize however, has lesser stability and binding affinity compared to E2-procaspase-8 hetero-complex. Taken together, we propose that surface defined by helices $\alpha 2/\alpha 3$ of E2 TAD enables it to act as a pseudo death-fold adaptor protein having potential to homodimerize and interact with typical death-fold domain of caspases, to induce caspase-8 oligomerization followed by its activation and cell death.

2.4 Effect of interface mutations on procaspase activation and cell death

Apoptosis assays were performed to determine percent cell death in presence of wild-type or the interface mutants. Measuring cell viability using flow cytometry and DNA ladder assay showed that there was significant reduction in cell death in mutants as compared to wild-type proteins.

Summary:

- Residues from $\alpha 2$ (R41, W42 and F51) of E2 and F122, L123 ($\alpha 2$) and D158, Q166 ($\alpha 5$) of procaspase-8 DED-B are essential for mediating E2-procaspase-8 interaction.

- The caspase-8 DED chain assembly involves hetero-oligomerization of DED-A and DED-B via LXXF and FL motif respectively.
- In classical death-fold pathway, the binding surface defined by $\alpha 2/\alpha 3$ or $\alpha 1/\alpha 4$ helices of FADD-DED engages with $\alpha 1/\alpha 4$ helices of procaspase-8 DED-A.
- HPV18 E2 TAD is weak dimer with the dissociation constant in the micromolar range and the dimerization interface comprises residues R41 and W42 from α helix-2.

Conclusion:

The high risk HPV18 E2 protein mediates caspase-8 dependent cell death probably by co-operating with FADD to induce caspase-8 activation. The finding that a non-death fold protein E2 could induce caspase-8 activation opens new avenues in apoptosis research. Furthermore, identification of the critical residues involved in this interaction might help in designing novel E2 analogs with desired characteristics so as to modulate procaspase-8 activation and hence cell death.

Part-II) Structural and functional characterization of serine protease HtrA2

HtrA2 (High temperature requirement A2) is a mitochondrial serine protease which antagonizes inhibitor of apoptosis proteins (IAPs) thereby activating caspases (8). It mediates apoptosis through classical pathways and also by a less understood caspase-independent mechanism (9). Looking into its complex trimeric three-dimensional structure (N-terminal domain, a serine protease domain and a PDZ domain) and trying to account for its low protease activity, researchers hypothesized a model which suggests intricate PDZ-protease coordination, rearrangement in their relative orientations and conformational changes at PDZ-protease interface as prerequisites for substrate cleavage (10). However, if this model is true, it might hold good only for subset of interactions since IAPs bind the N-terminal domain. This novel phenomenon unequivocally demonstrates existence of multiple or complex mechanism of HtrA2 activation involving PDZ and other regions of the protein.

Therefore, our aim was to address this complex behavior toward defining its global mode of regulation and protease activity.

Objectives:

1. To identify protein-protein interaction sites of HtrA2 and IAP and determine the role of IAPs in activation of HtrA2 protease activity
2. To compare specificity & catalytic efficiency of HtrA2 upon IAP protein/peptide and PDZ peptide binding.
3. To delineate the global conformational changes and active-site orientation following IAP protein/peptide binding.

Methodology:

Different constructs were generated using cloning and site-directed mutagenesis. Recombinant proteins were expressed, purified and characterized using biophysical probes such as circular dichroism, fluorescence spectroscopy. *In vitro* interaction analysis was carried out by pull-down assay, isothermal titration calorimetry and surface plasmon resonance techniques. Steady-state enzyme kinetic parameters were determined using FITC-labelled β -casein substrate, a well established model substrate of HtrA proteases.

Results:

1. Interaction analysis and elucidating the role of IAPs in activation of HtrA2

a) Determination of the proteolytic activity of HtrA2 in presence of IAPs

To assess any role of IAPs on HtrA2 activity, we determined the initial velocity (v_0) of HtrA2 as a function of increasing XIAP concentration, using fixed amount of enzyme and substrate β -casein. It was observed that in absence of XIAP, HtrA2 displayed maximum velocity (V_{max}) of $0.8 \times 10^{-9} \text{ Ms}^{-1}$ which significantly increased in presence of XIAP. We also tested the protease activity in presence of BIR2 or BIR3 as well as peptides corresponding to IBM (IAP-binding motif) groove of these two minimal binding domains. It was observed that in

presence of BIR2, the cleavage rate was similar to full length XIAP; however, the former achieved so at the expense of four-fold more protein concentration.

b) Interaction analysis

To quantify the kinetics of the complex formation, surface plasmon resonance technique was used. Kinetic analysis of the binding data for BIR2 and 3 demonstrated apparent equilibrium dissociation constants (K_D) of 54 ± 8 nM and 254 ± 18 nM respectively, suggesting that BIR2 has higher affinity compared to BIR3.

2 Determining specificity and catalytic efficiency of HtrA2

a) Steady-state kinetic parameters for HtrA2 in presence of XIAP

To determine the steady-state kinetic parameters, we measured initial rates of substrate cleavage in absence or presence of XIAP proteins/peptides. Upon pre-incubation of HtrA2 with XIAP, the catalytic efficiency increased about three-fold and Hill constant was reduced, suggesting greater stabilization of the relaxed state of protease upon IAP binding. It was observed that in presence of XIAP, there is an increase in V_{max} without significant alteration in K_m which indicates that HtrA2 likely follows ‘*V system*’ of allosteric modulation (11).

b) Comparison of the catalytic efficiency upon N- or C-termini induced activation

To understand whether XIAP-mediated activation of HtrA2 requires any direct involvement of PDZ domain, enzymatic parameters for PDZ lacking variant (HtrA2^{ΔPDZ}) and PDZ-peptide groove mutant were determined. The catalytic efficiency of these variants was lower compared to wild-type HtrA2 due to loss of C-terminal substrate binding, allosteric modulation and hence cleavage. However, interestingly, addition of XIAP to HtrA2^{ΔPDZ} or PDZ-groove mutant enhanced the cleavage rate to a fold increase similar to wild-type HtrA2 (~2.8-fold) with concomitant decrease in cooperativity. Thus, N-terminal ligand is capable of inducing conformational changes in protease domain to form an active conformer.

3. Delineating the global conformational changes following IAP binding

a) Monitoring conformational dynamics using in silico approach

With an aim to understand the conformational changes that govern HtrA2 activity upon IAP binding, we performed molecular dynamics (MD) simulation of BIR3_{pep}-HtrA2^{APDZ} bound and peptide-unbound forms. It was found that BIR3_{pep}-HtrA2^{APDZ} demonstrated significant difference in conformation and dynamics compared to unbound HtrA2. The rmsf plot of the trajectories for the bound form showed higher fluctuations in the regulatory loops LA, LD and L3. Further in-depth analyses of the structural changes were performed by superposition of energy minimized peptide-bound HtrA2 with the unbound form.

b) Structural evaluation of peptide bound and unbound HtrA2

Structural analyses showed that binding of peptide molecule led to stabilization of oxyanion hole and proper positioning of the catalytic triad (Asp65:His95:Ser173). The disorder-to-order transition of loops LD, L3 and L1 coordinate to bring about the observed concerted rearrangements in and around the active-site, which most likely causes activation of HtrA2 upon IAP binding. Based upon these observations we propose a simplistic model of HtrA2 activation. The model assumes that HtrA2 exists in an equilibrium of inactive (E) and most active (E*) states. The complex allosteric propagation mediated upon binding of N- and C-terminal ligands synergistically transforms the protease into the most active conformer.

Conclusion:

Based upon the structural and functional analyses of HtrA2, we clearly demonstrate how precise coordination between ligand binding and flexible loop movements at a site distal from catalytic pocket regulates HtrA2 proteolytic activity. Ligand binding at the N-terminal IAP-binding motif of HtrA2 triggers an allosteric signal essential for the formation of the most active state of the protease in a PDZ-independent yet synergistic manner. Moreover, the peptide-based activation process highlights the possibility of designing suitable peptidomimetics for manipulating HtrA2 functions, specifically for disease intervention.

References:

1. Thierry, F., and Demeret, C. (2008) *Cell Death Differ* **15**, 1356-1363
2. Suzuki, Y., Imai, Y., Nakayama, H., Takahashi, K., Takio, K., and Takahashi, R. (2001) *Mol. Cell* **8**, 613-621.
3. Kersse, K., Verspurten, J., Vanden Berghe, T., and Vandenabeele, P. (2011) *Trends Biochem Sci* **36**, 541-552
4. Yang, J. K., Wang, L., Zheng, L., Wan, F., Ahmed, M., Lenardo, M. J., and Wu, H. (2005) *Mol Cell* **20**, 939-949
5. Siegel, R. M., Martin, D. A., Zheng, L., Ng, S. Y., Bertin, J., Cohen, J., and Lenardo, M. J. (1998) *J Cell Biol* **141**, 1243-1253
6. Dickens, L. S., Boyd, R. S., Jukes-Jones, R., Hughes, M. A., Robinson, G. L., Fairall, L., Schwabe, J. W., Cain, K., and Macfarlane, M. (2012) *Mol Cell* **47**, 291-305
7. Majkut, J., Sgobba, M., Holohan, C., Crawford, N., Logan, A. E., Kerr, E., Higgins, C. A., Redmond, K. L., Riley, J. S., Stasik, I., Fennell, D. A., Van Schaeybroeck, S., Haider, S., Johnston, P. G., Haigh, D., and Longley, D. B. (2014) *Nat Commun* **5**, 3350
8. Suzuki, Y. T.-N., K., Akagi, T., Hashikawa, T., and Takahashi, R. (2004) *Cell Death Differ*. **11**, 208-216.
9. Vande Walle, L., Lamkanfi, M., and Vandenabeele, P. (2008) *Cell Death Differ* **15**, 453-460
10. Li, W., Srinivasula, S., Chai, J., Li, P., Wu, J., Zhang, Z., Alnemri, E., and Shi, Y. (2002) *Nat. Struct. Biol.* **9**, 436-441.
11. Mazat, J., and Patte, J. (1976) *Biochemistry* **15**, 4053-4058

Publications in Refereed Journal:

Published:

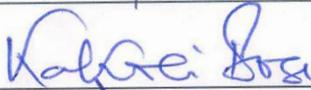
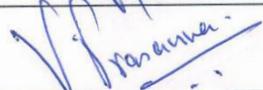
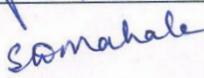
- 1) **Singh N**, D'Souza A, Cholleti A, Sastry GM, Bose K. Dual regulatory switch confers tighter control on HtrA2 proteolytic activity. *FEBS J.* 281 (2014) 2456–2470
- 2) Bejugam PR, Kuppili RR, **Singh N**, Gadewal N, Chaganti LK, Sastry M, Bose K. Allosteric regulation of serine protease HtrA2 through novel non-canonical substrate binding pocket. *PLoS One* 8(2):e55416, 2013
- 3) **Singh N**, Kuppili RR, Bose K. The Structural Basis of Mode of Activation and Functional Diversity: A Case Study with HtrA Family of Serine Proteases. *Archives of Biochemistry and Biophysics*, 516, 2011, 85–96 [Review article]

Date: 04/08/2014



Signature of Student

Doctoral Committee:

S. No.	Name	Designation	Signature	Date
1.	Dr. Rajiv Sarin	Chairman		4-08-14
2.	Dr. Kakoli Bose	Guide		01/08/14
4.	Dr. Ashok Varma	Member		4-08-14
5.	Dr. V. Prasanna	Member		4-08-14
6.	Dr. Smita Mahale	Invitee		4-08-14



Prof. S.V. Chiplunkar
Director, Chairman Academics and Training Program
ACTREC



Prof. K. Sharma
Director-Academics
Tata Memorial Centre

Prof. K.S. Sharma
DIRECTOR - ACADEMICS, TMC
Mumbai - 400 012

LIST OF ABBREVIATIONS

AFC	7-amino-4-(trifluoromethyl) coumarin
APAF1	Apoptotic protease-activating factor 1
BIR	Baculoviral IAP repeat
BPV	Bovine papillomavirus
CARD	CAspase Recruitment Domain
CD	Circular dichroism
cFLIP	cellular FLICE-like inhibitory protein
CIN	Cervical intra-epithelial neoplasia
DBD	DNA-binding domain
DD	Death domain
DED	Death effector domain
DEF	Death effector filament
DISC	Death-inducing signaling complex
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
<i>E.coli</i>	<i>Escherichia coli</i>
EDTA	Ethylene diamine tetraacetic acid
FADD	Fas associated death domain
FITC	Fluorescein isothiocyanate
GFP	Green fluorescent protein
GST	Glutathione S-transferase
HADDOCK	High Ambiguity Driven protein-protein DOCKing
HEK	Human embryonic kidney
HPV	Human Papillomavirus
HR	High risk
HtrA2	High temperature requirement A2
IBM	IAP-binding motif
IGFBP	Insulin growth factor binding domain
IMS	Intermembrane space
IP	Immunoprecipitation
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
KI	Kazal-type S protease inhibitor domain
LCR	Long control region
MBP	Maltose binding protein
MC159	Molluscum Contagiosum 159
MD	Molecular dynamics
MRE	Mean residual ellipticity
NES	Nuclear export signal
NLS	Nuclear localization signal
OMP	Outer membrane porins
ORF	Open reading frame
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PDZ	Postsynaptic density protein 95- <i>Drosophila</i> disc large tumor -Zona occludens1
PEA-15	Phosphoprotein enriched in astrocytes-15

PYD	PYrin Domains
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
RT	Room temperature
RU	Response units
SDM	Site-directed mutagenesis
SEC	Size-exclusion chromatography
SPR	Surface plasmon resonance
TAD	Transactivation domain
TM	Transmembrane domain
XIAP	X-linked inhibitor of apoptosis protein

LIST OF FIGURES

No.	Title	Page No.
2.1	Simplified linear organisation of the HPV18 genome	28
2.2	Diagram representating the differentiation programme in normal epithelium and viral activities in HPV-infected epithelium	31
2.3	Mitochondrion mediated intrinsic pathway of apoptosis	35
2.4	Receptor mediated extrinsic cell death pathway	36
2.5	Domain organizations of selective proteins containing death-fold superfamily domain	38
2.6	Ribbon diagrams of each sub-family of death-fold domain superfamily	39
2.7	Schematic representation of the domain organization of HtrA protein family	45
2.8	HtrA2 mediated apoptotic pathway	46
2.9	Crystal structure of HtrA2 (PDB id. 1LCY)	51
4.1.1	Pull-down assay of E2 TAD with procaspase-8 DED domains	107
4.1.2	Schematic representation of the deletion constructs for procaspase-8 DED-B	108
4.1.3	Pull-down assay of E2 TAD with procaspase-8 DED-B deletion constructs	108
4.1.4	Molecular modeling of procaspase-8 tandem DED	110
4.1.5	Molecular modeling of HPV18 E2 TAD	110
4.1.6	Interaction across HPV18 E2 TAD – procaspase-8 DED interface	111
4.1.7	MBP pull-down assay for E2 wild-type and mutants with procaspase-8 DED-B	113
4.1.8	Dilution experiment with the isolated E2 TAD – DED-B complex	114
4.1.9	Secondary and tertiary structural analyses of E2 TAD wild-type and mutant proteins	114

4.1.10	Expression and localization of procaspase-8 and HPV18 E2 in transfected cells	115
4.1.11	Immunoprecipitation of E2/procaspase-8 complex	116
4.1.12	MBP pull-down assay for wild-type E2 TAD and procaspase-8 DED-B mutants	116
4.1.13	E2-induced cell death analysis	118
4.1.14	E2-FADD interaction analysis	119
4.1.15	Sequence alignment of E2 proteins from different papillomavirus types and structure-based sequence alignment with HPV18	122
4.2.1	Molecular modeling and superposition of tandem DEDs	135
4.2.2	Inter-molecular interface of the procaspase-8 DED/DED complex	135
4.2.3	Mutational analysis of caspase-8 Death Effector Filament formation	139
4.2.4	Sequence alignment of DED-containing proteins and structure-based sequence alignment with FADD	140
4.2.5	Structural modeling of the interactions that may occur between the tandem DEDs to form DED chain	140
4.5.6	Subcellular localization of GFP DED-AB and individual DED domains in co-transfected cells	142
4.2.7	Oligomeric property of caspase-8 DED-AB wild-type and mutant proteins	144
4.2.8	Mutations disrupting the FADD DED filament formation	145
4.2.9	Characterization of FADD mutants	146
4.2.10	Structural modeling of FADD self-association	149
4.2.11	Two possible orientations of the FADD/procaspase-8 interaction	151
4.2.12	Localization of procaspase-8 and FADD in co-transfected cells	152
4.2.13	Localization of procaspase-8 DED-AB and FADD in co-transfected cells	153
4.2.14	Immunoprecipitation of FADD/procaspase-8 complex	153

4.2.15	Fas-induced cell death analysis	154
4.2.16	Proposed mechanism for death-fold mediated procaspase-8 activation	155
4.3.1	Validation of the E2 TAD dimer	169
4.3.2	Homodimeric model of HPV18 E2 TAD	170
4.3.3	Determination of the oligomeric state of E2 TAD wild-type and mutant proteins	172
4.3.4	Equilibrium unfolding of E2 TAD wild-type	175
4.3.5	Concentration dependence of the equilibrium unfolding of E2 TAD wild-type	177
4.3.6	Fraction of species as a function of urea concentration in the unfolding process	177
5.1	XIAP mediated activation of HtrA2	190
5.2	Mutational analysis of XIAP-HtrA2 interaction	192
5.3	SPR binding analyses of HtrA2 with BIR domain proteins and IBM groove peptides	193
5.4	Activity of wild-type HtrA2, its variant and mutant in presence or absence of activator	194
5.5	Determination of steady state kinetics for β -casein cleavage by HtrA2 and HtrA2 ^{APDZ} in presence of IBM peptides	195
5.6	Interaction analyses of PDZ _{opt} with HtrA2 wild-type and G230A mutant measured using ITC	197
5.7	Graphical representation of movements in peptide bound and unbound form of trimeric HtrA2 ^{APDZ}	201
5.8	Structural analysis of peptide bound and unbound HtrA2	203
5.9	Monitoring conformational transition and rearrangements in BIR peptide - bound and unbound trimer HtrA2 ^{APDZ}	205
5.10	Proposed simplistic model for HtrA2 activation	206

LIST OF TABLES

No.	Title	Page No.
3.1	List of oligos for site-directed mutagenesis and sub-cloning of different gene of interest	60
3.2	A typical PCR reaction composition	65
4.1.1	List of interactions across E2 TAD – procaspase-8 DED-B interface evaluated using the Protein Interactions Calculator with the default cut-offs	121
4.1.2	List of binding interface residue areas [\AA^2] (hydrophilic, hydrophobic, and total) calculated with the POPS server	121
4.2.1	List of interactions across DED-A – DED-B inter-molecular interface evaluated using the Protein Interactions Calculator with the default cut-offs	136
4.2.2	List of mutants characterized to identify the critical residues involved in DED/DED interaction	137
4.2.3	Predicted changes in the binding free energy ($\Delta\Delta G_{\text{bind}}$) for all the alanine mutations across the inter-molecular DED/DED binding interface	143
4.2.4	Secondary structural composition and melting temperature of FADD mutants	147
4.2.5	Detailed list of interactions at FADD DED-DED interface evaluated using the Protein Interactions Calculator with the default cut-offs	150
4.3.1	Dynamic light scattering analysis of E2 wild-type and mutant proteins	173
4.3.2	Detailed list of interactions at HPV18 E2 TAD-TAD interface evaluated using the Protein Interactions Calculator with the default cut-offs	174
4.2.3	Detailed list of interactions at HPV16 E2 TAD-TAD interface evaluated using the Protein Interactions Calculator with the default cut-offs	174

4.2.4	Mean parameters for the unfolding of E2 TAD wild-type, as monitored by its fluorescence intensity and calculating the average emission wavelength	178
4.2.4	Energy components contributing to the binding of the E2 TAD – E2 TAD homodimer and E2 TAD – DED heterodimer complexes in terms of MM-GBSA	178
5.1	Surface Plasmon Resonance analysis of the interaction between HtrA2 and XIAP-BIR proteins and peptides	195
5.2	Steady-state enzymatic parameters for HtrA2 and its variants	198
5.3	Comparison of distances between atoms of the catalytic triad in the BIR3 _{pep} -bound and unbound form of HtrA2 (in Å)	200

CHAPTER-1

Introduction

Introduction

In maintaining homeostasis, multicellular organisms tightly couple the rate of cell proliferation and cell death. Disruption of this fine balance due to altered regulation of apoptosis (programmed cell death) leads to several life-threatening diseases such as neurodegenerative disorders and cancer. Cancer, which is characterized by a breakdown in the cellular apoptotic machinery, might occur due to alteration of structural and functional properties of critical proteins in the apoptotic pathway and hence interaction between pro- and anti-apoptotic proteins is closely related to the genesis and progression of this deadly disease.

Tremendous progress in apoptotic research has occurred toward understanding the classical caspase-dependent apoptotic pathways with an aim at targeting them for disease intervention. However, complexity of cancer biology draws interest in identifying alternative mechanisms that can promote cell death. Recently, there have been reports of novel extrinsic adaptor-independent and caspase-independent mechanisms of apoptosis. These unique mechanisms are mediated by proapoptotic proteins, human papillomavirus E2 and serine protease HtrA2 respectively [1-3]. Therefore, characterizing the structural and functional properties of these proapoptotic proteins as well as identifying their binding partners will not only delineate their biological functions but will also help understand the apoptotic pathway better and the role of these proteins in cell death and cancer.

Part-I: Structural and functional characterization of human papillomavirus E2 protein

Human papillomaviruses (HPVs) are causative agents of cervical cancer which is the second largest cause of death in women worldwide, and is even a bigger challenge in developing countries such as in India due to their economic and social impediments. Out of hundred different human genotypes that have been identified, some such as HPV16 and 18 are associated with different stages of cervical cancer and are termed high risk types. On the other hand, the low risk

viruses, such as HPV6 and 11 cause benign warts and are rarely associated with malignant progression.

Papillomaviruses encode eight major proteins that regulate different viral functions. The early protein E2 acts as a central regulator of viral life-cycle regulating viral gene expression, replication as well as mitotic partitioning of viral genome, and thereby represents a pivotal factor for both the productive cycle and persistent infections by HPVs. Apart from these functions, several groups have demonstrated myriad of other functions that are independent of E2 binding to viral genome such as NF κ B activation, induction of apoptosis or regulation of host cell cycle [4]. In particular, proapoptotic activity has been demonstrated for the high risk E2 proteins such as HPV16 and 18 [5, 6]. Although, the exact mechanism of apoptotic induction by E2 is not fully understood, recent literature on HPV18 E2 suggest its involvement in direct interaction with procaspase-8 of external apoptotic pathway which might eventually lead to procaspase activation and hence initiation of caspase cascade [7, 8].

The extrinsic apoptotic pathway is triggered by ligation of cell-surface “death receptors” followed by formation of multiprotein death-inducing signaling complex (DISC). The DISC comprises oligomerized death receptors such as Fas, the adaptor protein FADD (Fas associated death domain), procaspase-8, and cellular FLICE – like inhibitory proteins (cFLIPs). In case of Fas mediated signaling, Fas ligand binds to Fas receptor leading to receptor trimerization followed by binding of FADD to the receptor through its death domain. FADD then interacts with procaspase-8 or -10 through homotypic interactions involving death effector domains (DED) of the partners which activate caspase-8 or -10 eventually leading to apoptosis. Interestingly, E2-procaspase-8 interaction has been proposed to bypass the requirement of upstream adaptor proteins which are essentially required for DISC formation, thereby representing a *novel adaptor-independent caspase activation pathway*. Based on these evidences we hypothesized that E2 influences external cell death pathway by interacting with procaspase-8 which might lead to a change in its conformation or stability in turn enhancing its capability to oligomerize and hence

promoting apoptosis. Therefore, our goal was to delineate the mechanism of this novel interaction between high-risk HPV18 papillomavirus E2 and procaspase-8 and how it aids in E2-induced apoptosis using interdisciplinary approach. This information would provide a comprehensive picture of the novel adaptor-independent mechanism of procaspase-8 activation and hence establish a model for E2-induced apoptosis in high risk HPV types. It might also be utilized in future studies to design E2 analogs so as to modulate procaspase-8 activation and hence apoptosis.

Part-II: Structural and functional characterization of serine protease HtrA2

HtrA2 (High temperature requirement A2) protein belongs to a unique family of serine proteases that are conserved from prokaryotes to humans. The most elaborately studied protein in this family is *E. coli* periplasmic protein DegP/HtrA that has a dual chaperone and temperature-dependent protease activity [9]. To date four human homologs (HtrA1-HtrA4) of DegP have been identified of which HtrA2 has proapoptotic activity while very little information is available on other human HtrAs [10].

HtrA2 undergoes maturation by autocatalytic N-terminal processing resulting in processed ~36 kDa protein that comprise an N-terminal domain, a serine protease and a PDZ domain (protein-protein interaction domain that primarily binds to the C-terminus of interacting proteins). Moreover, mature HtrA2, has an N-terminal IAP-binding motif (IBM) with which it interacts with the BIR domain of inhibitor of apoptosis proteins (IAPs) such as XIAP, cIAP1 and cIAP2 and relieves the inhibition on active caspases thus promoting apoptosis [11, 12]. Although, HtrA2 has primarily been identified as an IAP-binding proapoptotic protein, its other functions such as caspase-independent induction of apoptosis and serine protease activity are poorly characterized. Recent studies have identified a few antiapoptotic binding partners/substrates of HtrA2 such as PEA-15 suggesting its proapoptotic and proteolytic functions might converge [13]. Moreover, recently it has been observed that the IAPs are also substrates of HtrA2 although the binding

regions for these two proteins are completely different (N-terminal tetrapeptide for IAPs and C-terminal PDZ for PEA-15).

Looking into its complex trimeric three-dimensional structure and trying to account for its low protease activity and narrow substrate selectivity, researchers hypothesized a model which suggests intricate PDZ-protease coordination, rearrangement in their relative orientations and huge conformational changes at PDZ-protease interface are prerequisites for peptide binding and substrate cleavage by HtrA2 [14]. Although PDZ acts as a regulatory domain in all the members of HtrA family, uniqueness of HtrA2 is manifested by its ability to bind subset of proteins, such as IAPs, through its N-terminus and subsequently cleave them [3, 15]. This phenomenon emphasizes multiple modes of HtrA2 activation and regulation, the precise mechanism for which remains to be elucidated. Therefore, our aim was to delineate the structural correlates of HtrA2 activation as well as to develop a universal model for its mechanism of action. A clear understanding of the structural determinants of HtrA2 mediated substrate recognition and cleavage will help define ways of regulating its functions.

CHAPTER-2

Review of Literature

Part-I

Background:

Transforming viruses have the potential to change a normal cell into a cancerous cell during their normal life cycle. Persistent infections with these viruses have been identified to cause various types of cancer. These viruses have been implicated in the modulation of various biological processes, such as proliferation, differentiation and apoptosis [16]. The study of infections caused by oncogenic viruses has immensely helped in our understanding of several mechanisms that regulate cell growth, as well as the molecular alterations leading to cancer. Therefore, these viruses provide models of study that have enabled the advances in cancer research. Viruses with transforming abilities include different members of the Human Papillomavirus (HPV) family, Hepatitis C virus (HCV), Human T-cell Leukemia virus (HTLV-1), Epstein Barr virus (EBV) and Kaposi's sarcoma Herpes virus (KSHV) [17].

2.1 HUMAN PAPILLOMAVIRUS

HPVs are the small, non-enveloped, double stranded DNA viruses that infect the cutaneous and mucosal epithelium. Specific HPV types have been linked to the development of cervical carcinoma as well as benign genital warts. Over 120 different types of HPV have been identified to date with varied pathogenicity [18]. Based on their association with cervical cancer, HPVs are broadly classified into three groups:

1. Low risk HPV: They infect genital tract and bring about the formation of benign hyper proliferative lesions more commonly known as warts. To date, 12 types are identified under this group including HPV6, 11, 40, 42, 43, 44, 54, 61, 70, 72, 81 and CP6108 [19].
2. Probable high risk HPV: They are rarely involved in carcinogenicity. Three types are classified in this group including HPV26, 53 and 66 [19].

3. High Risk HPV: They infect the basal layer of cervical epithelium and account for more than 95% of all cases of cervical cancers. They are also found to be involved in genital malignancies as well as head and neck cancer. Several HPV types are included under high risk group including HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73 and 82 [20].

HPV16 and 18 are acknowledged to be associated with 70% of all cervical cancer cases. HPV18 is the second highest risk factor next to HPV16 and said to be clinically involved in more aggressive cervical cancer [21]. Intriguingly, HPV16 is integrated into the host genome in 72% of all invasive cancers while HPV18 shows 100% integration [22].

HPV has a circular genome approximately eight kilo bases (kb) in size and organized into eight open reading frames: a long local control region (LCR), six early proteins (E1, E2, E4, E5, E6, E7) and two late proteins (L1, L2) (**Fig. 2.1**). HPV early promoters transcribe the polycistronic transcripts which primarily encode E6, E7, and E2. The late promoter is activated during the productive phase of the viral life cycle which results in high level transcription of E1, E2, E4, E5 as well as the capsid proteins L1 and L2. Functions of E3 and E8 genes are not yet known [23].

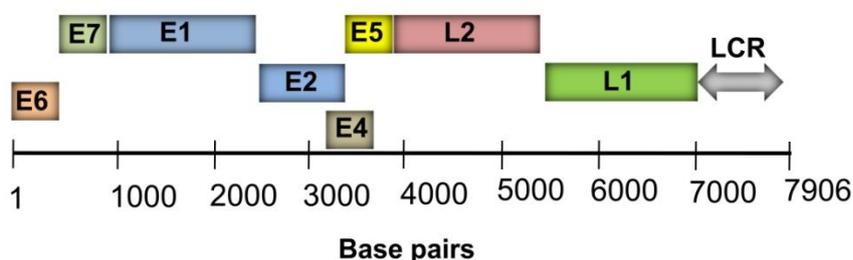


Figure 2.1 Simplified linear organisation of the HPV18 genome. Diagram representing HPV18 genome in its linearised state (scale bar is in kilobase pairs). The rectangles represent the positions of various open reading frames. The E genes encode proteins that are produced early in the infectious cycle (the non-structural proteins), whereas the L genes encode proteins that are produced late in infection (virion proteins that are necessary for virus assembly). LCR represent the long control region of the genome.

2.1.1 Functions of papillomavirus proteins

- 1) **E1 and E2:** are the two viral proteins that are required for viral DNA replication, together with the host cell DNA replication machinery [24].
- 2) **E4 and E5:** E4, although originally classified as early protein, is the most abundant protein produced in the viral life cycle. It interacts with cytoskeletal proteins and induces collapse of cytokeratin matrix, which is particularly compact in different keratinocytes [25]. E5 is a dimeric membrane protein of around 80 amino acids in length. It seems to activate the proliferative capacity of the differentiating cells, probably by increasing the cell sensitivity to growth factors [26].
- 3) **E6 and E7:** Products of E6 and E7 genes stimulate cell proliferation and promote immortalization and hence aptly termed as oncogenes. E6 is a zinc finger protein that mostly interacts with tumour suppressor protein p53 leading to its degradation via proteasome dependent pathway [27]. The oncogenic effect of E7 derives from its ability to interact with a number of cell cycle regulators, one of its common binding partner retinoblastoma tumour suppressor proteins and perturb their functions [28].
- 4) **L1 and L2:** These two proteins form the viral capsid; they are expressed late in infection, in the upper layers of the epithelium [29].

Part of the viral genome is a non coding region, called local control region (LCR), which contains most of the regulatory DNA sequences needed for proper replication of the viral genome (origin of DNA replication) and for the expression of the viral genes (enhancer and promoter regions) [20].

2.1.2 HPV life cycle and pathogenesis

HPVs are exclusively epitheliotropic, and their life cycle is intimately linked to the differentiation process of the host cells [30]. Normal squamous epithelia that grow as

stratified layers with those in the basal are able to actively divide. After division, one of the daughter cells migrates upward and begins to undergo differentiation while the other remains in the basal layer as a slow-cycling, self-renewing population. Productive papillomavirus infection begins when HPV virions infect the basal layers of the epithelia, most likely through micro-wounds, and enter cells via an unidentified receptor. Following migration to the nucleus, the viral genome is maintained in these cells as a stable episome at low copy number usually 20 - 100 copies per infected basal cell. The early HPV genes E1 and E2 support viral DNA replication and its segregation such that the infected cells can be maintained in the lesion for a long period. Following viral genome replication and cell division, one of the daughter cells migrates away from the basal layer and initiates a program of differentiation. The normal uninfected keratinocytes exit the cell cycle after leaving the basal layer, and the nuclei are degraded in many of these differentiating cells. In contrast, HPV-infected cells undergo differentiation but remain active in the cell-cycle. This allows highly differentiated super-basal cells to re-enter S-phase and support high-level productive viral DNA replication. The ability of differentiated basal cells to undergo cell cycle progression is mediated largely through the action of the viral E7 protein. Differentiation dependent amplification of viral genomes in super-basal cells coincides with activation of the late viral promoter. The late viral transcripts encode the capsid proteins L1 and L2 as well as two additional mediators' of late viral functions, E4 and E5. Progeny virions are assembled in highly differentiated cells and then released to the extra-cellular environment to re-initiate infection (**Fig. 2.2**).

HPV infections are found in almost all pre-invasive lesions, so called 'cervical intra-epithelial neoplasia' (CIN) and invasive cancers. Current epidemiological and fundamental data have confirmed infection with HPV to be the basic cause of the disease.

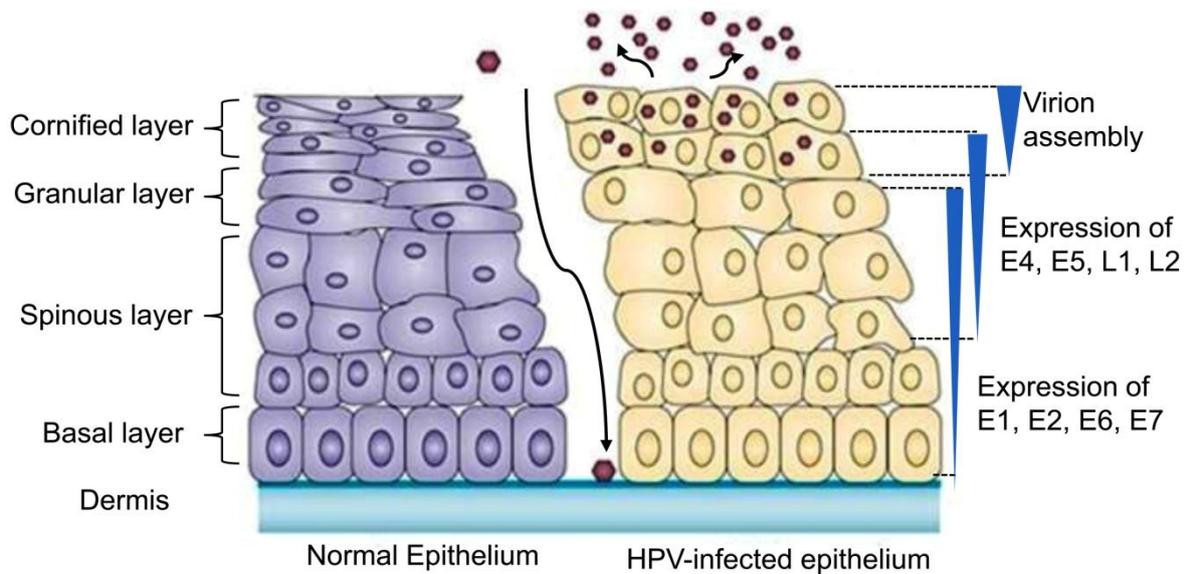


Figure 2.2 Diagram representing the differentiation programme in normal epithelium (left) and viral activities in HPV-infected epithelium (right). HPV virions infect the dividing cells of the basal layer. Unlike normal epithelium where cells exit the cell cycle upon initiation of differentiation, HPV-infected cells maintain cell division to allow genome amplification and virion production (modified from [31]).

2.1.3 Current cervical cancer therapies

The current treatment for cervical cancer includes surgery, radiotherapy or their combination. Pre-cancerous lesions can be treated using a variety of approaches including cryotherapy, laser therapy, electro-surgery, and surgical excision. Unfortunately, still this disease is responsible for over 1000 deaths per year and at best only around 70% of the women with this condition survive for 5 years after diagnosis. New approaches to the treatment and prevention of both cervical cancer and pre-cancerous lesions would obviously be of great potential benefit. Prophylactic vaccines although have shown promise will not be very helpful for the millions of women already suffering from this deadly disease and also for immunosuppressed individuals. Moreover, there are major economic and social challenges to widespread use of an expensive vaccine against a sexually transmitted disease in some populations such as in India. Hence, papillomavirus proteins that represent highly specific targets for therapeutic intervention have therefore been the focus of cervical cancer research.

Recent research has found several interesting targets for targeted therapy. The E1 helicase is the only HPV protein that has enzymatic activity and has therefore been the subject of intense investigation. Drugs that can inhibit E1 helicase based on selecting compounds that block ATP dependent DNA strand displacement have been identified [32]. However, unfortunately, these drugs bind to all ATPase active site and inhibit its activity thus limiting its applicability. Alternatively, the interaction of E1 and E2 provides another target for drug discovery. However, drugs that block a specific protein-protein interaction are difficult to select or design. Kasukawa et al. designed short peptides based on the amino acid sequence of the N-terminus of HPV16 E2 which blocked the interaction of E1 and E2 and the replication of HPV16 and HPV11 *in vitro* [33]. If drugs could be developed that mimic the activity of these peptides, they could be targeted in the treatment of pre-cancerous HPV-induced lesions. Since functional E1 protein is not required by HPV-induced cervical tumour cells, these drugs would probably not be useful in the treatment of cervical cancer. E6 and E7 are required for the growth of HPV transformed cells, therefore these proteins are the obvious targets for drug discovery.

At first thought, the E2 protein might seem as an improbable target for drug research since, like E1, it is nonessential for the growth of tumor cells. However, a drug that blocks E2 function associated with HPV replication might prevent HPV-induced tumorigenesis. Compounds that block the DNA-binding activity of HPV31 E2 protein have been identified and these could lead to the development of specific inhibitors of HPV replication [34]. In addition, it is reported that Bovine papillomavirus (BPV) E2 protein can suppress the growth of HPV-transformed cells [35], and the HPV16 and HPV18 E2 proteins can induce programmed cell death [5]. The dual anti-proliferative functions of E2 make this protein a putative therapeutic tool in cervical carcinomas. The use of E2 would have two main advantages. First, E2 could kill HPV-transformed cells by apoptosis. Growth suppression is

not very useful since without stable transformation, E2 protein expression would eventually be lost and growth could continue. The second advantage of using E2 is that it might evoke an immune response. E2 immunity is thought to play a role in the control of HPV infection, and induced immunity to E2 could be protective against further infection [36]. However, to be an effective treatment, either the action of the E2 protein must be specific to cancer cells, or it must be delivered only to cancer cells *in vivo*. Viral delivery of BPV E2 has been tested in an animal model and seems to suppress tumour growth effectively [37-39]. These recent works have unrevealed new avenues of research in the treatment of cervical cancer, and highlighted the potential of viral proteins including E2 in cervical cancer therapies.

2.2 HUMAN PAPILLOMAVIRUS E2 PROTEIN

The E2 protein plays critical roles in the viral life cycle, as it regulates viral DNA replication, transcription as well as long-term maintenance of the viral genome in dividing cells. Replication of the viral DNA is initiated by the cooperative binding of E1 and E2 proteins to their binding sites in the origin of replication to form an efficient replication initiation complex. A major role of HPV E2 protein, as transcription regulator, is repression of the transcription of E6 and E7 viral oncogenes by binding to the same sites as those used for DNA replication. E2 binding to these sites interferes by steric hindrance with the binding of cellular transcription factors such as SP1 and TFIID, which are essential for E6 and E7 transcription [7]. However, this function of E2 is alleviated due to disruption of the E2 gene by integration of the viral genome in the host genome during HPV-associated carcinogenesis. Integration typically occurs such that the E2 open reading frame (ORF) is disrupted, thereby eliminating E2-mediated transcriptional control of the early promoter resulting in increased production of the E6 and E7 oncoproteins. This up-regulated expression of the two viral oncogenes appears to play a key role in cellular transformation [40].

Apart from its role in viral life cycle, E2 proteins can also directly influence the host cell biology independently of their binding to the viral genome and of their transcriptional properties. The pro-apoptotic activity of E2 is one of the first described E2 functions independent of its binding to the viral genome [4]. Furthermore, E2 apoptotic activity is not cell-type specific, and is mediated by the amino-terminal domain independently of its associated transcriptional functions [41]. A striking aspect of the involvement of E2 in apoptosis is that it has been found to be specific for the high risk (HR) HPV E2 proteins [5]. Induction of apoptosis by HR-HPV E2 proteins is linked to their nucleo-cytoplasmic shuttling. However, inability of low risk HPV types to induce apoptosis might be attributed to their exclusive nuclear localization [42]. The fact that only the high risk E2 proteins have evolved a proapoptotic activity suggests that this activity could compensate for some of the transforming functions of the E6 and E7 oncogenes. The E6 and E7 proteins of HR-HPV interfere with cell cycle exit and block cellular differentiation, which are required for the viral vegetative cycle. These modifications of the host cell biology induce greater resistance to apoptosis [30]. It is speculated that the pro-apoptotic activity of E2 would be required to overcome these effects, contrary to the low risk infected cells where many of the E6 and E7 functions are not conserved, thereby facilitating virus release. The proapoptotic activity of high risk E2 proteins, however, could explain why the high risk viral genomes integrate into the host DNA preferentially with disruption of the E2 ORF, leading to higher transcription of the E6 and E7 oncogenes and subsequent neoplastic progression [22, 43].

2.2.1 HPV18 E2-induced apoptosis

APOPTOSIS

Cell death by apoptosis is a normal and energy dependent process initiated by cellular damage, stress or number of endogenous and extracellular stimuli. It is accompanied by decrease in cell volume, nuclear changes with chromatin condensation followed by DNA fragmentation and membrane blebbing finally leading to its disintegration into smaller apoptotic bodies [44].

Proteases play a key role in apoptosis through cleavage of several proteins that are critical for normal functioning of the cellular machinery. A major class of proteases called caspases regulates the apoptotic machinery. Upon receiving specific apoptotic signals, these proteases (either present or are released in the cytosol) get activated and lead to a cascade of events leading to proteolytic cleavage of key proteins that are required for normal cellular functions. There are two major pathways of apoptosis that are initiated upon activation of effector caspases by the formation of large multi-protein complexes.

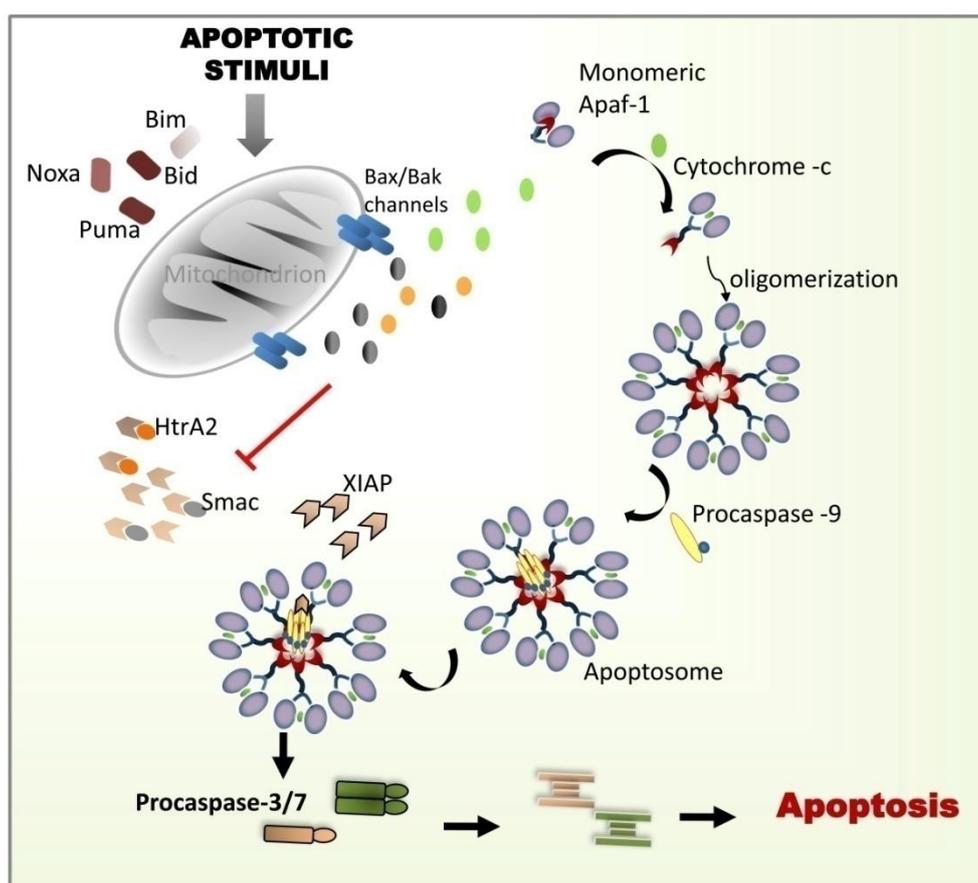


Figure 2.3 Mitochondrion mediated intrinsic pathway of apoptosis. Cellular stress activates p53, a cell-cycle check-point protein. p53 in turn initiates the intrinsic pathway by up regulating Bcl-2 family proteins, Puma and Noxa, which further activates Bax, Bak. Oligomerization of Bax-Bak complex at the outer membrane of the mitochondria results into membrane permeabilization, thereby resulting in efflux of cytochrome-c and other anti-apoptotic proteins such as HtrA2, Smac/Diablo. The released cytochrome-c binds an apoptotic protease-activating factor 1 (Apaf-1) monomer, leading to its oligomerization into a heptameric wheel-like structure called the apoptosome that later recruits and activates initiator procaspase 9. Active caspase-9 exposes the ATPF motif that binds to X-linked inhibitor of apoptosis protein (XIAP) preventing further activation. The mitochondrial protein

Smac/DIABLO augments apoptosis by binding XIAP and reversing their grip on active caspase-9. Activated caspase-9 then cleaves and activates effector caspases 3, 7 to trigger apoptosis.

i) Intrinsic pathway: This pathway is triggered due to a variety of stress conditions such as UV irradiation, cytotoxic drugs or growth factor withdrawal which induce mitochondrial perturbation resulting in the release of cytochrome c. The released cytochrome-c binds an apoptotic protease-activating factor 1 (APAF1) monomer, leading to its oligomerization into a heptameric wheel-like structure called '*apoptosome*' that later recruits and activates initiator procaspase 9. Activated caspase-9 further cleaves and activates effector caspases 3 and 7 to trigger apoptosis [45] (**Fig. 2.3**).

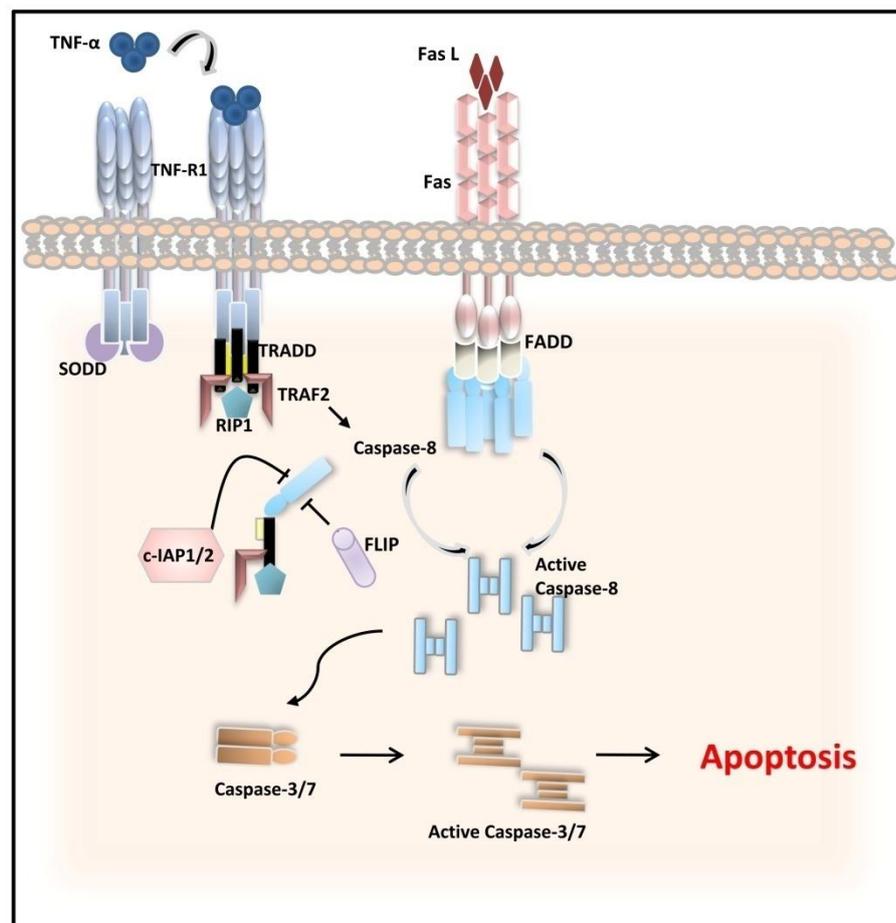


Figure 2.4 Receptor mediated extrinsic cell death pathway. In the extrinsic pathway, death ligands, such as FasL, TNF or TRAIL engage their cognate receptors Fas, TNFR or DR5 respectively. Engagement of death receptors with their ligands initiates the recruitment of adaptor proteins such as FADD or TRADD, which in turn recruit zymogenic initiator

procaspase-8 to form the DISC. DISC recruitment leads to auto-processing and thus activation of caspase 8. Active caspase-8 then proteolytically cleaves and activates executioner caspases-3, -6 and -7 that further culminates in substrate proteolysis and hence cell death. The pathway is tightly regulated with several inhibitory molecules at various stages of the signaling. In TNFR-induced mechanism, the activation of preassembled receptor is blocked by a complex called silencer of death domain (SODD). Fas mediated pathway has a negative regulator molecule FLIP (FLICE-like inhibitory protein). FLIP interferes with DISC functions by preventing the formation of catalytically active caspase-8.

ii) Extrinsic pathway: In the extrinsic pathway, death ligands, such as FasL, TNF or TRAIL engage their cognate receptors Fas, TNFR or DR5 respectively. Engagement of death receptors with their ligands initiates the recruitment of adaptor proteins such as FADD (Fas associated death domain), which in turn recruits zymogenic initiator procaspase-8 to form huge protein complex called the DISC. This increases the local concentration of these monomeric proenzymes through proximity induced by the adaptor molecules leading to their dimerization, removal of prodomain and subunit rearrangements to form active caspase 8. The active protease further cleaves and activates executioner caspases 3, 6 and 7 that later culminates in substrate proteolysis and hence cell death [46] (**Fig. 2.4**).

Several interactions between HR-HPV E2 and mediators of both the extrinsic and the intrinsic pathways have been identified over the years, notably involving p53 [47], caspase-8 [48] and cFLIP [49]. Although, the exact mechanism of apoptotic induction by E2 is yet not fully understood, recent literature suggests it involves direct protein-protein interaction of HR-HPV18 E2 with procaspase-8 through a unique pathway that bypasses the requirement of homotypic interactions between death-fold domain adaptor proteins in DISC [1].

2.2.2 Structure and function of typical death-fold domain

The 102 members discovered in humans so far belonging to the death-fold superfamily are categorized in four subfamilies: 39 Death Domains (DDs), eight Death Effector Domains (DEDs), 33 Caspase Recruitment Domains (CARDs) and 22 Pyrin Domains (PYDs). Domain organization of the members of each sub family is shown in **Fig. 2.5**. This superfamily is

hallmarked by its classic structure, the so-called death-fold, which consists of a globular architecture wherein six amphipathic α -helices are arranged in an antiparallel α -helical bundle with Greek key topology (**Fig. 2.6**).

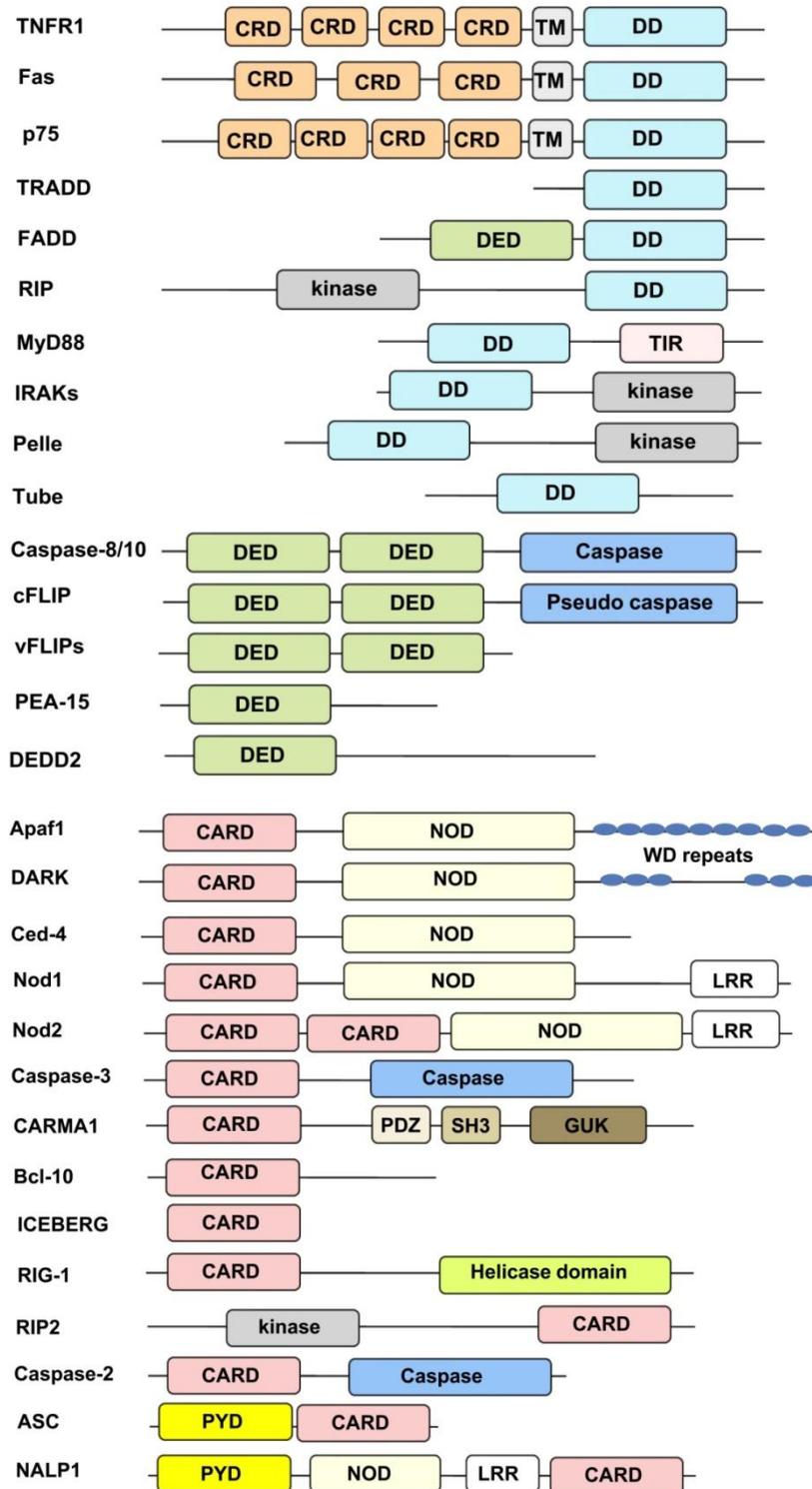


Figure 2.5 Domain organization of selective proteins containing death-fold superfamily domains (modified from [50]).

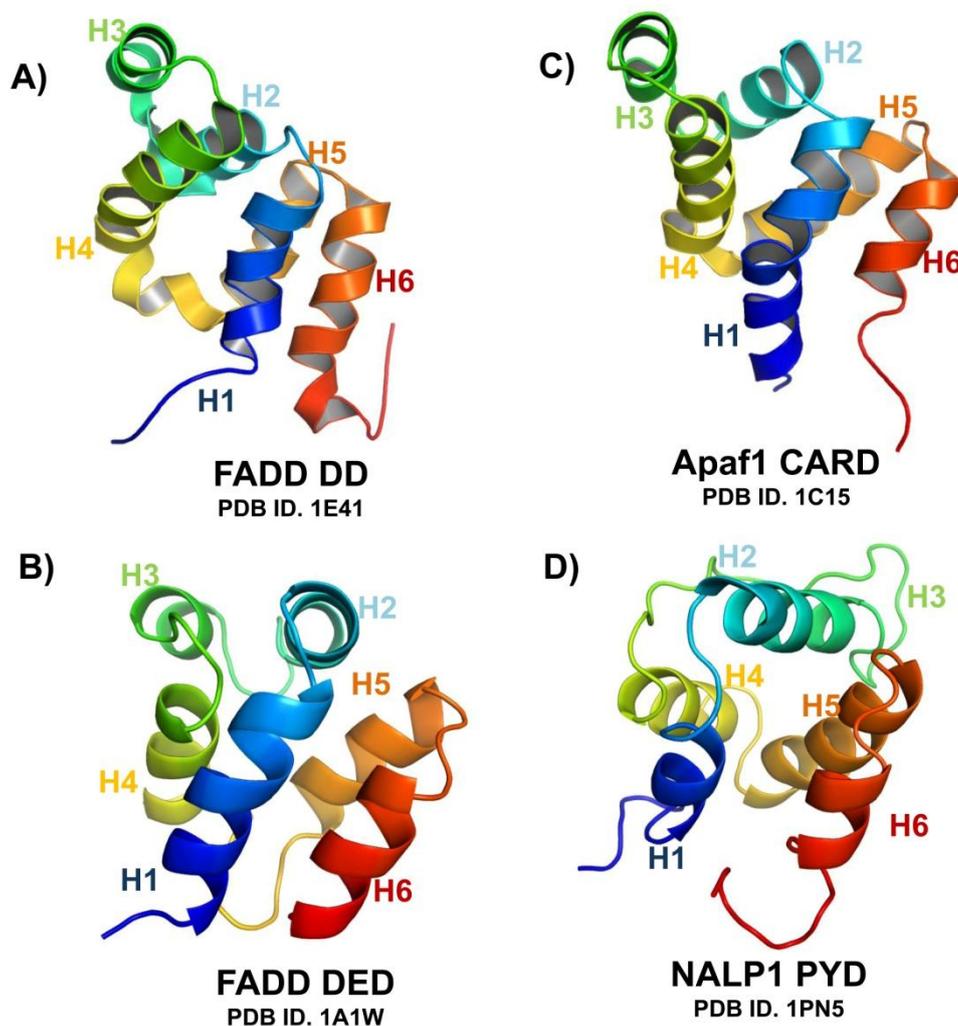


Figure 2.6 Ribbon diagrams of each sub family of death-fold domain superfamily. A) FADD DD, B) FADD DED, C) Apaf1 CARD, D) NALP1 PYD. All the six amphipathic α -helices are represented in different colors.

Structural differences between individual domains are the result of alterations in the length and orientation of the α -helices, and the distribution of charged and hydrophobic residues along the surface. Functionally, death-fold domains mediate the assembly of large, oligomeric signaling platform in which crucial effectors, such as caspases gain activity. The ability of death-fold domains to mediate the formation of intricate, higher order complexes resides in their ability to mediate homotypic interaction within the members of the sub families.

Firstly, DDs assemble into oligomeric complexes, which vary in stoichiometry depending on the particular DD and signaling pathway. This is crucial for the second role carried out by CARD and DED modules, which is recruitment of downstream effector proteins. CARDS and DEDs homotypically interact with other CARD or DED present on the effector proteins such as caspases and kinases, bringing them sufficiently close for zymogen processing and subsequent activation of the apoptotic pathway [51]. Since E2 is reported to influence the extrinsic death receptor pathway, in this section of the review, we have highlighted only the functional context and structural organization of the death domain and death effector domains.

2.2.2.1 Death domain (DD) structures

Currently, structures of six isolated DDs are available, which includes four NMR structures for Fas DD [52], FADD DD [53], TNFR1 DD [54] and p75 DD [55], and two crystal structures of IRAK4 DD [56] and RAIDD DD [57]. Since these domains are involved in protein-protein interactions and have a tendency to aggregate, most of these structures were solved under non-physiological conditions such as extreme pH and/or with 'de-aggregating' mutations. While all these DDs exhibit the characteristic six-helical bundle fold, differences exist in the length and direction of the helices. Although structurally quite similar, their low sequence homology among DDs thereby the surface features of these domains are also entirely different, which may be responsible for their specificity in protein-protein interactions.

DD/DD interaction

Most knowledge concerning interactions of death-fold domains in the context of oligomeric protein assemblies are obtained from crystal structures containing DDs. The structural basis of DD/DD interactions, such as those in the Fas DD: FADD DD complex [52], the MyD88 DD: Tube DD complex [58] and the PIDD DD: RAIDD DD complex [59], have recently been

illuminated. Regardless of the details of the interaction, the biggest surprise is perhaps the asymmetry of the interaction, considering what might have been expected for homotypic interactions. The ability of death-fold domains to mediate the formation of higher order complexes have been shown to reside in their ability to engage in three distinct types of asymmetric interactions, type I, II and III. A type I interaction is formed when residues from helices 1 and 4 of one death-fold domain interact with residues from helices 2 and 3 of another (Figure 2a). A type II interaction involves residues from helix 4 and the loop between helices 4 and 5 of one domain interacting with residues of the loop between helices 5 and 6 of the other death-fold domain. A type III interaction is formed when residues from helix 3 interact with residues located on the loops between helices 1 and 2 and helices 3 and 4 of the other death-fold domain. The involvement of multiple interaction surfaces in DD complexes is supported by mutational data, which showed wide spreads of residues important for binding and/or function on the surfaces of Fas [60], FADD [61], TRADD [62] and TNFR1 [63].

2.2.2.2 Death effector domain (DD) structures

Currently only three DED structures are available, which include the NMR structures of FADD DED [64] and PEA-15 DED [65] and the crystal structure of the tandem DED of MC159 viral FLICE - like inhibitory protein (vFLIP) [66]. The NMR structure of FADD DED is also available in the context of its full length structure comprising both a DED and a DD [67]. While FADD is a component of the DISC, PEA-15 appears to participate in MAP kinase activation through a non-homotypic interaction with the kinase ERK [68]. MC159 v-FLIP is shown to inhibit caspase activation at the DISC [66]. Consistent with their assignment to the DED subfamily, the structures show the conserved six-helical bundle fold; however they are more similar to each other than to other members of the death-fold domain superfamily.

Surface features

DEDs share two prominent conserved surface features, which distinguish them from the other members of the death-fold domain superfamily. The first feature is the presence of a conserved hydrogen-bonded charge triad interaction revealed by the high resolution structure of MC159 vFLIP DED [66]. The charge triad is formed by the E/D-RxDL motif, involving an acidic residue in helix H2 and the Arg and Asp residues in the RxDL motif in helix H6. Extensive hydrogen bonding interactions are observed among the charged side chains with the Arg residue situated in between the two acidic residues. It is speculated that these hydrogen bonds likely help to maintain a precise organization of death-fold structure. The charge triad is found highly conserved in most single and tandem DEDs. However, except p75 DD, this motif is absent in all other members of the death-fold domain superfamily suggesting that it is a characteristic feature of DEDs alone. The second surface feature is the conserved hydrophobic patch formed mostly by residues on H2. This was first observed in the NMR structure of FADD DED and later shown to be conserved in most tandem DEDs as well.

DED/DED interaction as seen in tandem DED

The structure of MC159 vFLIP presented the first and the only available glimpse of a DED/DED interaction [66]. Instead of beads on a string, DED1 and DED2 interact with each other intimately to form a rigid, dumbbell shaped structure. The two DEDs are related approximately by a translation across the contact interface so that one side of DED1 is contacting the equivalently opposite side of DED2. The translational relationship between DED1 and DED2 is made possible by helix H7 of DED1. This interaction between DED1 and DED2 differs from the known Pelle DD: Tube DD interaction [69] and the Apaf-1 CARD: caspase-9 CARD interaction [70]. The interaction at the DED1/DED2 interface is mostly hydrophobic, mediated by helices H2 and H5 of DED1 and helices H1 and H4 of DED2. The

interfacial residues, especially those which contribute large surface areas and are completely buried, such as F30, L31, F92, L93 and R97 are mostly conserved in tandem DEDs. This suggests that all known tandem DEDs may form a similar rigid compact structure as MC159.

2.2.3 Structure and function of HPV E2 protein

E2 has a modular structure comprising a conserved N-terminal transactivation domain or TAD (~ 201 amino acids) that regulates both viral transcription and DNA replication, followed by an unstructured flexible hinge region and a DNA binding domain (DBD) of around 80 amino acids. The three-dimensional structures of the TAD and the DBD of E2 have been reported. The DBD is composed of a compact anti-parallel β -barrel dimer that carries a pair of surface-exposed α helices which make specific contacts with the base pairs in two successive major grooves of the DNA. It recognizes inverted repeats with the consensus sequence 5'- AACCGN4CGGTT-3', where N represents any base pair. The TAD contains two structural sub-domains: an N-terminal region comprising three anti-parallel α -helices and a β -sheet rich C-terminal domain. These sub-domains are separated by two consecutive single turns of helical structure which packs tightly against the two domains. Interestingly, the proapoptotic E2 protein with no classical death-fold domain structure directly binds to the DED containing prodomain of caspase 8 promoting its activation [1].

Despite the fascinating progress in understanding caspase-8 activation, several issues have not been well understood yet. Specifically, the molecular and structural basis of DISC assembly has mostly remained unsolved. Moreover, the regulatory switch for procaspase-8 activation at DISC remains yet to be identified. To add further to the complexity, the finding that viral E2 protein could induce caspase-8 activation through direct interaction raises intriguing questions pertaining to the structural determinants of E2 – procaspase-8 interaction and how a non-death-fold domain protein mediates this interaction.

Part-II

Background:

The HtrA (high-temperature requirement protease A) family of serine proteases was initially identified in *E. coli* as a heat shock-induced envelope-associated serine protease [71]. At low temperatures it generally acts as a molecular chaperone, while, with increase in temperature, it exhibits the proteolytic activity [72, 73]. In humans, HtrAs are involved in numerous cellular processes, ranging from maintenance of mitochondrial homeostasis to cell death in response to stress-inducing agents. In particular, emerging evidence support the role of human HtrAs in multiple pathways of programmed cell death and chemotherapy-induced cytotoxicity. This multifaceted ability associates them with different pathological conditions such as cancer, neurodegenerative disorders and arthritis including myocardial ischemia/reperfusion injury hence making them therapeutically important [2, 74-76].

HtrA proteins belong to a unique family of oligomeric serine proteases (S1B class) characterized by the presence of a trypsin-like fold domain, and have been found to be highly conserved from prokaryotes to humans [77]. Until now, four human homologues of *E. coli* HtrA have been identified: HtrA1 (L56 or PRSS11) [78], HtrA2/Omi [79], HtrA3 [80] and HtrA4 [81] (**Fig. 2.7**). All mammalian HtrA proteins share a highly conserved chymotrypsin-like serine protease domain and one PDZ domain. PDZ is an acronym combining the first letters of three proteins — post synaptic density protein (PSD95), *Drosophila* disc large tumor suppressor (Dlg1), and zonula occludens-1 protein (zo-1) — which were first discovered to share the domain. Although there is an overall structural conservation, significant diversity at the N-terminus construes their cellular localization and functionality. Strikingly, the structure of N-terminal regions of mammalian HtrA1, 3 and 4 are distinct from that of HtrA2/Omi. The N-termini of HtrA1, 3 and 4 contain predicted signal peptides as well as domains that are recognized as IGF binding and protease inhibitor domains, which include

an N-terminal signal peptide (SP), an insulin growth factor binding domain (IGFBP) and a kazal-type S protease inhibitor domain (KI) [82]. Mammalian HtrA2 lacks all these N-terminal domains but has a transmembrane segment or regulatory domain (TM). It is interesting to make a note that of the four human HtrA family members, HtrA2/Omi is the only member of the family with a clear intracellular localization and shown to be processed to reveal a Reaper-like amino-terminal motif [10].

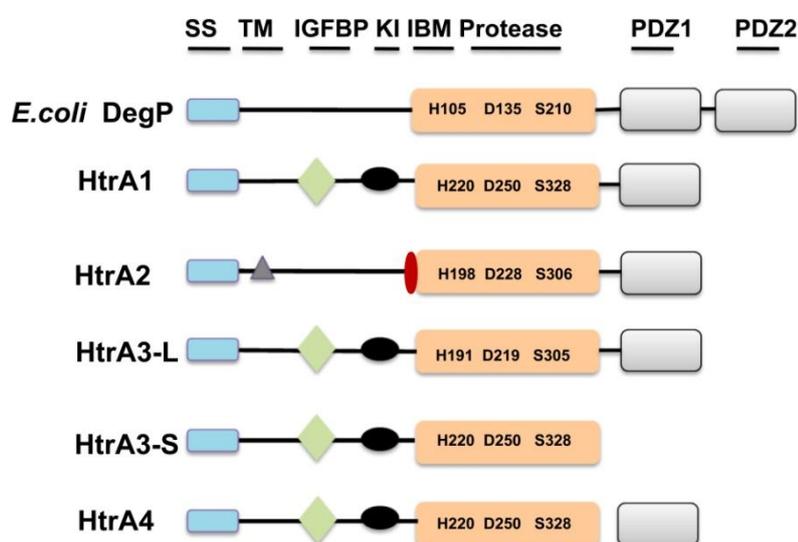


Figure 2.7 Schematic representation of the domain organization of HtrA protein family. The protease domain is in orange rectangles, PDZ1 and PDZ2 domains in grey rectangles, SS (signal sequence) is shown as blue rectangles, TM (transmembrane domain) as black triangle, IGFBP (Insulin-like growth factor binding) in green diamond, IBM (IAP-binding motif) and KI (Kazal protease inhibitor domain) as red and black ovals respectively. *E. coli* represents *Escherichia coli*, HtrA-high temperature requirement protease A. Catalytic triad residues are represented as H (histidine), D (aspartate), and S (serine).

HtrA1 and the closely-related HtrA3 and HtrA4 are secretory proteins [83-85]. HtrA1 has been implicated in a number of malignancies [74, 84, 86] while HtrA3 has been found to be up-regulated in the placenta [87]. Both HtrA1 and HtrA3 are primarily involved in protein quality control in the extracellular matrix. Among the HtrA homologues, HtrA2 is most well characterized member of this family.

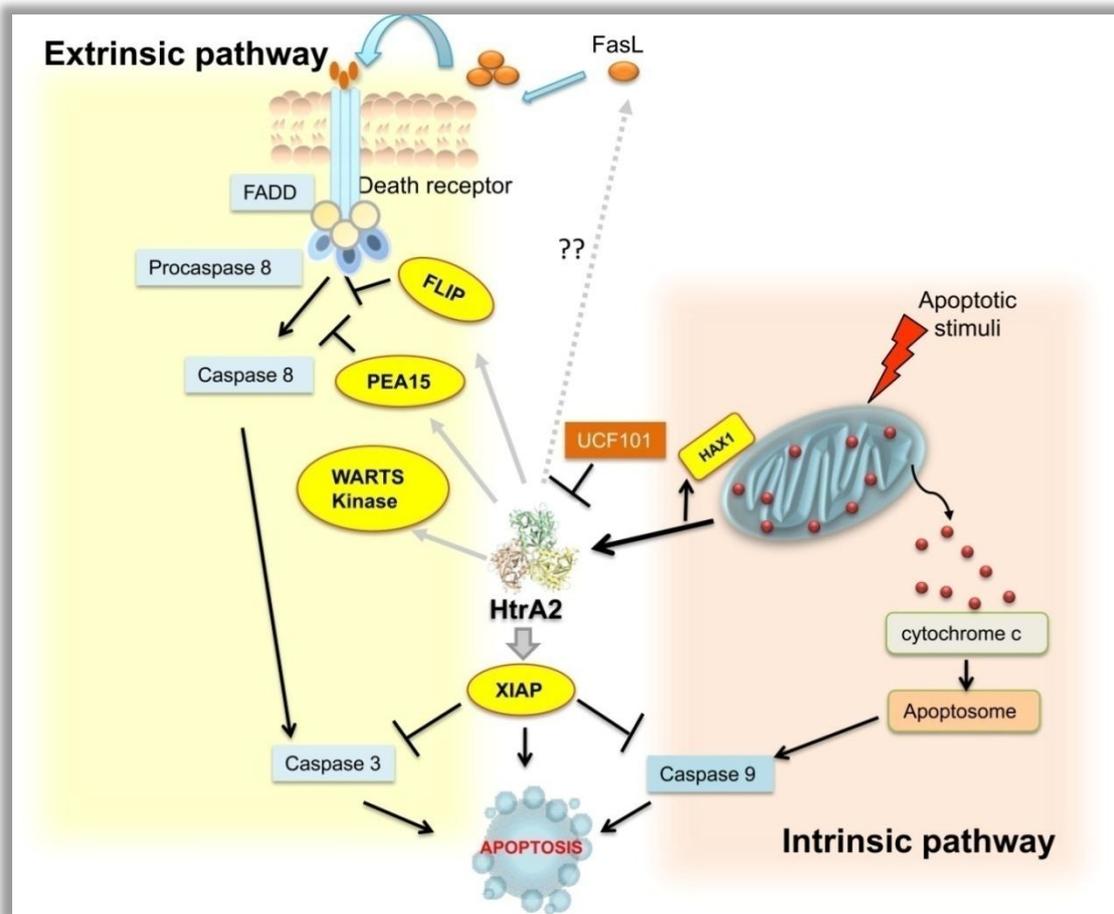


Figure 2.8 HtrA2 mediated apoptotic pathway. Upon apoptotic stimuli, mature HtrA2 is released from the mitochondrial intermembrane space into the cytosol. HtrA2 interacts with anti-apoptotic protein, XIAP thus relieving its inhibitory action on caspases, thereby facilitating caspase-dependent apoptosis. Homotrimeric HtrA2 via its PDZ domain interacts with a trimeric assembly of Fas, thus initiating death-receptor pathway of apoptosis. HtrA2 through its serine protease activity also induces caspase-independent cell death by binding and cleaving anti-apoptotic proteins FLIP, PEA-15 and HAX1, WARTS kinase, etc.

2.3 HtrA2 - AN OVERVIEW

HtrA2 has been described as a stress-induced serine protease in mammalian cells with multiple functions. It is predominantly localized in the mitochondria [88, 89]; the proform, possessing the transmembrane region is anchored in the inner membrane while the proteolytic and PDZ domains are exposed to the intermembrane space (IMS) [90]. Mature form of HtrA2 (lacking the TM region) largely resides in the IMS as a soluble protein. Under physiological conditions, HtrA2 serves as a protein quality control factor in mitochondria and promotes cell survival. Loss of HtrA2 causes accumulation of unfolded proteins in mitochondria,

dysfunction of the mitochondrial respiration, and generation of reactive oxygen species thus resulting in loss of mitochondrial competence [91-93]. However, in response to apoptotic stimuli such as cellular stresses, UV exposure [13], oxidative stress [94], etc., it transforms from a protector into a pro-apoptotic factor. It can induce cell death in both caspase-independent and caspase-dependent manner (**Fig. 2.8**). It accomplishes the former by means of neutralizing anti-apoptotic proteins such as Inhibitor of Apoptosis Proteins (IAPs) and the latter by serine protease activity.

2.3.1 Caspase-dependent mechanism

HtrA2 was first identified as an Inhibitor of Apoptosis binding protein due to its IAP recognizing Reaper-like motif (AVPS) similar to human Smac/DIABLO and *Drosophila* death proteins Reaper, Grim, Hid and Sickie [2, 12, 95]. It is synthesized as a 458 amino acid precursor protein comprising N-terminal region, serine protease and a C-terminal PDZ domain. Upon translocation to mitochondria, HtrA2 undergoes processing to remove first 133 residues comprising transmembrane anchor and a mitochondrial localization signal. The proteolytic maturation exposes the internal tetrapeptide motif (AVPS), exclusive to this member of HtrA family. IAPs are known to antagonize cell death by acting as inhibitors of activated caspases. Several members of human IAPs, including XIAP (X-linked Inhibitor of Apoptosis Protein), cellular IAP1 (cIAP1) and cIAP2, are the endogenous inhibitors of caspase-3 -7 or -9 [96-98]. Among these, XIAP is the most potent inhibitor of caspases and apoptosis. In response to various apoptotic stimuli, mature HtrA2 is released into the cytosol [95], where it interacts with XIAP through the N-terminal tetrapeptide and relieves the inhibition of IAPs on active caspases thus propagating caspase cascade and hence cell death.

The homotrimeric HtrA2 has been reported to interact via its PDZ domain with a trimeric assembly of TNFR1 or Fas [99]. The Fas ligand-induced trimerization activates the

'death domain' present in the cytoplasmic region of each Fas monomer thereby initiating caspase-8-dependent apoptotic pathway [100]. This hypothesis is further supported by observation of increased FasL expression after myocardial ischemia/reperfusion [101].

2.3.2 Caspase-independent apoptosis

HtrA2 via its serine protease activity induces cell death, independent of molecules such as caspase-9 and Apaf-1. Once HtrA2 is released into the cytosol, the enzyme cleaves one or more cellular substrates through its serine protease activity, thereby propagating atypical caspase-independent apoptosis. Following are some of the well characterized substrates of HtrA2:

2.3.2.1 Inhibitor of Apoptosis Proteins

IAPs, originally found in baculoviruses, have been found to be conserved across the species, from insects to humans [11]. Throughout the evolution, they play a significant role in regulation of apoptosis. The structures of XIAP, cIAP1 and cIAP2 are characterized by three tandem repeats of the Baculoviral IAP repeat (BIR) domain at its N-terminus and a RING finger domain towards its C-terminus. The BIR domains of IAP are majorly responsible for inhibiting caspase activity. Several studies indicated that HtrA2 mediated caspase-independent apoptosis is mainly elicited due to proteolytic cleavage of IAPs [3, 12, 15, 88, 102], thus irreversibly triggering apoptosis as this removes the endogenous anti-apoptotic 'blocks' in place.

2.3.2.2 PED/PEA-15

The DED containing anti-apoptotic protein PED/PEA-15 is also reported to be a substrate of recombinant HtrA2 *in vitro* and the presence of HtrA2 inhibitor Ucf-101 prevented PED/PEA-15 degradation in UV-irradiated 293T and HeLa cells [13]. Furthermore, as PED/PEA-15 interfered with XIAP binding on HtrA2 and prevented UV-induced caspase-3

activity, it was proposed that the cellular PED/ PEA-15 levels modulate the ability of HtrA2 to relieve XIAP-mediated inhibition of caspases [13]. Collectively, it suggests, in part, that the caspase independent cell death by HtrA2 is linked to the degradation of this anti-apoptotic molecule.

2.3.2.3 FLIP

Recently, it was found that similar to PED/PEA-15 the cellular Fas-associated death domain-like interleukin-1 β -converting enzyme-inhibitory protein (cFLIP) is HtrA2 substrate *in vivo*, and HtrA2 inhibitor Ucf-101 enhanced cFLIP expression in myocardial ischemia/reperfusion induced injury [103, 104]. It is hypothesized that cytosolic HtrA2 directly interact with the DED domain of FLIP, promote FLIP degradation, and activate apoptosis, however, the exact molecular mechanism ought to be further explored.

2.3.2.4 BRUCE/Apollon

BRUCE/Apollon, unlike other IAPs, can bind to procaspase-9 and inhibit its cleavage to the mature form. It also inhibits the activity of mature caspase-9 but not the effector caspase-3. A recent report showed that expression of catalytically active HtrA2 induced apoptosis in apollon-expressing cells. HtrA2 mediated two cleavage fragments of 180 and 150 kDa could be detected with the anti-apollon antibody [105]. It is proposed that cleavage of apollon causes irreversible activation of caspase cascade [106].

2.3.2.5 WARTS Kinase

WARTS (WTS)/large tumor-suppressor 1 mitotic kinase is a serine - threonine kinase playing important roles in mitotic and post-mitotic cell cycle regulation [107]. Studies have shown that WTS interacts with the PDZ domain of HtrA2 and promotes its protease activity to induce cell death. siRNA-mediated knockdown of the WTS protected HeLa cells against HtrA2-induced cellular toxicity and XIAP degradation [108]. Notably, the kinase activity of

WTS was required for its association with the PDZ domain of HtrA2 and the consequent increase in HtrA2 protease activity and apoptosis.

2.3.2.6 Others

Besides these targets, recently Pag, a potential candidate to act as an apoptotic factor through protein – protein interaction has been found to interact with HtrA2. The C-terminal region of Pag specifically interacts with the PDZ domain of HtrA2, which causes an increase of the protease activity [109]. Suzuki et al. suggested that HtrA2 indirectly regulates the translocation of Bax by cleaving Bmf, an endogenous inhibitor of Bax normally sequestered within the actin cytoskeleton-based myosin V motor complex [110]. Moreover, upon induction of apoptosis, HtrA2 is translocated to the nucleus where it cleaves p73. Proteolytically modified p73 thereby stimulates transcription of the BAX gene, whose protein product exhibits proapoptotic function [111]. Furthermore, caspase-independent proapoptotic property is manifested by its ability to cleave important cytoskeletal proteins such as actin, α -, β -tubulin, and vimentin culminating into cell destruction [10].

2.3.3 Alterations in cancer

Regulation of cell death is linked with cancer development. In cancer, activation of anti-apoptotic survival pathways and suppression of pro-apoptotic pathways mitigate apoptosis. Indeed, there are many reports of HtrA2 involvement in oncogenesis. Although HtrA2 mRNA is widely expressed in the cancer cell lines, the intensity of HtrA2 expression is quite cell-type specific. Lee et al. showed that reduction of HtrA2 expression can be correlated with advanced gastric adenocarcinomas [112]. Furthermore, a diminished expression of HtrA2 was observed in ovarian cancer and adult male germ cell tumor compared to normal ovaries or testis respectively [113]. Two independent analyses of gene expression in breast cancer indicate that HtrA2 expression was reduced with increasing tumor staging [114, 115].

In contrast, HtrA2 was found to be highly expressed in chronic prostatitis patients [116]. In these cases, increased HtrA2-induced apoptosis severely affected male fertility. Collectively these results highlight the important role of HtrA2 in cancer progression, and provide a rationale of targeting this HtrA family member for therapeutic intervention. However, before using HtrA2 as a therapeutic target, its mechanism and mode of activation should be fully understood.

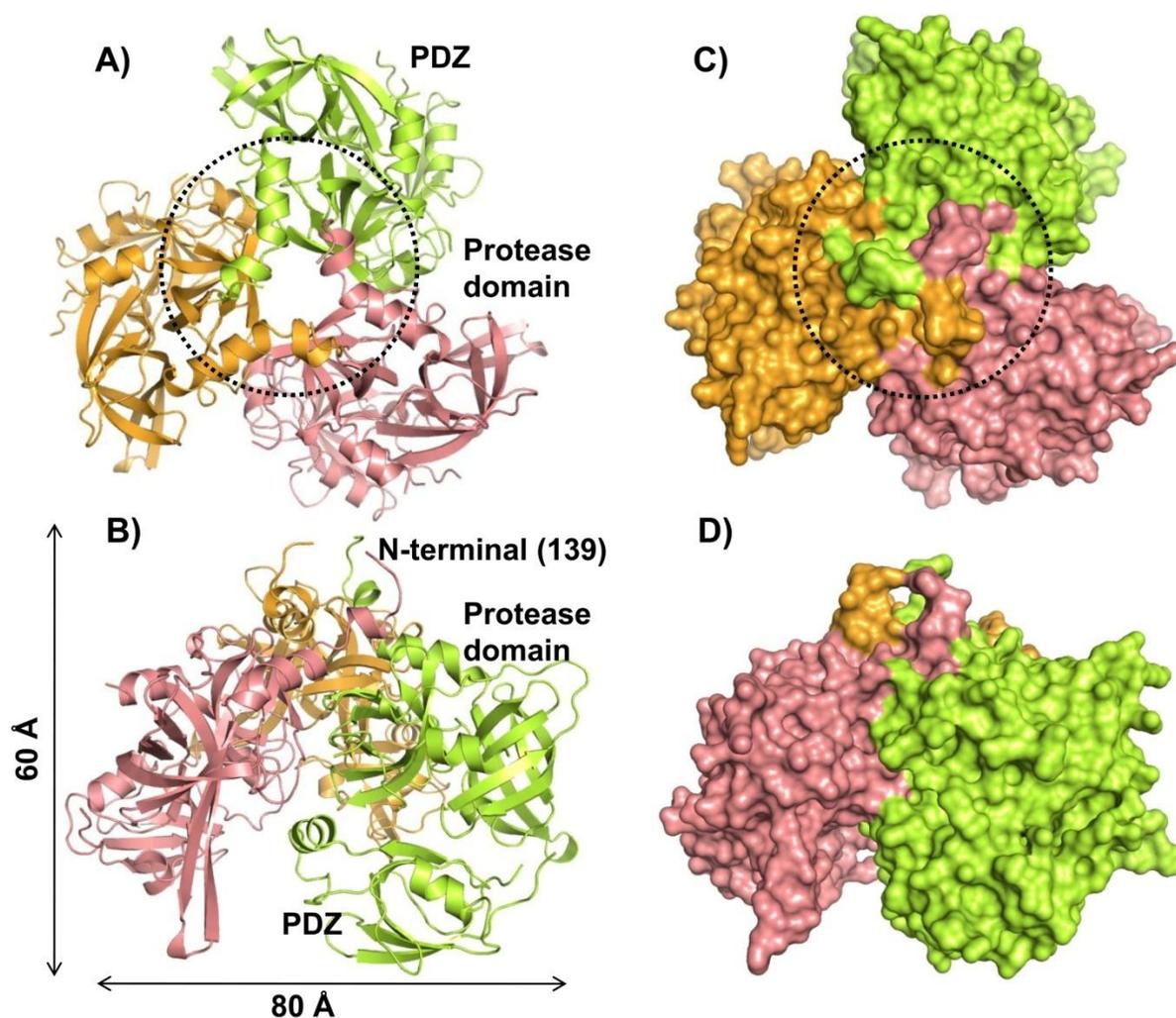


Figure 2.9 Crystal structure of HtrA2 (PDB id. 1LCY). A and B) Ribbon diagram of HtrA2 trimer. Each monomer is represented in different colors. The HtrA2 trimer is viewed either along (A) or perpendicular to (B) the three-fold symmetry axis. Trimerization is mediated exclusively by the serine protease domain. The N-terminal IAP-binding tetrapeptide motif is located at the top of the pyramid, and the PDZ domain is at the base. C and D) Surface representation of HtrA2 structure. All the images were generated using PyMol.

2.3.4 Structural details

Crystal structure of mature HtrA2 (133 - 458 residues) in a substrate unbound form was solved at 2.1 Å [14], which provides a broad overview of the global structural organization of the inactive protease. It has trimeric pyramidal architecture with the short N-terminal region at the top and PDZ domains residing at the base of the pyramid. Trimerization is mediated through extensive intermolecular hydrophobic and van der Waals interactions involving aromatic residues primarily from the N-terminal region. The protease domain is of the chymotrypsin type and is composed of two six-stranded β -barrels; the active-site consisting of the amino acid triad His198-Asp228-Ser306 is located at the interface of the two perpendicularly arranged β -barrels. Apart from the β -strands, the domain contains several regulatory loops that are named according to the chymotrypsin nomenclature, LA (residues 37-41), L1 (169-173), L2 (190-196), L3 (residues 142-162), and LD (residues 126-140) which are important for proteolytic activity and its regulation [9]. The active-site pocket surrounded by the regulatory loops is buried in the hydrophobic core of the protease domain and is arranged 25 Å above the base of the pyramid. This arrangement along with its trimeric architecture restricts entry of substrate molecules to the active site thus leading to low protease activity.

PDZ domains are well characterized protein-protein interaction modules known to recognize specific hydrophobic residues in the C-termini of binding partners. They have a canonical binding site comprising highly conserved 'G- Φ -G- Φ motif', where Φ denotes hydrophobic residues. The first Gly residue is highly variable among PDZ domains while second and fourth residues are hydrophobic (Val, Ile, Leu, or Phe) [117]. This recognition sequence is represented by YIGV in HtrA2, which is highly buried in the intimate interface between the PDZ and the protease domains [14]. The PDZ domain packs against the protease domain through van der Waals contacts, and the hydrophobic residues on strands β 11 and

β 12 of the protease domain interact with the hydrophobic residues from strand β 14 and helix α 5 of the PDZ domain. Therefore, the PDZ binding groove remains unavailable for interaction with other proteins in this 'closed' conformation.

2.3.5 Mechanism of activation

Based upon the substrate unbound structure which represents the resting form of HtrA2, Li et al. developed a model of HtrA2 activation. According to this model, in the basal state, the PDZ domains keep the protease activity of HtrA2 in check. Substrate or ligand binding at 'YIGV' groove induces a huge conformational change at the PDZ-protease interface which removes the inhibitory effect of PDZ from the active-site [14]. This structural rearrangement leads to significant increase in activity, emphasizing intramolecular PDZ-protease crosstalk to be pivotal in HtrA2 activation. Recently our group revisited the model for HtrA2 activation and highlighted the importance of PDZ domain in not only interface dynamics and initial substrate binding but also in intermolecular cross-talk for efficient substrate hydrolysis [118, 119].

To date, wide repertoire of proteins binding to the C-terminal PDZ domain have been found to stimulate the protease activity. Gupta et al. (2004) demonstrated that a peptide corresponding to the cytoplasmic C-terminal tail of presenilin-1 increased the proteolytic activity of HtrA2 towards generic serine protease substrate β -casein [120]. Furthermore, it has been shown that peptides binding to the PDZ domain bring about increase in HtrA2 activity. Ligand specificity at PDZ domain was determined to characterize its binding properties using the peptide libraries fused to the C- or N- terminus of a phage coat protein. Series of peptides binding to the isolated PDZ domain were selected and it was found that the peptide GQYYFV (termed PDZ_{opt}), which binds efficiently to PDZ, was able to stimulate

HtrA2 activity measured with synthetic substrate peptide [121]. Thus, engagement of binding partners with the PDZ domains results in opening up of access to the catalytic site.

Although PDZ acts as a regulatory domain in all the members of HtrA family, uniqueness of HtrA2 is manifested by its ability to bind subset of proteins, such as IAPs, through its N-terminus and subsequently cleave them [15, 95, 122]. This phenomenon emphasizes involvement of multiple modes of HtrA2 activation and regulation, involving PDZ as well as other regions of the protein. In agreement with this hypothesis, previous studies suggested that interaction with IAPs is not the only way that HtrA2 induces apoptosis. The serine protease activity of exogenously expressed HtrA2 when localized in the cytosol is sufficient to cause caspase-independent cell death [12, 95]. This therefore highlights the possibility that the protease activity of HtrA2 plays a significant, conceivably even dominant, role in its regulation of cell death. However, the role of PDZ domain in regulating its protease activity although highlighted by many research groups does not hold universally true such as in the case of N-terminal IAP binding and cleavage suggesting either the mechanism of substrate binding and recognition by HtrA2 is far more complicated. Therefore, it is important to understand the mode and mechanism of HtrA2 activation so as to develop a model for its substrate-binding and cleavage. A clear understanding of the structural determinants of HtrA2-mediated substrate recognition and cleavage will help define ways of regulating its functions for implicating in therapeutic strategies.

AIMS AND OBJECTIVES

In the present study, our goal is to characterize structural and functional properties, and dissect the mechanism of these unique proapoptotic proteins involved in alternate pathways of apoptosis.

AIM-I

High risk HPV18 E2 protein has been found to induce apoptosis via direct interaction with procaspase-8 through a unique pathway that bypasses homotypic interactions between death domains of proteins in death inducing signaling complex. Our aim is to delineate the mechanism of interaction between HPV18 E2 and procaspase-8 as well as to understand how it aids in E2-induced apoptosis.

Objectives:

- To identify minimal binding regions and critical residues involved in the E2-procaspase-8 interaction
- To design interface mutants and study the effect on procaspase activation and apoptosis
- Delineate the mechanism of procaspase-8 activation and hence cell death

AIM-II

The mechanism of HtrA2 mediated proteolysis and hence its role in caspase-independent apoptosis is poorly understood due to the limited number of identified substrates as well as its structural complexity and plasticity. Although C-terminal PDZ acts as a regulatory domain in all the members of this family, HtrA2 exhibits an additional level of functional modulation

through is unique N-terminus that is evident from IAP binding and cleavage. Therefore, our goal is to understand the structural determinants of protease activity of HtrA2 in presence of N-terminal ligand and to address this complex behavior toward defining its global mode of regulation and protease activity.

Objectives:

- To identify protein-protein interaction sites of HtrA2 and IAP and determine the role of IAPs in activation of HtrA2 protease activity
- To compare specificity & catalytic efficiency of HtrA2 upon IAP protein/peptide and PDZ peptide binding
- To delineate the global conformational changes and active-site orientation following IAP protein/peptide binding

CHAPTER-3

Materials and Methods

3.1 CLONING, EXPRESSION AND PURIFICATION OF RECOMBINANT PROTEINS

3.1.1 Materials

3.1.1.1 Bacterial strains

Cloning host used is *E. coli* DH5 α ultra competent cells, expression host included *E. coli* strain BL21 (DE3) and Rosetta (DE3) (Novagen).

3.1.1.2 Bacterial cell culture medium

- **2 % Luria Broth (LB) medium**

20 g Luria Broth powder (Himedia) was dissolved in deionized water and the final volume made to 1000 ml. The medium was sterilized by autoclaving at 15 lbs for 20 min.

- **2 % Luria Broth agar medium**

2.0 % Luria Broth medium, 1.5 % Agar (Himedia RM 301): 15 g of agar was added to the LB medium and the mixture was autoclaved at 15 lbs for 20 min.

- **Antibiotics**

Ampicillin sodium salt (Sigma A-9518):

Stock: A solution of 100 mg/ml was made in deionized water and filtered through a 0.22 μ m filter. LB media containing 100 μ g/ml ampicillin was used for growing cultures with plasmid harboring ampicillin resistance gene.

Kanamycin sulfate (Sigma K4378):

Stock: A solution of 50 mg/ml was made in deionized water and filtered through a 0.22 μ m filter. 34 μ g/ml kanamycin was used in LB media.

Chloramphenicol (MP Biomedicals):

A solution of 34 mg/ml was made in deionized water and filtered through a 0.22 μ m filter. LB media containing 34 μ g/ml chloramphenicol was used for growing Rosetta 2(DE3).

3.1.1.3 Reagents for preparation of E. coli DH5- α ultra competent cells

- **SOB medium**

2 % bacto-tryptone (Himedia RM 014), 0.5 % yeast extract (Himedia RM 027), 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, pH 6.7 to 7.0

To make SOB medium, 6 g bacto-tryptone, 1.5 g yeast extract, 175.32 mg NaCl, 55.91 mg KCl were dissolved in 297 ml deionized water and autoclaved. After autoclaving, 3 ml of 2 M Mg solution was added prior to use.

- **2 M Mg solution**

1 M MgSO₄·7H₂O (6.15 g in 25 ml deionized water),

1 M MgCl₂·6H₂O (5.075 g in 25 ml deionized water)

25 ml of 1 M MgSO₄·7H₂O and 25 ml MgCl₂·6H₂O were mixed to make 2 M Mg solution, and it was filtered through 0.22 μ m filter.

- **TB buffer**

10 mM PIPES, 15 mM CaCl₂, 250 mM KCl, 55 mM MnCl₂

393 mg PIPES, 286.65 mg CaCl₂, and 2.42 g KCl were dissolved in 120 ml deionized water and pH was adjusted to 6.7 with 5 N NaOH. After adjusting pH, 1.42 g MnCl₂ was added to the solution and volume was made up to 130 ml with deionized water. The solution was filtered through 0.22 μ m filter.

- **Dimethyl Sulfoxide (DMSO) (Sigma D-8779)**

3.1.1.4 Plasmids

CloneJET PCR cloning kit (Thermo Scientific), pMALc5-E (NEB), pGEX2T (GE Healthcare), pEGFPn1 and pEGFPc1 (Clontech), pmCherry (a kind gift from Dr. Bhattacharya lab, ACTREC), pcDNA3-HA (Life Technologies)

Polymerase Chain Reaction (PCR)

High fidelity DNA polymerase-Pfu Turbo (Chem-agilent), 10X Pfu Turbo buffer, 25 mM MgCl₂, dNTPs mix (2 mM) from Stratagene

3.1.1.5 Primers

PCR primers were commercially synthesized from Sigma Genosys. The primers were designed using DNA star software.

For the complete list of primers please find **Table 3.1**. The lyophilized primers were reconstituted in autoclaved deionized water to obtain final concentration of 1000 ng/μl. A working dilution of 125ng/μl was prepared for further experiments.

Table 3.1 List of oligos for site-directed mutagenesis and sub-cloning of different gene of interest.

HPV18 E2 mutants	
Q35A, Q38A	CAG CCA AAT AGC GTA TTG GGC ACT AAT ACG T
R41A	GCA ACT AAT AGC TTG GGA AAA TG
W42A	CAA CTA ATA CGT GCG GAA AAT GCA
R41A, W42A	CAACTAATAGCTGCGGAAAATGC
F48A	GCA ATA TTC GCT GCA GCA AGG
E51A, H53A	GCA GCA AGG GCA GCT GGC ATA CAG
Procaspase-8 mutants	
L42A	CAA GGA TGC CGC GAT GTT ATT CC
F45A	CCT TGA TGT TAG CCC AGA GAC TC
L42A, L45A	CAA GGA TGC CGC GAT GTT AGC CCA GAG ACT C
F122A	GTC TTT TAA GTT TGC TTT GCA AGA GG
F122Y	GAGGTCTTTTAAGTATCTTTTGCAAGAGG
L123A	GTC TTT TAA GTT TGC TTT GCA AGA GG
L123A, Q125A	CTT TTA AGT TTG CTT TGG CAG AGG AAA TCT C
F122A, L123A	GTC TTT TAA GGC TGC TTT GCA AGA GG
Y8G	CAG AAA TCT TGG TGA TAT TGG GGA AC
R5A, Y8A	GGA CTT CAG CGC AAA TCT TGC TGA TAT TGG
S119A, E126A	CAG AAT TGA GGG CTT TTA AGT TTC TTT TGC AAG CGG AAA TCT C

D158A	GGA AAG TTG GCC ATC CTG AAA
R162A	CAT CCT GAA AGC TGT CTG TGC C
Q166A	GTCTGTGCCGCAATCAACAAGAG
K161A, R162A	TGG ACA TCC TGG CAG CAG TCT GTG CC
Q166A, N168A	GTC TGT GCC GCA ATC GCC AAG AGC CTG CTG
C360A	TTA TTC AGG CTG CTC AGG GGG ATA ACT
FADD mutants	
V6A	CGT TCC TGG CGC TGC TGC AC
H9A	GTG CTG CTG GCC TCG GTG TCG
H9G	CTG GTG CTG CTG GGC TCG GTG TCG TC
T21A	GCG AGC TGG CCG AGC TCA AG
F25Y	GAG CTC AAG TAC CTA TGC CTC GG
L28A	GCT CAA GTT CCT ATG CGC CGG GCG CGT GG
K33R	GGC GCG TGG GCG AGC GCA AGC TGG AG
R34A	GCA AGC GCA GGC TGG AGC GCG TG
K35R	GCA AGC TGG CGC GCG TGC AG
E37A	GTG GGC AAG GCC AAG CTG G
R38A	CGC AAG CTG GAG GCC GTG CAG AGC GGC C
L43A	CAG AGC GGC GAT GAC CTC TTC
L43D	CAG AGC GGC GCA GAC CTC TTC
D44A	CAG AGC GGC CTA GCC CTC TTC TCC ATG C
E51A	CCA TGC TGC TGG CGC AGA ACG ACC TGG AGC
L62A	GCA CAC CGA GGC CCT GCG TGA GCT G
L63A	CAC ACC GAG CTC GCG CGT GAG CTG CTC
A68F	CGT GAG CTG CTC TTC TCC CTG CGG CGC
HtrA2 mutants	
G230A	GTC CCA GCG GCG CTA CAT TGC GGT GAT GAT G
AVPS	GAG ATA TAC ATA TGG TCG CTA GCC CGC CG
XIAP mutants	
E219R, H223V	GTG CCT GGT CAC GAC ACA GGC GAG TCT TTC CTA ATT G
Q319R, W323V	CTT GGG AAC GAC ATG CTA AAG TGT ATC CAG GGT GC
Procaspase-8 DED-B deletion constructs	
Helix-1	GGGAGAATCTTTATTTTCAGGGCGGATTCATGGAAGTG AGCAGATCAGAATTGAG
Helix-2	GAGAATCTTTATTTTCAGGGCGGATCCATGGAAATCTC CAAATGCAAACCTGGATG
Helix-3	GAGAATCTTTATTTTCAGGGCGGATCCATGGATGACAT GAACCTGCTGGATATTTTCA
Helix-4	CAAATGCAAACCTGGATGATGACATGTGAGAATTCGAA

	GCTTGGCACTGGC
Helix-5	GTCATCCTGGGAGAAGGAAAGTTGTGAGAATTCGAAG CTTGGCACTGG
Helix-6	GAGTCTGTGCCCAAATCAACAAGTGAGAATTCGAAGCT TGGCACTGG

For all these mutants, the reverse primer is the complementary sequence.

Mammalian procaspase-8 constructs	
DED-A (F)	CGGGAATTCATGGACTTCAGCAGAAATC
DED-A (R)	CGGTGGATCCCGCCTGTAGGCAGAAATTTGAGC
DED-B (F)	CTCAAGCTTCGAATTCATGGTCATGGTCATGCTCTAT CAGATTC
DED-B (R)	CGGTGGATCCCGGATTGTCATTACCCACAC
FADD (F)	GAATCCATGGACCCGTTCTGGTGCTG
FADD (R)	GGA TCC TCA GGA CGC TTC GGA GGT AGA TGC G
Caspase-8 full length (R)	GGATCCTCAATCAGAAGGGAAGACAAGTTTT
Bacterial procaspase-8 constructs	
DED-AB (F)	CGCGGATCCATGGACTTCAGCAGAAATC
DED-AB (R)	CCGGAATTCTCATTCTTCATAGTCGTTGA
DED-A (R)	CCGGAATTCTCAGTTTAGGTAGGTAATCAGC
DED-B (F)	GGATCCATGGCCTACAGGGTCATG
DED-B (R)	CTGGATCCCGCTCTCTTTGCTGAATTC
Bacterial HPV18 E2 constructs	
HPV18 E2 (1-201)(F)	CGCGGATCCATGGAGACACCGAAGGAAAC
HPV18 E2 (1-201)(R)	CCGGAATTCTCAATTACAATCAATTACATTATTTCC
HPV18 E2 full length (F)	GGATCCATGGAGACACCGAAGG
HPV18 E2 full length (R)	GAATTCTTACATTGTCATGTATCCCACCA
XIAP	
BIR2 (124-240) (F)	GAATTCCATATGAGAGATCATTTTGCCTTA
BIR2 (124-240) (R)	GGATCCCTCGAGTTCACTTCGAATATTAAG
BIR3 (241-356) (F)	GAATTCCATATGTCTGATGCTGTGAGTTCT
BIR3 (241-356) (R)	GGATCCCTCGAGAGTTCTTACCAGACACT
Sequencing oligos	
CMV promoter	CGCAAATGGGCGGTAGGCGTG
T7 promoter	TAATACGACTCACTATAGGG
T7 terminal	GCTAGTTATTGCTCAGCGG
MBP reverse	GGCCCAGTCTTTTCGACTGAG
MBP forward	GATGAAGCCCTGAAAGACGCGCAG

3.1.1.6 Restriction Digestion and Ligation

Fast Digest DpnI, BamHI, EcoRI, NcoI, NdeI, XhoI and 10X Fast digest buffer supplied along with the restriction enzymes (Fermentas)

T4 DNA ligase and ligase buffer from NEB

3.1.1.7 Agarose Gel Electrophoresis

- **10 X TBE buffer**

0.89 M Tris base, 0.89 M Boric acid, 0.02 M EDTA:

108 gm Tris base, 55 g boric acid and 7.44 g EDTA were dissolved in deionized water to a final volume of 1 litre.

- **Ethidium bromide (Sigma E-8751)** – Stock of 10 mg/ml in deionized water
- **Agarose (Sigma A-6013)** – 1% Agarose in 1X TBE buffer
- **1kb DNA ladder** – 0.5 µg/µl (Fermentas SM0311)
- **Loading dye (6X)** – 0.25 % Bromophenol blue, 0.25 % Xylene cyanol, 30 % Glycerol

3.1.1.8 Reagents for expression and purification of recombinant GST fusion proteins

- **1M Isopropyl-β-D-thiogalactopyranoside (Sigma I-5502)**

1.19 g Isopropyl-β-D-thiogalactopyranoside (IPTG) was dissolved in 5 ml deionized water, filter sterilized and stored at -20 °C.

- **Lysis buffer for bacterial cells**
- **Protease inhibitor cocktail (Sigma P8340)**

Resins – Ni-IDA (BIOTEX), GST-sepharose (Novagen), Amylose (NEB), Pre-packed gel filtration column: Superdex 200 and 75 – 16/600 and 10/300 from GE Healthcare

3.1.2 Methods

3.1.2.1 Preparation of *E.coli* DH5- α ultra competent cells

The procedure followed was adopted from *Inoue et al* with some modifications [123]. *E.coli* DH-5 α cells were streaked from the glycerol stock on LB agar plate and were incubated overnight at 37 °C. A single colony of cells were picked up and grown in 250 ml of SOB medium at 18 °C until OD₆₀₀ reached to 0.6. The cells were then immediately kept on ice for 10 min and harvested by centrifugation at 6,000 rpm for 10 min at 4 °C using a SS-34 rotor in a Sorvall RC5C centrifuge. After centrifugation, cells were re-suspended in 80 ml of pre-chilled TB buffer and kept on ice for 10 min. The cell suspension was then centrifuged at 5000 rpm for 10 min at 4 °C. Post centrifugation, cells were re-suspended gently in 20 ml ice-cold TB buffer and 1.4 ml DMSO was added to attain final concentration of 7% (v/v). Cells were kept on ice for 10 min and aliquots of 100 μ l each were made and snap frozen in liquid nitrogen. The ultra competent cells were then stored at -80 °C.

3.1.2.2 Strategy for cloning of genes

There are different kinds of gene cloning methods such as homology based cloning, ligation-independent cloning and restriction enzyme based cloning. In the present study, we employed restriction enzyme based cloning which is described in detail in the following section:

Restriction enzyme based cloning

Here, both expression vector and target genes are subjected to digestion with the same set of sticky ends generating restriction enzymes that enables insertion of target gene into a plasmid backbone.

a) Preparation of expression vector: The target genes were cloned into different bacterial and mammalian expression vectors to perform various *in vitro* and *ex vivo* experiments,

respectively. The plasmid DNA was extracted and purified as per the manufacturer's protocol (GeneElute Plasmid Miniprep Kit from Sigma). The isolated plasmid vector was digested with the respective restriction endonucleases and the digestion was carried out at 37 °C for 1 h. The digested plasmid was resolved on 1% agarose gel followed by purification using Gel Extraction Kit (Sigma) according to manufacturer's protocol.

b) Preparation of target gene: The gene of interest was PCR amplified using gene specific forward and reverse primers. The primers were designed using Primer-X software online tool. The parameters taken under consideration for designing primers included GC content: 40-60%, T_m: 50-65 °C, presence of 3'-GC clamp. Detailed list of oligos synthesized for amplifying specific gene of interest is given in **Table 3.1**.

Table 3.2 A typical PCR reaction composition.

Reagents	Volume (µl)
PCR Water	15.5
10X Pfu Turbo buffer	2
50 mM MgCl ₂	0.5
Template DNA (80ng)	0.5
10mM dNTPs	0.5
(125 ng/µl) Forward primer	0.5
(125 ng/µl) Reverse primer	0.5
Pfu Turbo	0.5
Total volume	20

The typical program set for PCR was as follows:

1. Denaturation – 95 °C for 5 min
2. Denaturation – 95 °C for 30 sec
3. Annealing – 53 °C for 50 sec
4. Extension – 72 °C for min (depending upon rate of amplification and no. of bases)
5. Repeat steps from 2 to 4 for 25 cycles

6. Final Extension – 72 °C for 10 min

7. Hold at 4°C forever

The amplified PCR product was purified using PCR clean-up kit from Sigma by following the manufacturer's protocol. The purified PCR product was subjected to restriction digestion with the respective endonucleases used for preparing the vector backbone. The digested PCR product was further purified by gel extraction technique.

c) **Ligation:** The digested vector and PCR product was mixed in a molar ratio of 1:3 or 1:6, and then ligated using T4 DNA ligase according to the manufacturer's protocol. The ligation reaction was carried out at 22 °C for 2 h. The ligation mixture was then transformed into ultra-competent DH5α cells.

The plasmid DNA can be incorporated into the competent cells in the process of bacterial transformation. The ultra competent *E.coli* DH5α cells were taken out from -80 °C and thawed on ice. 5 µl of ligation mixture was added to the cells and the suspension was incubated for 30 min. Cells were given heat shock at 42 °C for 90 sec and immediately kept on ice for 2-5 min. After heat shock, 750 µl of sterile LB media was added and the mixture was incubated at 37 °C for 30 min on shaker. Post incubation, cells were pelleted down by centrifugation at 5,000 rpm for 3 min at RT. The supernatant was discarded and fresh 100 µl of LB was added in the tube. Cells were re-suspended and spreaded on LB-agar containing appropriate antibiotic for selection. The plate was incubated at 37 °C for 16-18 h.

d) **Screening of potential clones:** The colonies were screened by isolating the plasmid DNA and subjecting it for restriction digestion using the respective enzymes. The digested product was resolved on 1% agarose gel to check the release of target gene from the vector backbone.

The positive clones are further confirmed by DNA sequencing using specific sequencing oligos listed in **Table 3.1**.

3.1.2.3 Site-Directed Mutagenesis (SDM)

The method relies on amplification of entire plasmid using mutagenic primers and DNA polymerase lacking primer displacement activity. SDM involves point mutation at a particular position in the gene of interest and comprises following three major steps.

a) Overlapping PCR: It requires two site-specific primers with the incorporated mutation and should be self complementary. The desired mutation should preferably be at the centre of the primer sequence of about 25-45 bases in length. The PCR reaction can be setup according to the reaction conditions described in **Table 3.2**.

b) Dpn I digestion: DpnI is an endonuclease which specifically targets the methylated DNA strands. Most of the plasmid DNA synthesized in *E. coli* is dam methylated and is therefore susceptible to DpnI digestion. One Unit of DpnI is added to 30 µl of reaction mixture and incubated for about 1 h at 37 °C.

c) Transformation of SDM PCR product and sequencing: The Dpn I digested product is transformed into competent *E.coli* DH5α cells and selected on LB agar plate containing appropriate antibiotic for selection. Single colony is inoculated in 10 ml of LB broth supplemented with the respective antibiotic and grown overnight in a shaker incubator at 37 °C. The plasmid DNA isolated from these colonies was sequenced to confirm the presence of mutation.

3.1.2.4 Bacterial Protein Expression

a) Selection of expression host: Different bacterial strains having specific or unique properties can be utilized for protein expression. Primarily, the choice of bacterial host strain depends on level of gene expression. For routine purposes, *E. coli* BL21 (DE3) strain, a lysogen of lambda-prophage (DE3), is used because it lacks the Lon and OmpT proteases. It carries 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. *E. coli* Rosetta 2(DE3) is a derivative of BL21 (DE3) which can be used for the expression of eukaryotic proteins having rare codons. The strain is equipped with additional tRNAs that can recognize codons such as AUA, AGG, AGA, CUA, CCC, and GGA encoded on a compatible chloramphenicol-resistant plasmid call pRARE. Also, a modified version of BL21 (DE3) called BL21 pLysS can be used for expression of toxic genes. The strain contains a chloramphenicol resistant pLysS plasmid which encodes for T7 lysozyme to prevent basal or leaky expression.

b) Preparation of seed culture: A single transformed colony of the expression host is inoculated in 20 ml of sterile LB broth containing suitable antibiotic. The culture is allowed to grow overnight at 37 °C with continuous agitation at 200 rpm.

c) Scale up and induction: The scaling up of protein expression is carried out by inoculating the starter culture in 1000 ml LB broth in a ratio of 1:100. The culture is grown at 37 °C till the OD₆₀₀ reaches between 0.6-0.8. Depending on the expression vector, the protein is induced by adding IPTG (at a concentration of 0.4 mM) or arabinose (2 mg/ml). The culture flasks are incubated at 18 °C for 16-18 h under constant shaking of 200 rpm. Cells are then

harvested by centrifugation at 5,000 rpm for 10 min at 4 °C. The bacterial pellet is stored at -80 °C until further use.

3.1.2.5 Recombinant Protein Purification

Protein purification procedure utilizes differences in physical properties of the bio-molecule such as protein size, charge, ligand specificity and biological activity. Commonly, purification steps contain one or more chromatographic separation techniques. These techniques that exploit different physiochemical properties are listed below:

- Affinity chromatography – Based on ligand specificity or bio recognition
- Ion exchange chromatography – Dependent upon charge of the protein
- Gel filtration chromatography – Separation on the basis of molecular size/mass
- Hydrophobic interaction chromatography – Relies on overall hydrophobicity

In the present study, we followed two step purification using affinity chromatography followed by gel filtration.

a) Affinity chromatography: Affinity chromatography separates proteins on the basis of a reversible interaction between a protein and a specific ligand coupled to a chromatography matrix. It provides highest specificity and selectivity for purification of bio-molecules. The interactions between ligand and target protein can be a result of electrostatic or hydrophobic interactions, van der Waals' forces and/or hydrogen bonding. To elute the target molecule from the affinity medium the interaction is reversed, either specifically using a competitive ligand, or non-specifically, by changing the pH, ionic strength or polarity. In a single step, affinity purification offers immense time-saving over the less selective multi-step procedures. The affinity matrix, used in our experiments is listed below:

- Glutathione Sepharose – glutathione-S-transferase for GST fusion proteins

- Metal ions (Ni-IDA) – Poly (His) fusion proteins, native proteins with histidine
- Amylose resin – Proteins fused to maltose-binding protein (MBP)

b) Gel filtration chromatography: Gel filtration also called size-exclusion chromatography is the separation technique based on the molecular size and volume of the components. The molecule with larger molecular hydrodynamic volume migrates faster compared to the smaller one. Separation is achieved by the differential exclusion of the sample molecules from the pores of the packing material. The principle feature of this technique is its gentle non-adsorptive interaction with the sample, enabling high retention of bimolecular activity and achieving homogeneous purified proteins. It can be used for varied applications including protein purification, buffer exchange and desalting.

3.1.2.6 General protocol for protein purification

The cell pellet is re-suspended with pre-chilled lysis buffer (10 ml for 1 litre of pellet). The cells are lysed by ultra-sonication under ice-cold condition at 50 pulse rate and 50 power for 10 cycles of 1 min each. The lysed homogenous suspension is centrifuged at 18,000 rpm for 20 min. For soluble proteins, collect the supernatant and discard the cell debris. All the proteins analyzed or characterized in the present study was purified using two stages of purification. The pre-cleared lysate is carefully added on the equilibrated affinity resins and incubated at room temperature (RT) for an hour. After binding, the resin is extensively washed with the lysis buffer to remove non-specific or bacterial proteins. Depending upon the tag, the protein is eluted using a gradient of imidazole (25-500 mM) for His-tag fusion protein or maltose (2-20 mM) for MBP-fused proteins. The purity of proteins is checked on 12-15% SDS-PAGE based on the molecular weight of the proteins being analyzed. Further, the desired fractions are concentrated by ultra-filtration method using Amicon Ultra

centrifugal filter devices of 3-30 kDa molecular weight cut-off. The second step of purification involves gel filtration chromatography using Superdex 200 or 75 matrixes. 1 ml of concentrated protein is injected in the pre-equilibrated column (instrument-AKTA purification system). Elution of the protein is monitored by recording the absorbance at wavelength 280 nm. The purity of the eluted fractions is checked on 12-15% SDS-PAGE depending on the molecular weight of the proteins being analyzed. The detailed purification protocols for different proteins have been mentioned in the respective chapters.

3.2 ANIMAL CELL CULTURE

3.2.1 Routine maintenance of cell lines

3.2.1.1 Materials

- **Medium**

Powdered Dulbecco's modified Eagle's medium (DMEM) from GIBCO was dissolved in 1 litre of deionized water (3.7 g of sodium bicarbonate per litre was added and the pH of was adjusted to 7.4). The medium was filtered using millipore assembly – 0.22 µm membrane filter (Whatman). One ml of the filtered medium was added to the sterility test medium and incubated at room temperature for 6 days under observation. Sterility test medium (14.9 g of Fluid-thioglycolate was dissolved in ~250 ml of water. The volume was made up to 500 ml in measuring flask and boiled. After aliquoting 6 ml of the medium in glass tubes a pinch of CaCO₃ was added to each tube and autoclaved.)

- **Phosphate Buffered Saline (PBS)** - 150 mM NaCl, 2 mM KCl, 8 mM Na₂HPO₄, 1 mM KH₂PO₄)

- **0.4% Trypan Blue solution** - Prepared in 0.81% NaCl and 0.06% KHPO₄, dibasic

- **Trypsin-EDTA** (0.025% Trypsin, 0.2 mM EDTA)

- **Complete medium**

DMEM with 10% Fetal bovine serum and 1% antibiotic solution (Amphotericin B 20 µg/ml, Penicillin 2500 Units/ml, Streptomycin 800 µg/ml)

- **Freezing mixture** (90% Fetal bovine serum, 10% DMSO)
- **Cell lines** – Human Embryonic Kidney (HEK) 293, HeLa

3.2.1.2 Protocols

a) Revival of cells: Revival of frozen cell cultures was carried out by thawing the freezing vials with the cells into a glass beaker containing water at 37 °C. Immediately the thawed cell suspension was added drop wise into a sterile centrifuge containing 5 ml of complete medium. The cell suspension was centrifuged for 10 min at 1,000 rpm. The supernatant was discarded and the cell pellet was washed with 1X PBS twice for complete removal of freezing mixture. Finally, the cell pellet was re-suspended in 3 ml of complete medium. The cell suspension was mixed gently with a Pasteur pipette to remove any clumps and finally transferred in a culture plate or flask. The cells were incubated in a humidified CO₂ (5%) incubator at 37 °C and their growth was observed each day under inverted microscope.

b) Subculture/trypsinization and transfer of cells: Cells with 70-80% confluency were washed with 1X PBS twice and 1 ml of trypsin-EDTA was added to the culture plate. Excess trypsin-EDTA was discarded and the plate was incubated till the cells partially detached. Complete medium was added into the plate and the resulting cell suspension was mixed with pipette to make a single cell suspension. Total cell count and the percent viability were calculated by dye exclusion method using Trypan Blue dye on a haemocytometer using an inverted microscope. Appropriate amount of cells depending upon cell type were seeded in culture plates. The plates were further incubated in humidified CO₂ incubator at 37 °C.

c) Freezing and cryopreservation of cells: For freezing the cells, 70-80% confluent culture flask or dish was subjected to trypsinization. After noting the total cell count, the cell suspension was spun at 1,000 rpm for 10 min. The supernatant was discarded and the pellet was dislodged by tapping the tube gently. One ml of freezing mixture was added drop wise and mixed gently with pipette. The cells suspension was then transferred to freezing vials. The freezing vials were labelled and suspended in a cylinder with liquid nitrogen vapour and cooled gradually before plunging it into liquid nitrogen for long term storage.

3.2.2 Transient transfection

Reagents: Plasmid DNA, 0.5 M CaCl₂, 2X BBS (50 mM BES, 1.5 mM Na₂HPO₄ and 250 mM NaCl), autoclaved water, Lipofectamine 2000 (Invitrogen)

Note: Working aliquot of calcium chloride and water were stored at 4 °C for not more than a month. BBS was stored at -20 °C. About half an hour prior to use, calcium chloride and sterile water were warmed to 37 °C and BBS was thawed at RT.

Protocol: Cells at a confluency of 40-60% were transfected by calcium phosphate precipitation protocol as described [124] or with Lipofectamine 2000 as per the manufacturer's protocol. Following is the composition of calcium phosphate transfection mixture in different culture dishes-

Diameter of culture dish	DNA	Water	CaCl ₂	2X BBS	Total volume
35 mm	5 µg	45 µl	50 µl	100 µl	200 µl
60 mm	10 µg	90 µl	100 µl	200 µl	400 µl
100 mm	25 µg	225 µl	250 µl	500 µl	1000 µl

3.2.3 Live cell confocal imaging

Material: Optimum Minimum Essential Media (Opti-MEM), buffered with HEPES and sodium bicarbonate, 35 mm glass-bottomed dishes (Cell E&G, USA)

Protocol: Cells were grown in the glass-bottomed dishes and were transfected at the confluency of 40-50 % with the desired plasmids. Post 24-30 h of transfection, the complete medium was replaced with minimally fluorescent Opti-MEM. To monitor colocalization, confocal imaging was performed with Zeiss LSM 510 META equipped with 100× or 63× 1.4 NA (numerical aperture) objectives. Single- or dual-colour images were obtained using separate excitation in 12-bit format and with line averaging of two. GFP fluorescence was excited with a 488 nm argon laser and collected between 495 and 550 nm. Similarly, mCherry was excited with a 543 nm helium-neon laser and emission was captured between 650 and 760 nm. Transmitted light images were captured in the blue channel. Z-stacks were collected at intervals of 2–4 seconds and laser illumination was minimized to limit photodamage. Images acquired were further processed using LSM 510 image examiner software.

3.2.4 Preparation of whole cell lysate post transfection

Reagents: SDS lysis buffer (5 mM EGTA, 5 mM EDTA, 0.4% SDS and protease inhibitor cocktail in 25 mM Tris-HCl pH 7.2) and non-denaturing lysis buffer (20 mM Tris-HCl pH 8.0, 137 mM NaCl, 0.5% Nonidet P-40, 1 mM EDTA)

Protocol: Post 30 h of transfection, the cells were spun at 1,000 rpm for 10 min at 4 °C. Cells were suspended in SDS or denaturing lysis buffer and incubated on ice for 40 min. The cell suspension was then subjected to ultra-sonication for 1 min, followed by centrifugation at 13,000 rpm for 15 min to separate cell debris. The supernatant and pellet were loaded on SDS-PAGE followed by Western blot analysis using respective antibodies.

3.2.5 Protein estimation using Peterson method

Reagents: Working stock of BSA (1mg/ml) prepared in sterile distilled water, Copper Tartarate Carbonate (CTC) reagent: CTC reagent was prepared by adding 20% Na₂CO₃, 0.2% CuSO₄, and 0.4% Sodium potassium tartarate in the ratio of 1:1:2.

Solution A: Prepared by mixing CTC, 10% SDS, 0.8 N NaOH, and deionised water in the ratio of 1:1:1:1 and Solution B was prepared by diluting Folin Ciocalteau's Phenol Reagent with distilled water in the ratio of 1:6.

Protocol: One ml of 5 – 25 µg /ml of BSA standard tubes were prepared in duplicates along with the blank. For estimating the concentration of unknown proteins, 5 µl of the lysate was added to 995 µl of distilled water. 1 ml of solution A was added and the tubes were vortexed followed by incubation at 37 °C for 10 min in dark. To each tube add 500 µl Solution B followed by incubation at room temperature (RT) for 30 min in dark. The absorbance was read at 750 nm, and the concentration of the unknown lysate was calculated by extrapolating the standard curve.

3.2.6 SDS-PAGE

Reagents: 30% Acrylamide (29.2% acrylamide (w/v) and 0.8% bis acrylamide(w/v)), 1.5 M Tris-HCl pH 8.8, 1 M Tris-HCl pH 6.8, 10% SDS, 10% APS, TEMED and 6X SDS sample buffer (62.5 mM Tris-HCl pH 6.8, 25% Glycerol, 2% SDS, 0.5% bromophenol blue), Electrode buffer (5X): 15.1 g Tris base, 94 g glycine and 5 g of SDS were dissolved in deionized water to a final volume of 1 litre.

- **Molecular weight markers** – Fermentas page ruler prestained protein ladder-SM 0671
- **Coomassie staining solution** – 0.25% coomassie brilliant blue R 250, 45% methanol and 10% acetic acid in distilled water
- **Destainer** – 45% methanol (v/v) and 10% acetic acid (v/v) in distilled water

Protocol: The samples were separated on 6-15% SDS-PAGE depending on the molecular weight of the proteins being analyzed with 5% stacking gel.

3.2.7 Western Blotting

Reagents: 1X Transfer Buffer (190 mM glycine, 20% methanol, 0.05% SDS, 25 mM Tris base), Tris-buffered saline (TBS) – 150 mM NaCl, 10 mM Tris-HCl pH 8.0, Tris-buffered saline with 0.1% TWEEN-20 (TBST), Ponceau staining solution (0.2% ponceau stain in 5% acetic acid), Blocking buffer (3% BSA or 5% Milk in TBS), Antibodies dilutions were made in 1% BSA in TBS, ECLprime Kit (GE Healthcare), X-ray film (Kodak), Antibodies- anti-His antibody (Abcam, USA), anti-HPV18 E2 (sc-26939), anti-caspase-8 (1C12, Cell signaling), anti-FADD (sc-56093), APO 1–3 anti-Fas (clone CH11, 05-201, Millipore), anti-GFP (sc-9996), anti- β actin (AC-74, Sigma), secondary anti-HRPO mouse, rabbit, goat from Millipore and Sigma

Protocol: Proteins resolved on SDS-PAGE were transferred to the activated polyvinyl difluoride (PVDF) membrane placed in form of the sandwich and wet electro-blotting was performed at constant voltage of 100V for 2 h. The transfer of proteins was visualized using reversible Ponceau-S staining. The membrane was incubated in blocking solution for 1 h at RT on a rocker. The blots were rinsed in TBST to remove residual BSA. After blocking, the membrane was incubated with diluted primary antibody for 1 h at RT on the rocker. The blot was then washed at least thrice with TBST for 5 mins each. It was then incubated with horseradish peroxidase (HRPO) conjugated secondary antibody (anti mouse/ anti rabbit/ anti goat) for 1 h at RT. The unbound secondary antibody was removed and the blot was washed thrice with TBST for 10 mins each. The blots were developed using ECL prime chemiluminescence reagent according to the manufacturer's protocol and the signal captured onto X-ray films.

3.2.8 Cell death assays

a) Cell viability analyses by flow cytometry: The estimation of the viability of a cell population by flow cytometry is based on a simple yet powerful principle, that the dead cells are leaky. All the different methods for evaluating viability are based on either direct leak detection or measurement of a direct consequence of this leakage. Dyes are used that either do not leak into or out of viable cells. These methods can clearly identify whether the cells are damaged or dying but not completely dead.

Propidium iodide (PI) is a membrane impermeant dye that is generally excluded from viable cells. It binds to double stranded DNA by intercalating between base pairs. PI is excited at 488 nm and, with a relatively large Stokes shift, emits at a maximum wavelength of 617 nm. Because of these spectral characteristics, PI can be used in combination with other fluorochromes with same excitation wavelength such as green fluorophores including fluorescein isothiocyanate (FITC) and phycoerythrin (PE).

Reagents: 1X PBS, PI staining solution - 1mg/ml PI (Sigma) in PBS.

Protocol: Post transfection, cells were trypsinized, and washed twice with PBS. Approx. one million cells were re-suspended in 500 μ l PBS containing 4 μ g/ml PI, and analyzed by FACS Calibr (BD Bioscience) for GFP fluorescence and PI content. Percent of dead cells are the % of GFP and PI-positive cells /GFP-positive cells.

b) Caspase 8 activity assay: Caspase-8 is activated early in apoptosis and is involved in the proteolysis and activation of downstream executioner procaspase-3. To identify and quantify caspase-8 activity in apoptotic cell lysates, an upstream sequence of the site recognized by active caspase-8, IETD (Ile-Glu-Thr-Asp), is utilized as a basis for the highly specific caspase-8 substrate. The caspase-8 fluorimetric assay is based on the hydrolysis of the

peptide substrate N-Acetyl-Ile-Glu-Thr-Asp-7-amino-4-(trifluoromethyl) coumarin (Ac-IETD-AFC) by caspase-8 resulting in the release of a 7-amino-4-(trifluoromethyl) coumarin (AFC) moiety. AFC release can be monitored in a spectrofluorometer at an excitation wavelength of 400 nm and an emission wavelength range of 480-520 nm.

Reagents: Caspase assay buffer (20 mM HEPES pH 7.7, 100 mM NaCl, 1 mM DTT, 1 mM EDTA, 0.1% CHAPS, 10% sucrose, 0.5 % NP-40), Ac-IETD-AFC (Enzo Life Sciences)

Protocol: Cells were extracted in caspase assay buffer containing 0.5% NP-40. Caspase activity was measured in cell lysate containing 20 - 30 µg proteins with 0.002 µM of Ac-IETD-AFC in 100 µl of caspase assay buffer for 2 h at 37 °C. Ac-IETD hydrolysis was measured by monitoring emission at 510 nm on excitation at 405 nm.

c) DNA ladder assay: Apoptosis is characterized by the activation of endonucleases with subsequent cleavage of chromatin DNA into inter-nucleosomal fragments of roughly 180 base pairs (bp) and multiples thereof (360, 540 etc.). Separation of DNA fragments by agarose gel electrophoresis and subsequent visualisation by ethidium bromide staining results in a characteristic "ladder" pattern. While most of the morphological features of apoptotic cells are short-lived, DNA laddering can be used as final state read-out method and has therefore become a reliable method to distinguish apoptosis from necrosis.

General Protocol

1. Collect culture media; add 1 ml of trypsin to cell monolayer on 100-mm dishes, scrape the cells, harvest cells (culture media and cell monolayer) by centrifugation (2,500 rpm, 5 min), and wash cell pellets with 1X PBS
2. Add 100 µl of lysis buffer (1% NP-40 in 20 mM EDTA, 50 mM Tris-HCl pH 7.5)

3. Centrifuge at 3,000 rpm for 5 min and obtain the supernatant
4. Add 10 μ l of 10% SDS solution to pooled supernatant (final: 1% SDS), treat with 10 μ l of 50 mg/ml RNase A (final 5 μ g/ μ l) and incubate for 2 h at 56 °C
5. Add 10 μ l of 25 mg/ml Proteinase K (final 2.5 μ g/ μ l) and incubate for 2 h at 37 °C
6. Add 1/2 vol. (65 μ l) of 10 M ammonium acetate and 2.5 vol. (500 μ l) of ice-cold ethanol and mix thoroughly. Stand for 1 h in -80 °C (“ethanol precipitation”)
7. Centrifuge for 20 min at 12,000 rpm, wash the white pellet with 200 μ l of 80% ice cold ethanol followed by air-drying for 10 min at room temperature
10. Dissolve the pellet in 50 μ l of TE buffer
11. Determine the DNA concentration (Abs 260 nm), and resolve about 4 μ g of DNA on 2% agarose gel.

3.3 Protein-protein interaction analyses

There are several techniques by which protein-protein interaction can be studied; however, the requirements of each of these methods are unique and dependent on many properties of the ligand and analyte. Usually a combination of techniques is necessary to validate, characterize and confirm protein interactions. Following were the techniques used for measurement of protein-protein interaction in our study:

3.3.1 Pull-down assay

Pull-down assays are commonly used for qualitative measurement of interactions between two or more proteins. A “bait” protein is tagged and captured on an immobilized ligand (support beads) through an affinity tag such as GST, MBP or 6X His. This forms a complex, a “secondary affinity support”, which is then incubated with cell lysate or [³⁵S] methionine

labeled or unlabeled proteins. Protein complexes are either eluted or directly loaded onto SDS-PAGE and are analyzed by staining, western blotting or autoradiography.

3.3.2 Co-Immunoprecipitation

Co-immunoprecipitation (co-IP) is similar in methodology to pull-down assays because of the use of beaded support to purify interacting proteins. The difference between these two approaches, though, is that while pull-down assays use a bait protein to purify any proteins, co-IP uses antibodies to capture protein complexes in a lysate. In a typical experiment, cells are lysed and a whole cell extract is prepared under non-denaturing conditions. It is essential to use non-denaturing conditions in order to maintain any interactions that occur. An antibody specific to the bait is then added to the extract, forming a primary complex. This complex is then immobilized on protein A or protein G sepharose beads. Proteins that do not bind are removed by a series of washes. The protein complex is then eluted from the beads and dissociated by SDS sample buffer. Samples are then evaluated by SDS-PAGE followed by western blotting with specific antibodies against the binding partners.

3.3.3 Isothermal Titration Calorimetry

Isothermal Titration Calorimetry (ITC) is the most quantitative probe available for measuring the thermodynamic properties of a protein-protein interaction. It measures the binding equilibrium directly by determining the heat evolved (exothermic) or absorbed (endothermic) on association of a ligand with its binding partner. In a single experiment, the values of the binding constant (K_a), stoichiometry (n), and the enthalpy of binding (ΔH) are determined. The free energy (ΔG) and entropy (ΔS) of binding is then calculated using the equation:

$$\Delta G = -RT \ln K_a = \Delta H - T\Delta S \quad (\text{Eq.1})$$

(where R is the gas constant and T is the absolute temperature in Kelvin)

An isothermal titration calorimeter consists of two identical cells (reference cell and sample cell) composed of highly thermal conducting and chemically inert substance (Hastelloy alloy or gold), surrounded by an adiabatic jacket. The temperature difference between the two cells containing the macromolecule is detected and maintained using sensitive thermocouple circuits. A known concentration of ligand (loaded in the syringe) is titrated into the sample cell, in precisely known aliquots, which upon interaction results into heat dissipation or generation. Heat change during the reaction is quantified in terms of time-dependent input of power required to maintain equal temperatures between the sample and reference cells. In an exothermic reaction, feedback power of the sample cell is decreased because of increase in the temperature of sample cell while opposite occurs in case of an endothermic reaction.

We used an iTC200 system from Microcal (GE Healthcare), one of the most sensitive isothermal titration calorimeters available. The sample cell is only 200 μl . It can directly measure sub-millimolar to nanomolar binding constants (10^{-7} to 10^{-9}M^{-1}).

3.3.4 Surface Plasmon Resonance

Surface Plasmon Resonance (SPR) is one of the fastest methods that allow real-time, label free detection of bimolecular interactions. SPR phenomenon occurs when a polarized light, under conditions of total internal reflection, strikes an electrically conducting surface at the interface between two media. This generates electron charge density waves called plasmons, reducing the intensity of reflected light at a specific angle known as the resonance angle; in proportion to the mass on a sensor surface. The basic principle involves the changes in the reflective index as a result of absorption of molecules onto prefixed ligand.

The molecular interaction studies were performed using Biacore 3000 (GE Healthcare). In Biacore instruments, the sensor surface comprises a micro-fluidic flow cell ($\sim 20 - 60 \text{ nl}$).

Through this flow cell an aqueous solution (running buffer) is passed under a continuous flow rate (1 - 100 $\mu\text{l}/\text{min}$). To allow the detection of an interaction, desired molecule (termed as the ligand) is immobilized onto the sensor surface. Its binding partner (analyte) is then injected in aqueous solution (ideally with the same components and composition as the running buffer) through the flow cell, under continuous flow. As the analyte interacts with the ligand the accumulation of mass on the sensor surface causes a change in refractive index. This is measured in real time, and the result is plotted as response units (RU) versus time (termed a sensorgram). Importantly, a signal (background response) will also be generated if there is a difference in the refractive indices of the running and sample buffers. This background response must be subtracted from the sensorgram to obtain the actual binding response. The background response is recorded by injecting the analyte through a control or reference flow cell, which has no ligand or an irrelevant ligand immobilized to the sensor surface. The real time measurement of association and dissociation of a binding interaction allows for the calculation of association (k_a) and dissociation rate constants (k_d) and the corresponding equilibrium dissociation constants (K_D) can be determined from the ratio of the dissociation and association rate constants (k_d/k_a).

Detailed experimental procedures for each of these techniques are given in the respective chapters.

3.4 Biophysical characterization

Various biophysical techniques are used to structurally characterize a given protein in solution. The techniques used in the present work are described below:

3.4.1 Circular Dichroism Spectroscopy

Circular Dichroism (CD) spectroscopy measures differences in the absorption of left-handed versus right-handed polarized light which arise due to structural asymmetry. Once the protein

molecule differentially absorbs the left and right handed light, the emerging light is elliptically polarised. CD spectra in the far UV range (260-180 nm) can be analyzed for the different secondary structural components of the protein such as alpha helix, parallel and antiparallel beta sheet, turn, and others. Absorption minima at λ 208 nm and 222 nm indicate α -helical structure, whereas a minimum at λ 218 nm is a characteristic of β -sheets. The disordered secondary structural elements or random coil protein are characterised by a low ellipticity at 210 nm and negative band near λ 195 nm. Based on the high agreement between secondary structures derived from CD and X-ray crystallography, several algorithms have been developed to provide an estimation of the secondary structure composition of proteins from CD data. Widely used algorithms include SELCON (self-consistent), VARSLC (variable selection), CDSSTR, K2D and CONTIN. An online server DICHROWEB that provide flexibility of analyzing data by various algorithms and databases as well as using several reference sets is used for the data analysis. CD spectra therefore can rapidly determine the overall folding and secondary structural elements of the protein. The method is also very reliable for monitoring changes in the conformation of proteins under different conditions such as denaturation studies, unfolding experiments, mutational analysis, etc.

3.4.2 Fluorescence spectroscopy

Fluorescence spectroscopy, also known as fluorometry, exploits the phenomenon of fluorescence in which a molecule absorbs a lower wavelength photon, undergoes electronic excitation, and then emits longer wavelength. The fluorescence of a folded protein is a combination of the signal from individual aromatic residues with the major contribution from tryptophan residue. It can be an indirect probe of the proteins environment, describing the solvent-solute as well as inter- and intra-molecular interactions. After exciting a sample, a

'Fluorometer' allow us to obtain and measure various attributes including intensity, spectral wavelengths and lifetime.

Protein emission scans are acquired from λ 310 to 400 nm, following excitation at λ 280 or 295 nm. The λ_{280} monitors the environmental changes associated with tryptophan and tyrosine side chains due their absorption wavelengths at 280 and 275 nm, respectively. However, emission spectrum corresponding to λ_{295} selectively follows tryptophan emission because of least contribution from tyrosine. The fluorescence intensity and λ_{max} are monitored to study the protein-folding, ligand-binding and so on.

Using these techniques structural alterations, protein unfolding was monitored and the melting temperature was determined for different proteins.

a) Secondary structural analysis

Material: Phosphate buffer (20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ buffer pH 7.8 with 20 μM β -mercaptoethanol), Quartz cuvette (1 mm), CD Polarimeter (Jasco, J815)

Protocol: Far-UV CD data of 10 μM protein in phosphate buffer were collected between λ 250 and 190 nm (Settings: Scan speed 20 nm/s, accumulation-3, data pitch 0.1, and temperature 25 ° C).

Data analysis

CD data from at least three independent experiments were used for data analysis. The spectra were plotted as ellipticity (θ) on Y-axis and wavelength (nm) on X-axis. Data is represented in the form of the Mean Residual Ellipticity (MRE or $[\theta]$) given as $\text{deg.cm}^2.\text{mol}^{-1}$, since the value is concentration independent and constant for a protein. Ellipticity was converted to mean residue ellipticity using the formula, $[\theta]_{\text{MRE}} = (\theta * \text{MRW}) / (10 * c * d)$

where, MRW (Mean residue weight) = Molecular weight / (number of amino acids -1)

'c' is concentration of protein (mg/ml), d is the pathlength in cm

The data is then saved in Dichroweb format and subsequently analyzed by Dichroweb server (<http://dichroweb.cryst.bbk.ac.uk>).

b) Thermal Denaturation

Thermal stability of protein is an index of its overall stability. It can be assessed by monitoring the CD spectrum with increasing temperature. Far-UV CD region can be used to evaluate the secondary structural changes in proteins. The co-operative nature of melting curve shows that the protein is well folded.

Reagents: Phosphate buffer (20 mM Na₂HPO₄/NaH₂PO₄ buffer pH 7.8 with 20 μM β-mercaptoethanol)

Protocol: Thermal denaturation of wild-type and mutant proteins was done simultaneously using multi-cell cuvette holder. A Far-UV CD spectrum (λ 250 to 190 nm) was collected in a temperature range of 10 °C to 90 °C with an increment of 1 °C/min. At each data point, the sample was equilibrated for 5 mins.

Data analysis

Ellipticity corresponding to 222 nm at different temperatures was obtained for calculation of melting temperature (T_m). Firstly, ellipticity of fully folded (θ_f) and unfolded forms (θ_u) were estimated using nonlinear regression (GraphPad Prism). This was then used to calculate fraction folded at any temperature (α) with the following formula,

$$\alpha = [F] / ([F] + [U]) = (\theta_t - \theta_u) / (\theta_f - \theta_u)$$

where [F] and [U] are concentration of folded and unfolded forms respectively, and θ_t is the observed ellipticity at a given temperature. To calculate the T_m, the fraction folded at given temperature was further analyzed using nonlinear regression (Igor Pro).

c) Equilibrium unfolding studies

Reagents: Phosphate buffer (20 mM Na₂HPO₄/NaH₂PO₄ pH 7.8, 20 mM β-mercaptoethanol), 10 M urea stock solution prepared in phosphate buffer. The solutions were made freshly for each experiment and were filtered (0.22 μm pore size) prior to use.

Protocol: The experiments were performed as described previously [125]. Briefly, a protein stock solution (ten times the final concentration to be used in the experiment) was prepared in phosphate buffer. Phosphate buffer, urea from 10 M stock solution, and protein solutions were mixed in 2 ml siliconized microcentrifuge tubes to obtain final urea concentration ranging between 0 M and 8 M. For renaturation experiments, the protein was incubated in 8 M urea-containing buffer. After incubation for 1 h at 25 °C, the protein was diluted in phosphate buffer such that the final urea concentration was as indicated in the figure legends. The samples were incubated for 20 h at 25 °C prior to data collection. The incubation time was such that it was sufficient to allow all samples to equilibrate.

Fluorescence emission at each urea concentration was measured using JOBIN YVON Horiba Fluorolog 3 spectrofluorometer. The measurements were acquired at excitation wavelengths of 280 nm and 295 nm with fluorescence emission in the range of λ 310 and 400 nm. Raw data corrected for buffer background was normalized between zero (unfolded) and one (native), as shown in Eq. 2, in order to visualize different spectroscopic signals on a single scale.

$$Y_{\text{Normalized}} = (Y_x - Y_U) \div (Y_N - Y_U) \quad . \text{ (Eq. 2)}$$

(where Y_x is the signal being normalized, Y_U is the signal of the unfolded protein, and Y_N is the signal of the native protein).

Circular dichroism at 228 nm was measured with a Jasco J-815 spectropolarimeter using a cuvette of 1 mm pathlength. Both instruments were equipped with thermostated cell holders, and the temperature was held constant at 25 °C using a circulating water bath.

Data analysis

For equilibrium studies using fluorescence spectroscopy, average emission wavelength was determined for emission scans using the following equation,

$$\langle \lambda \rangle = \sum_{i=1}^N (I_i \lambda_i) / \sum_{i=1}^N (I_i) \quad (\text{Eq. 3})$$

where $\langle \lambda \rangle$ is the average emission wavelength and I_i is the fluorescence emission at wavelength λ_i .

Fluorescence and circular dichroism data set were fitted using the three state (dimeric intermediate) or two state equilibrium model using Igor Pro (6.03A).

i) Three state dimeric intermediate model



In this model, the protein is assumed to be either in the native homodimeric state (N_2), a non-native dimeric state (I_2), an unfolded monomeric state (U) and K_1 , K_2 are the equilibrium constants for the two steps, respectively. If we consider the total molar concentration of the polypeptide chain as P_T , as shown in equation 5,

$$P_T = 2[N_2] + 2[I_2] + [U] \quad (\text{Eq. 5})$$

the mole fraction of each species can be defined as shown in equations 6-8.

$$f_{N_2} = \frac{2N_2}{P_T} \quad (\text{Eq. 6})$$

$$f_{I_2} = \frac{2I_2}{P_T} \quad (\text{Eq. 7})$$

$$f_U = \frac{U}{P_T} \quad (\text{Eq. 8})$$

The sum of all fractions is equal to unity as shown in equation 9.

$$f_{N_2} + f_{I_2} + f_U = 1 \quad (\text{Eq. 9})$$

The equilibrium constants K_1 and K_2 are related to the mole fraction of each species and to P_T , as shown in equations 10 and 11.

$$K_1 = \frac{f_{I_2}}{f_{N_2}} \quad (\text{Eq. 10})$$

$$K_2 = \frac{f_U}{f_{I_2}} \quad (\text{Eq. 11})$$

Equating equations 9, 10, substituting in terms of f_U , and rearranging yields the following equation,

$$\frac{2f_U^2 P_T}{K_1 K_2} + \frac{2f_U^2 P_T}{K_2} + f_U - 1 = 0 \quad (\text{Eq. 12})$$

By solving the quadratic equation 11, the fraction of each species is obtained, as shown in equation 13-15

$$f_U = \frac{-K_1 K_2 + \sqrt{(K_1 K_2)^2 + 8P_T(K_1 K_2 + K_1^2 K_2)}}{4P_T(1 + K_1)} \quad (\text{Eq. 13})$$

$$f_{I_2} = \frac{2f_U^2 P_T}{K_2} \quad (\text{Eq. 14})$$

$$f_{N_2} = \frac{f_{I_2}}{K_1} \quad (\text{Eq. 15})$$

From equations 12-14 and eq. 1, one can calculate the equilibrium constant and the values of ΔG at each urea concentration. We assume the free energy change for each step in the reaction to be linearly dependent on denaturant concentration as described earlier (Equations 16 and 17).

$$\Delta G_1 = \Delta G_1^{\text{H}_2\text{O}} - m_1 [\text{denaturant}] \quad (\text{Eq. 16})$$

$$\Delta G_2 = \Delta G_2^{\text{H}_2\text{O}} - m_2 [\text{denaturant}] \quad (\text{Eq. 17})$$

where $\Delta G_1^{\text{H}_2\text{O}}$ and $\Delta G_2^{\text{H}_2\text{O}}$ are the free energy changes in the absence of denaturant corresponding to K_1 and K_2 respectively, and m_1 and m_2 are the cooperativity indices associated with each step. The amplitude of the spectroscopic signal determined at each urea concentration is assumed to be a linear combination of the fractional contribution from each species (equation 18),

$$Y = Y_{N_2} f_{N_2} + Y_{I_2} f_{I_2} + Y_U f_U \quad (\text{Eq. 18})$$

where Y_{N_2} , Y_{I_2} and Y_U are the amplitudes of the signals for the respective species.

In order to determine the unknown parameters $\Delta G_1^{\text{H}_2\text{O}}$, $\Delta G_2^{\text{H}_2\text{O}}$, m_1 and m_2 , the data sets shown in figures 4.3.4 and 4.3.5 were fit simultaneously using Igor Pro.

ii) Two state monomer model

In this model, the protein is assumed to be either in the native monomeric state (N) or in an unfolded monomeric state (U), and K is the equilibrium constant for the unfolding process.



Therefore the total protein concentration is,

$$P_T = [N] + [U] \quad (\text{Eq. 20})$$

The equilibrium constant K is related to the mole fraction of both the species and is given as,

$$K_1 = \frac{[U]}{[N]} \quad (\text{Eq. 21})$$

The mole fraction of each species can be defined as shown in equations 22 and 23,

$$f_N = \frac{1}{1+K} \quad (\text{Eq. 22})$$

$$f_U = \frac{K}{1+K} \quad (\text{Eq. 23})$$

Therefore, the fractional contribution from each species is given by the equation,

$$Y = Y_N f_N + Y_U f_U \quad (\text{Eq. 24})$$

(where Y_N and Y_U are the amplitudes of the signals for the respective species)

To determine the unknown parameters $\Delta G^{\text{H}_2\text{O}}$ and m_1 , the data set was fitted simultaneously using Igor Pro.

3.5 Protein oligomerization and size characterization

An oligomer is a macromolecular complex formed mostly due to non-covalent bonding of the molecules, and may or may not have biological significance. Homo-oligomerization involves

assembly of identical molecules together, whereas, hetero-oligomer involves two or more different macromolecules. Protein oligomerization can be detected and characterized using different methodologies. However each technique has its own pros and cons, therefore several approaches should be combined to draw a conclusion. The methodologies used in the present work are described below:

3.5.1 Size-exclusion chromatography

Size-exclusion chromatography (SEC) is an analytical technique that separates dissolved macromolecules based on their size and shape (hydrodynamic radius). There are two basic types of size exclusion chromatography. One is gel permeation chromatography, which uses a hydrophobic column packing material and a non-aqueous mobile phase (organic solvent) to measure the molecular weight distribution of synthetic polymers. The other is gel filtration chromatography, which uses a hydrophilic packing material and an aqueous mobile phase to separate, fractionate, or measure the molecular weight distribution of molecules soluble in water, such as polysaccharides and proteins. Separation is carried out by means of a porous separation matrix with different sized cavities. Completely solvated polymer chains diffuse according to their hydrodynamic volume into the cavities of the separation matrix and are sorted in terms of their size.

The molecular weight of an unknown protein is calculated based on the time taken to move through the gel column (Superdex 75 or 200) as compared to the time by the mixture of standard known proteins such as alcohol dehydrogenase, bovine serum albumen, lysozyme, and MBP. The proteins with the higher molecular weights travel faster as they are excluded from the gel compared to low molecular weight proteins. Elution volume (V_e)/void volume (V_0) vs. log of molecular masses of standards was plotted to generate the calibration curve. The molecular weight of the unknown protein is 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).

3.5.2 Glutraldehyde cross-linking

Crosslinking is the process of chemically joining two or more molecules by a covalent bond. Crosslinking reagents such as glutaraldehyde contain two or more reactive ends capable of chemically attaching to specific functional groups (primary amines, sulfhydryls, etc.) on proteins. Two physically interacting proteins can be covalently cross-linked and the formation of bridge between two proteins suggests their existence in a close proximity.

Reagents: Freshly prepared 0.23% glutaraldehyde stock, phosphate buffer (pH 7.5), 2 M Tris-HCl (pH 8.0).

Protocol: Reaction mixtures with 20 μ g of purified protein in 10 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ buffer (pH 7.5) was treated with freshly prepared solution of glutaraldehyde (final concentration 0.1%) for 2 min at 37 °C. This reaction was terminated by adding 5 μ l of 2 M Tris-HCl, pH 8.0. Cross-linked product was mixed with equal amount of Laemmli SDS sample buffer and analyzed on 12% SDS-PAGE.

3.5.3 Dynamic light scattering

Dynamic Light Scattering (Photon Correlation Spectroscopy) is one of the most popular technique for measuring the size and distribution of molecules and particles typically in the submicron region. The Brownian motion of particles or molecules in suspension causes laser light to be scattered at different intensities. Analysis of these intensity fluctuations yields the velocity of the Brownian motion and hence the particle size using the Stokes-Einstein

relationship. Most of the "particle size analyzed" operate at 90° and use red light of wavelength 675 nm. DLS is also capable in measurements of several parameters such as molecular weight, radius of gyration, translational diffusion constant etc.

The molecular size estimation in the present work was done using Wyatt technology DynaPro particle size analyzer.

Reagents: Protein solution (1 mg/ml) in phosphate buffer

Protocol: Protein and buffer solution was filtered (0.44 µm pore size) and degassed prior to measurement. 1mg/ml protein in 10 mM Na₂HPO₄/NaH₂PO₄ buffer (pH 7.5), 100 mM NaCl was loaded into a 45 µl quartz cuvette. Measurements were performed at temperature 25 °C and at least 30 – 40 measurements each of 12 s duration were collected. The refractive index and viscosity values were taken for the water as provided by the software. The translational diffusion coefficient of the protein was calculated from the autocorrelation of scattered light intensity. Histogram analyses of DLS results were carried out using the software DYNAMICS v.6.0

3.6 Serine protease enzymatic assay

The protease activity of wild-type HtrA2 and its variants were determined using a generic substrate for serine proteases, FITC (fluorescein isothiocyanate) labelled β-casein (Sigma). FITC-fluorescence was monitored in a multi-well plate reader (Berthold Technologies) using excitation and emission wavelengths of 485 and 545 nm respectively. Reaction rates (v_0) were calculated using linear regression analysis. The steady-state kinetic parameters were calculated from the reaction rates by fitting the data to Hill form of Michaelis-Menten equation, $velocity = V_{max}/(1+(K_{0.5}/ [substrate])^n)$, where ' V_{max} ' is the maximum velocity and

$K_{0.5}$ is substrate concentration at half maximal velocity using KaleidaGraph (Synergy software). All the experiments were done independently in triplicate.

3.7 Molecular Modeling, Docking and Simulation

Molecular modeling involves representation of three-dimensional structures of bio-molecules and their physico-chemical properties. It can be obtained using *in-silico* tools which produce, estimate and predict reasonable atomic level structures and related characteristics of molecules. It is based on the principle of molecular mechanics which is combined with various computational tools to obtain energetic and structural information of target bio-molecules using theoretical and experimental data. The entire essence of molecular modeling resides in the relationship between the *microscopic* and the *macroscopic* observations that belong to statistical mechanics. Macroscopic observations defined by solvation energy, hydrogen-hydrogen bond distance, affinity between two bio-molecules and molecular conformation show a great harmonization with average of observable over selected microscopic states. Thermodynamically determined macroscopic behavior of a system is equivalent to a quantity called the partition function (Z). The partition function is a rather very complex to compute therefore, generally a numerical approximations is considered.

$$Z = \sum_i (e^{-\beta}) E_i$$

Numerical approximations involve the following contents:

- 1) Calculation of the system energy for microstate i - performed using semi-empirical force fields by GROMOS /Amber/ CHARMM.
- 2) Sampling of the microstates accessible in a given macroscopic state, i.e. micro-canonical sampling, canonical sampling and isothermic-isobaric sampling for fixed N, V, E ; N, V, T and N, P, T systems respectively.

3.7.1 Types of Interactions: Different kinds of non-bonded interactions that are present in interacting molecules are described below:

1) Electrostatic interactions

Coulomb law: The attraction and repulsion between charged bodies is directly proportional to product of their charges and inversely to distance between them which is given as,

$$V_{Ele} = \sum_{i>j} \frac{q_i q_j}{4\pi\epsilon r_{i,j}}$$

where, ϵ represents dielectric constant (value 1 for vacuum, 4 - 20 for protein core and 80 for water), q_i and q_j are the charge on interacting bodies and r is the distance between them.

2) van der Waals interactions

It involves *attractive part* which exist due to induced-dipole/dipole and *repulsive part* due to Pauli's Exclusion Principle. It is usually represented by the Lennard-Jones potential which is obtained from the single atom parameter ϵ and σ .

3) Hydrogen bonds

Interaction of the type is D-H \cdots A. The origin of this interaction is a dipole-dipole attraction.

Typical ranges for distance and angle: 2.4 - 4.0 Å (D-A) and 90°-180° (D-H \cdots A).

4) Hydrophobic interactions

This interaction exists due to combined effect of contact between water, polar medium and hydrophobic groups, and is usually present with in a distance of 4.7 Å.

Docking: The interaction between bio-molecules such as carbohydrates, lipids, nucleic acids and proteins play an essential role in cellular processes. Furthermore, the binding orientation of the two molecules may affect the type of signal produced such as agonism and antagonism. Docking is a computational simulation method of predicting the preferred orientation of one molecule to a second in a stable complex. The relative preferred orientation may be used to

determine the strength of association between two molecules. It aims to reach an optimized conformation and relative orientation for both the protein and ligand such that the system attains a lowest energy state.

Docking efficiency can be evaluated from a particular pose by counting the number of favorable intermolecular interactions such as hydrophobic contacts and hydrogen bonds. This method is most widely used for classifying a ligand which interacts favourably to a particular receptor, based on the predicted free-energy of binding. Docking is also very frequently used to predict the association of drug and receptor targets in order to determine the binding affinity and molecular kinetics.

Docking model: Earlier assumption about docking was a *lock-and-key* mechanism between receptor and ligand in which only the favorable conformation of both interacts and form a stable complex. However, soon the concept adapted to a more feasible aspect of *induced fit* model in which the ligand and the protein adjust their conformation to achieve an overall "best-fit". Due to the flexible nature of both the ligand and protein, the later model is widely accepted but not necessarily applicable to all proteins and ligand.

Molecular dynamics (MD): It is a computer simulation that allows interaction of atoms and molecules for a period of time, providing a view of their physical movements. The trajectories of atomic and molecular motion are determined by considering forces between the particles and potential energy as defined by molecular mechanics force fields. MD simulation use numerical methods to find the properties of molecular systems. In simulation of docking process, the protein and the ligand are physically separated, and then ligand is allowed to find its position into the protein's active conformation. The moves incorporate internal changes to the ligand's structure, translations and rotations. The advantage of

docking simulation is that it more accurately models reality, and ligand flexibility is easily incorporated.

3.7 2 General steps for molecular modeling, simulation and docking:

1. Crystal structures retrieval from Protein Data Bank (PDB) (<http://www.rcsb.org/pdb/>) which can be used as a model template.
2. Modeling: Homology modeling or *ab-initio* modeling
3. Model energy minimization: Energy level is the most basic property of molecules that can be calculated using three major theoretical methods, (i) molecular mechanics, (ii) semi-empirical (quantum mechanics), (iii) *ab-initio* (quantum mechanics) approach. Energy minimization is the first step carried out for geometry optimization of the molecular structure.
4. Dynamics simulation and conformation search: Integration of molecular dynamics (solving Newton's law of motion for the nuclei) for all atoms in the system generates molecular trajectories. Conformation search of bio-molecules is carried out by repeating the procedure of rotating dihedral angles to achieve lowest energy conformations of molecular systems.
5. Calculation of molecular properties: Some physicochemical properties such as thermodynamic quantities, solubility, molar volume, heat capacity, molar refractivity density, dipole moment, magnetic susceptibility, partial atomic charge, ionization potential, electrostatic potential, solvent accessible surface area and van der Waals surface area are computed.
6. Structure superposition and alignment: It involves comparison of series of homologous molecules to get the best energy minimized model, which requires superposition or alignment of structures.

7. Molecular interactions, docking: In intermolecular interaction studies, usually the receptor (e.g., protein) is kept rigid or partially rigid while the conformation of ligand/protein is dynamic. Various docking software available includes AUTODOCK, HAADOCK, PATCHDOCK, ClusPro, etc.

CHAPTER-4

Human papillomavirus E2 protein

Introduction:

Viruses have evolved complex strategies to manipulate cellular apoptotic machinery of infected cells for their own benefit. They selectively inhibit or induce apoptosis for efficient viral replication and evasion of host defence mechanisms or for virus propagation. Oncogenic high risk HPV displays anti-apoptotic activities mainly mediated by oncoprotein E6, which interferes with both intrinsic and extrinsic pathways by targeting p53 [126] and FADD [127] for degradation and the pro-apoptotic role has been attributed to a transcription factor E2 protein [7, 128]. However, this activity of E2 is only associated with high risk viruses, such as HPV16 and 18, while the low risk HPV types 6 and 11, which only cause benign lesions, do not induce cell death [129]. Furthermore, several studies have shown that E2-mediated apoptosis is not specific to transformed cells and occurs in primary epithelial as well and is independent of its virus associated functions including transcription and replication [48]. Different mechanisms have been proposed for different types of high risk HPVs. It has been reported that HPV18 E2 activates extrinsic death receptor pathway, via physical interaction with caspase-8, a component of 'Death Inducing Signaling Complex'. This interaction bypasses the requirement of upstream death-fold adaptor proteins such as FADD thereby representing a novel *adaptor-independent caspase activation pathway* [1]. On the other hand, HPV16 E2 involves p53-mediated apoptosis or abrogation of apoptosis-inhibitory function of cFLIP thereby sensitization to the extrinsic apoptotic pathway [49, 130]. Nevertheless, induction of apoptosis by high risk HPVs has been linked to their cytoplasmic accumulation due to an active nucleo-cytoplasmic shuttling and activation of apical caspase-8 [131].

Formation of the death-inducing signaling complex (DISC) is a critical step in extrinsic cell death pathway [132]. The current model for DISC formation proposes that the ligation of Fas or TRAIL-R1/TRAIL-R2 receptors trigger the recruitment of death-fold domain adaptor proteins such as FADD via homotypic interaction between death domain (DDs) within the

receptor and FADD. The death effector domain (DED) of FADD then recruits DED-only proteins including procaspase-8 or c-FLIP, forming an active DISC. This enables activation of the initiator zymogenic procaspase-8, a process that requires both proximity induced dimerization of the catalytic domain and autoproteolytic cleavage [133]. Once activated, activate caspase-8/caspase-10 initiate the caspase cascade, directly through cleavage of downstream executioner caspases, ultimately resulting in apoptosis [134]. This suggests caspase-8 strictly depends on homotypic binding to death-fold domain adaptor proteins such as FADD to oligomerize and to self-activate [67, 135, 136]. Only few adaptors for caspase-8 activation have been described so far, and they all contain DED or pseudo-DEDs [137, 138]. Therefore, it is intriguing to understand how E2 a non-death-fold domain protein physically interacts with caspase-8 to promote apoptosis and how is it different from the classical FADD – procaspase-8 interaction.

Here, we have characterized the molecular basis of E2 – procaspase-8 interaction and defined the surface for the physical interaction between them. Furthermore, we present a fascinating insight into the mechanism of classical FADD – procaspase-8 interaction and attempted to establish a relationship between these two mechanisms of apoptosis.

The present chapter is divided into three sub-chapters:

- 4.1 Understanding protein-protein interactions involved in non-death-fold E2-induced apoptosis
- 4.2 Molecular details of DED assembly and its role in death-fold mediated caspase-8 activation
- 4.3 Biochemical and biophysical characterization of E2 transactivation domain

CHAPTER-4.1

4.1 Understanding protein-protein interactions involved in non-death-fold E2-induced apoptosis

4.1.1 Experimental Procedures

4.1.1.1 Plasmids – c-DNA encoding the wild-type caspase-8 cloned in pET29b caspase-8 and pET29b FADD constructs were provided kindly by Prof. Guy Salvesen (Burnham Medical Research Institute, USA) and pEGFPc1 HPV18 E2 full length construct was obtained as a kind gift from Dr. F. Thierry (Institut Pasteur, France). For co-expression of E2 and procaspase-8 in *E. coli*, we sub-cloned the c-DNA encoding HPV18 E2 transactivation domain (residues 1-201) in chloramphenicol resistant pACYC-vector (Clontech) using restriction enzymes NcoI and HindIII. A C-terminal His-tag was introduced in this vector by site-directed mutagenesis, to facilitate protein purification. The full length procaspase-8 and its prodomain DED-A (1-80), DED-B (100-181) and DED-AB (1-181) were subcloned in ampicillin resistant pMALc5-E-TEV vector digested with restriction enzymes BamHI and EcoRI. For expression in mammalian cells, procaspase-8 full length cDNA were sub-cloned in pEGFPn1, pmCherry-n1 and pcDNA3.0 vectors. The mutants for all the proteins were generated using QuickChange site-directed mutagenesis kit (Stratagene). Sequences of all constructs were verified by automated DNA sequencing (ACTREC sequencing facility).

4.1.1.2 Protein co-expression, purification and *in vitro* complex formation assay – For protein co-expression, the plasmids were transformed in *E.coli* BL21 (DE3) cells. The cells were grown in the medium supplemented with 2 µg/ml of arabinose (to induce expression of pACYC-E2 TAD) at 37 °C till the optical density ($A_{600\text{ nm}}$) reaches 0.5. Cells were further induced with 0.4 mM IPTG to induce expression of MBP-tagged constructs and cultured at 18 °C for 20 h. After co-expression of the proteins, the cell lysate (in 20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$, pH 7.8 containing 200 mM NaCl, 0.1 mM DTT, and 1% glycerol) was

passed through pre-equilibrated Ni-IDA or amylose resin for 1 h at 4 °C. Beads were washed three times with binding buffer (20 mM Na₂HPO₄/NaH₂PO₄, pH 7.8 containing 200 mM NaCl, 1 mM DTT, 0.1% TritonX-100) and the bound proteins were later eluted using imidazole or maltose respectively. The complexes were separated on 12% SDS-PAGE and were immunoblotted with 1:1000 dilution of anti-E2 antibody (sc-26939, Santa Cruz Biotechnology, Inc.).

4.1.1.3 Analysis of the stability of E2-DED complex – The protocol was adapted as described earlier for Fas-FADD complex [52]. To evaluate the stability of the complex, isolated His₆ tagged E2 TAD – MBP fused DED-B complex was diluted to different final concentrations in the range (500 nM - 20 µM). The diluted complexes were then mixed with 50 µl of amylose resin and incubated for 1 h at 4 °C. The beads were given a rapid wash and the bead-bound proteins were eluted with Laemmli SDS sample buffer. Protein was visualized on 12 % SDS-PAGE and the E2 TAD to MBP DED-B ratio was determined by densitometric analysis of background corrected band intensity using the ImageJ software.

4.1.1.4 Co-immunoprecipitation assays – HEK 293 cells were maintained in DMEM media supplemented with 10% foetal bovine serum. Cells were seeded in 6-well plates and transfected with 1.5 µg of pEGFPc1 HPV18 E2 and 500 ng of pCDNA3.0 caspase-8 expressing plasmids using Lipofectamine 2000 (Life Technologies). After incubation for 24 h, the transfected cells were lysed in 1% CHAPS buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 1% CHAPS, 0.1% Triton-X 100, 5% sucrose, 1X protease cocktail inhibitor). The cell lysate was centrifuged at 10,000 rpm for 15 min at 4 °C and the resulting soluble cell lysate was pre-cleared using Protein G sepharose beads (GE Healthcare). The pre-cleared cell lysate was incubated with 1 µg of anti-E2 antibody (clone vW-20, Santa Cruz Biotechnology, Inc.)

at 4 °C overnight. 100 µl of Protein G sepharose beads were added to each sample, incubated for 2 h at 4 °C and then washed three times with 1% CHAPS buffer. Bead-bound proteins were eluted with Laemmli sample buffer. Eluted proteins and cell lysates were separated by SDS-PAGE and immunoblotted using antibodies to HA-tag (Sigma-Aldrich) and developed using enhanced chemiluminescence ECLprime Kit (GE Healthcare).

4.1.1.5 Confocal imaging – Post 24-30 h of transfection, GFP-tagged E2 and mCherry fused procaspase-8 were imaged using a confocal system with an Argon 488-nm and helium/neon 543-nm lasers. All the images were captured using Plan-Apochromat 63×1.4 NA objective in 12-bit format using Zeiss LSM 510 Meta confocal laser scanning microscope.

4.1.1.6 Molecular modeling and docking – Procaspase-8 DED (1-190) was modeled using homology modeling by *Modeller 9 v7* [139]. For this, the 1.4 Å resolution structure of viral FLIP (MC159) from Molluscum contagiosum virus (PDB id. 2BBR) was taken as template. Sequence for the DED was taken from the UniProt Knowledgebase (UniProt ID. Q14790). Using Modeller, 20 models of the protein were generated and then evaluated using the Discrete Optimized Protein Energy (DOPE) scores. The best model was chosen as the one with the lowest DOPE score. In order to assess the quality of the model, PROCHECK tool was used [140]. The homology modeled structure was then subjected to energy minimization for 2000 steps and equilibration using CHARMM27 force field in NAMD2.9 Molecular Dynamics Package (<http://www.ks.uiuc.edu/Research/namd/>). Particle Mesh Ewald was used to treat long range electrostatics and SHAKE algorithm was used to constrain hydrogen containing bonds. The system was equilibrated in water at 300 K and 1 atm pressure for 20 ns. The protein was then further simulated in water for 22 ns and the final structure at the end

of the simulation was then used for docking studies. The structure and simulations were visualized and analyzed using Visual Molecular Dynamics (VMD) software. Similarly, the protein sequence of HPV18 E2 TAD was retrieved (UniProt ID. P06790). Although the crystal structure of E2 protein TAD domain (PDB id. 1QQH) was available in the literature at resolution of 2.1 Å, the first 65 residues were missing [141]. Therefore, protein blast search was carried out against PDB database (<http://www.rcsb.org/pdb>) to get most suitable template for homology modeling of the complete TAD structure. The first 65 residues of HPV18 E2 TAD was modeled using HPV16 E2 TAD (PDB id. 1DTO) as a template [142]. It was followed by a brief equilibration (as done earlier) of the protein in water for 16 ns after addition of hydrogens. The final energy minimized structure was used for docking studies.

The HPV18 E2 TAD structure was docked on to procaspase-8 DED using HADDOCK [143]. To predict the best binding mode between the two proteins, the active residues were assumed to be the set of accessible residues from the two proteins which were evaluated using WHATIF [144]. The tool was allowed to define the passive residues automatically. A total of 1000 complexes were sampled for rigid body docking, followed by 200 for semi-flexible refinement and 200 for explicit solvent refinement. The final structures after explicit solvent refinement were clustered using a RMSD cut-off of 7.5 Å. The HADDOCK predicted complex of DED and HPV18 E2 TAD with the best score was then energy minimized for 2000 steps and equilibrated in water for 70 ns at 300 K and 1atm pressure in order to achieve further refinement of the docked complex. The complex was then simulated for 40 ns and the final structure of the complex at the end of the simulation was then used for further analyses. The binding interface residues were evaluated using the PDBsum generate server [145, 146] and Protein Interactions Calculator [147] with the default cut-offs.

4.1.1.7 CD and fluorescence spectroscopy – Far-UV CD scans were acquired using a JASCO J 815 spectropolarimeter (Jasco, Easton, MD, USA) using a quartz cell with 1-mm path length. The spectra were recorded for 10 μ M proteins at 25 °C between 190 and 260 nm using a scan speed of 20 nm/min for total 5 accumulations. The mean residue ellipticity was calculated as described in chapter-3. Fluorescence emission was measured for protein solutions (2 μ M) using a Fluorolog-3 spectrofluorometer (Horiba Scientific, Edison, NJ, USA) with 280 nm excitation followed by emission between 310 and 400 nm.

4.1.1.8 Cell death assays – IETD hydrolysis: E2 wild-type and mutant transfected cells were lysed in caspase assay buffer (20 mM HEPES pH 7.7, 100 mM NaCl, 1 mM DTT, 1 mM EDTA, 0.1% CHAPS, 0.5% NP-40, 10% sucrose). Caspase activity was measured in crude extracts containing 30 μ g proteins with 200 nM of Ac-IETD-AFC in 100 μ l of caspase assay buffer for 2 h at 37 °C. Ac-IETD-AFC hydrolysis was measured by monitoring emission at 510 nm with 405 nm excitation. For each experiment, cells expressing wild-type GFP-E2 incubated with pan caspase inhibitor (z-VAD-fmk) was kept as a control. Cell viability analyses by flow cytometry and DNA ladder assay were performed as described in chapter-3.

4.1.2 Results

4.1.2.1 Mapping the minimal binding region involved in HPV18 E2 – procaspase-8 interaction

It has been previously demonstrated that the amino-terminal transactivation domain (TAD) of HPV18 E2 binds directly to the prodomain of caspase-8 to induce apoptosis [8]. In order to identify the minimal binding domain for E2 – procaspase-8 interaction, we developed a co-expression system, owing to the difficult nature of DED proteins in solution, to perform

binding analysis. For this, we cloned the individual genes in two compatible plasmid sets with different origin of replication, antibiotic selection, inducer and affinity tag. HPV18 E2 TAD was cloned in an ara-operon plasmid (pACYC-vector) facilitating expression of His₆ tagged E2 TAD. Similarly, lac-operon constructs encoding MBP fusion procaspase-8 full length or DEDs were cloned in pMALc5-TEV empty vector. The constructs were co-expressed in *E.coli* BL21 (DE3) cells and pull-down assay with appropriate controls was performed as described in the *experimental section*. The pull-down of tandem DED or individual DEDs with E2 TAD showed that E2 binds very weakly to DED-AB or DED-A, while it strongly associates with DED-B (**Fig. 4.1.1**). This result suggests that HPV18 E2 is probably involved in heterotypic interaction with DED-B of procaspase-8.

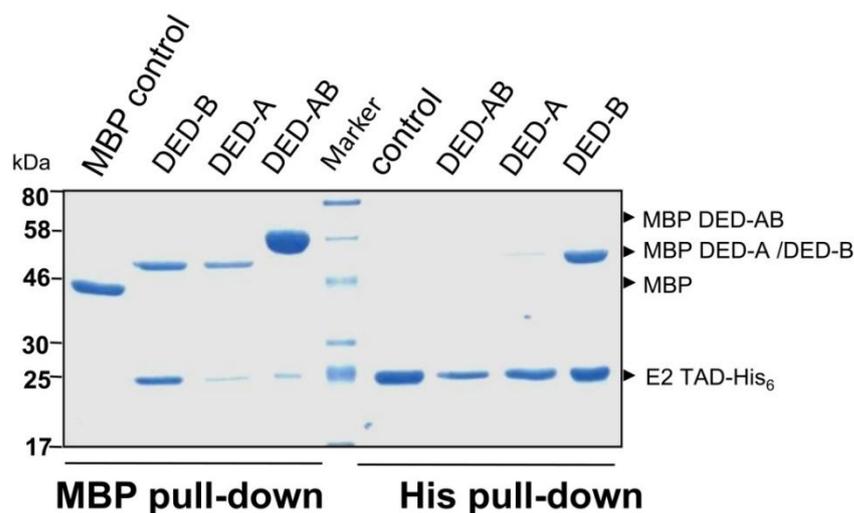


Figure 4.1.1 Pull-down assay of E2 TAD with procaspase-8 DED domains. MBP tagged DED-AB, DED-A or DED-B were tested for their binding with E2 TAD-His₆. Lanes 1-4 represent MBP pull-down using amylose resin while lanes 6-9 is for reverse His pull-down with Ni²⁺-IDA resin. In lane-1, MBP tag was checked for its binding to E2 TAD-His₆; and in lanes 2-4, MBP fused DED-B, DED-A and DED-AB acted as baits to test their complex formation with E2 TAD. In lanes 6-9, E2 TAD-His₆ acted as a bait to monitor the binding of MBP, MBP-tagged DED-AB, DED-A or DED-B respectively. The image is of 12% SDS-PAGE coomassie stained gel. It is interesting to note that DED-B interacts with E2 while DED-AB, which although contains domain-B, does not show interaction.

In order to map the minimal binding region in DED-B, we generated 12 deletion constructs comprising different combination of six α -helices of DED-B represented in **Fig. 4.1.2**, and performed the pull-down assay. We observed that deletion of helices α 2 and α 5

disrupted the interaction with E2 TAD (Fig. 4.1.3), indicating that plausibly these helices are important for mediating E2 – procaspase-8 interaction.

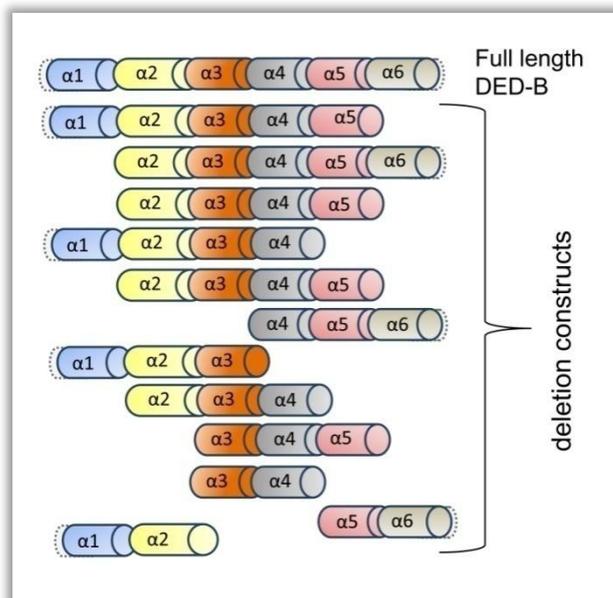


Figure 4.1.2 Schematic representation of the deletion constructs for procaspase-8 DED-B. The individual helices are shown in different colors.

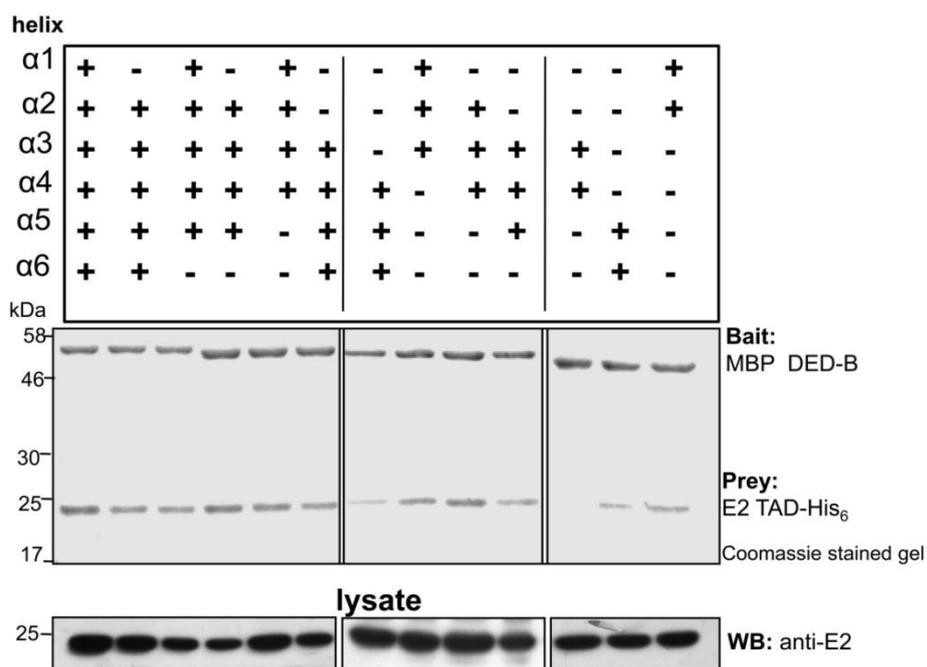


Figure 4.1.3 Pull-down assay of E2 TAD with procaspase-8 DED-B deletion constructs. MBP pull-down was performed for 12 deletion constructs of MBP DED-B as the bait to test their interaction with E2 TAD-His₆. 20 μ g of the total cell lysate was immunoblotted with anti-E2 antibody to confirm the presence of E2 protein in all the samples.

4.1.2.2 *In silico* prediction of E2 – procaspase-8 binding interface

The crystal structure of viral FLICE-like inhibitory protein (vFLIP) from *Molluscum contagiosum* 159 (MC159) provides an excellent template to study the structure function relationship of tandem DED containing proteins [66]. The procaspase-8 tandem DED is 23% identical in sequence to the MC159 vFLIP. Therefore, a homology model of caspase-8 DED was constructed using MC159 vFLIP structure (PDB entry 2BBR) as a template (**Fig. 4.1.4A**). The final energy minimized modeled structure of caspase 8 DED had 0.6% residues in the disallowed region of the Ramachandran plot as evaluated using PROCHECK tool. Each of the two DED domains showed the presence of classic death-fold architecture with six α helices and both the domains connected by a linker containing a single helix (**Fig. 4.1.4B**).

The HPV18 E2 TAD structure was modeled for the first 65 missing residues in crystal structure (**Fig. 4.1.5A**) using HPV16 E2 TAD as the template. It has an overall sequence identity of 47% with HPV16 and superposition of (66-201) residues demonstrates a root mean square difference of less than 1.0 Å. The simulated structure of HPV18 E2 TAD contains three helices and a β -strand containing domain connected by two consecutive helical turns (**Fig. 4.1.5B**). Both these two modeled structures were then subjected to docking for identification of the binding interface.

The inter-molecular interface between DED and E2 TAD docked complex had 1667\AA^2 of extensively buried surface area and notably involved residues F122, L123, D158, K161, R162, Q166 in DED-B and Q35, Q38, R41, W42, F48, D52, I77, Q80, M81, Q84 in HPV18 E2 TAD (**Fig. 4.1.6**). **Table 4.1.1** provides a detailed list of the interactions. Strikingly, 80% of the interactions in DED involved residues from helices $\alpha 2$ and $\alpha 5$ of DED-B, which very well corroborates with our deletion mutagenesis studies. For E2 TAD, 60% of the interactions majorly involved residues from helix $\alpha 2$ highlighting the importance of this helix

in E2-mediated apoptosis. This observation is in support of a previous finding where deletion of helix $\alpha 2$ from E2 showed diminished apoptosis [1].

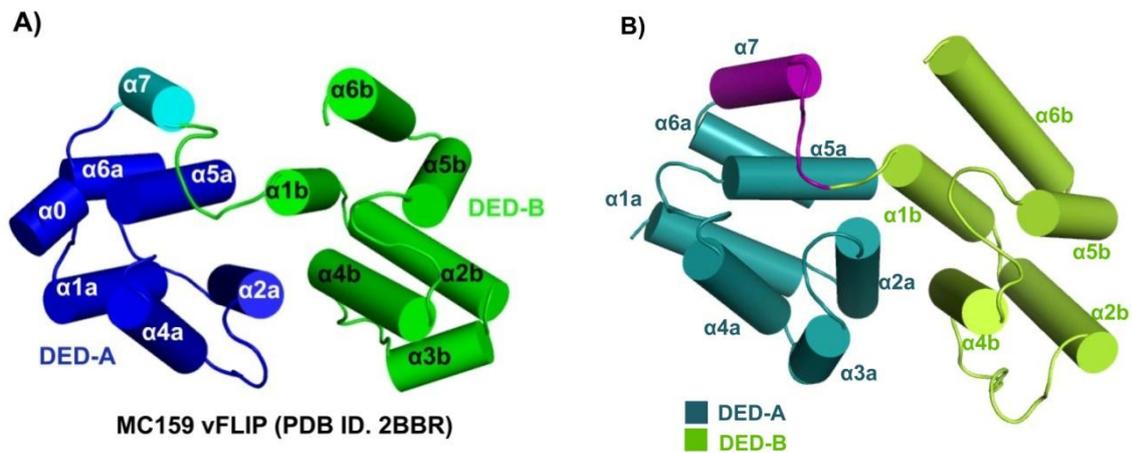


Figure 4.1.4 Molecular modeling of procaspase-8 tandem DED. A) Cartoon representation of MC159 vFLIP X-ray structure (PDB id. 2BBR). The domains DED-A and DED-B are shown in blue and green colours respectively. B) Homology model of human procaspase-8 DED-AB (residues 1-185).

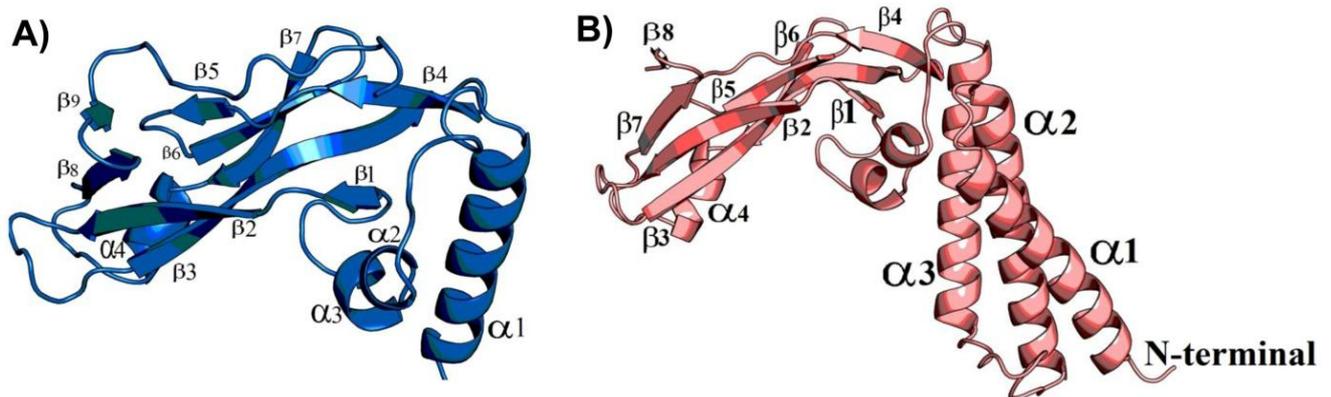


Figure 4.1.5 Molecular modeling of HPV18 E2 TAD. A) Crystal structure of HPV18 E2 TAD (residues 66-208), PDB id. 1QQH. B) Model of HPV18 E2 TAD built on the template 1QQH. The missing residues 1-65 were homology modeled using HPV16 E2 TAD (PDB id. 1D7O).

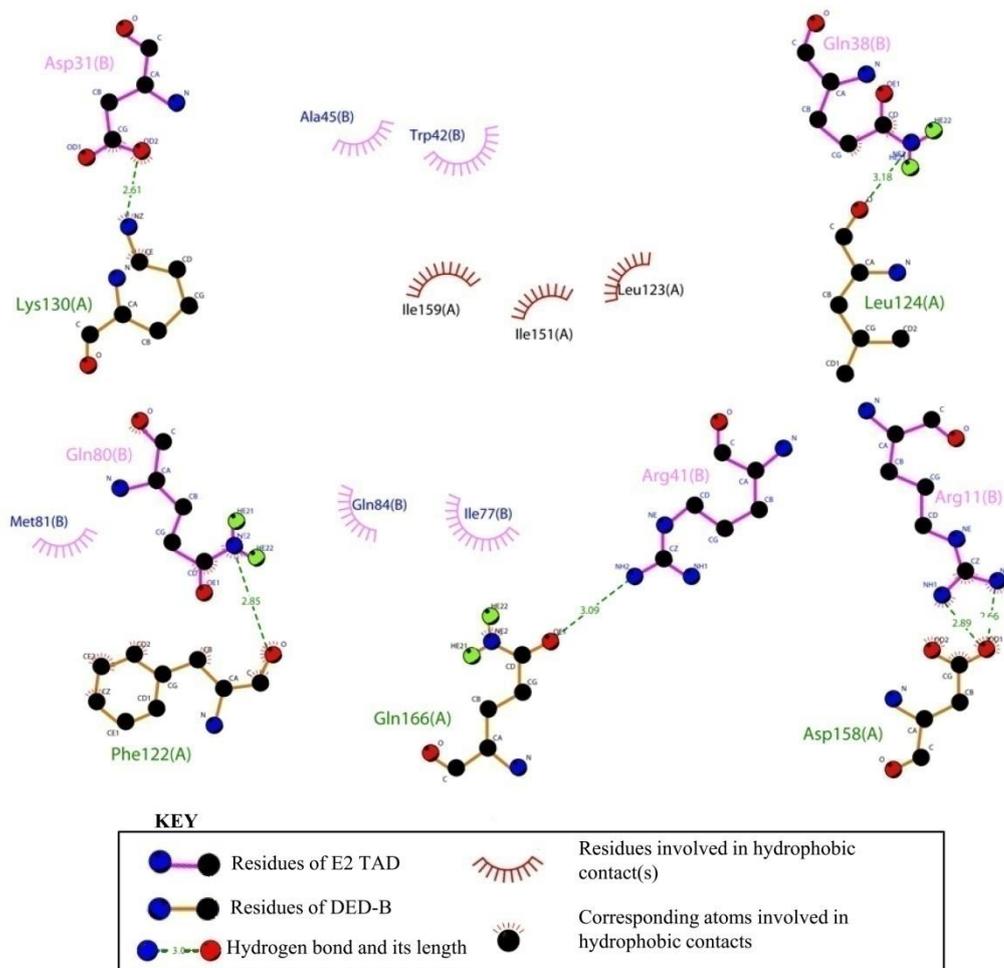


Figure 4.1.6 Interaction across HPV18 E2 TAD – procaspase-8 DED interface. Dimplot of the docked E2 – procaspase-8 complex. Residues from E2 TAD and procaspase-8 DED-B are shown in pink and orange colors respectively.

4.1.2.3 Identification of procaspase specific binding site on E2

To identify the key residues involved in the interaction, we introduced series of mutations in E2 TAD followed by co-expression and pull-down assay of mutant E2 with wild-type DED-B. This analysis identified R41, W42 and F48 as potential participants in the interaction with procaspase 8. They reside on $\alpha 2$ helix of E2 and mutations of these residues together were found to significantly affect binding with procaspase-8 DED-B (**Fig. 4.1.7A**). We also tested the single mutants of these residues and observed that neither of these point mutations alone inhibited the complex formation (**Fig. 4.1.7B**). Since the biochemical analysis of DED complex is limited by its poor solubility, we carried out a dilution experiment with the

isolated E2 – procaspase-8 DED-B complex as previously demonstrated for characterization of Fas – FADD death domain complex [52]. It was observed that E2 – DED-B wild-type complex showed cooperative dissociation of the complex below concentration as low as 500 nM, based upon plot derived from quantitative SDS–PAGE analysis from various E2 – DED complex dilutions (**Fig. 4.1.8**), suggesting high affinity between the two proteins. Also the complex was found to remain stable in low to high ionic strength buffers. In contrast, the E2 TAD triple mutant showed remarkably reduced binding and could not form a stable complex. The mutations disrupt the function of any protein by two distinct possibilities: either by destabilizing the structure or affect a functional surface that is essential for binding to other protein(s). Notably, these potential residues in E2 TAD are solvent exposed (**Table 4.1.2**), and are thus unlikely to affect the structure. To confirm this, secondary and tertiary structural analyses was performed using Far-UV CD and fluorescence spectroscopic studies respectively (**Fig. 4.1.9A and B**), which showed that all the TAD mutants have overall structural conformation similar to the wild-type protein.

To further analyze the relationship between E2 and caspase-8, and the effect of substituting the interface residues in E2, we carried out co-transfection in HEK 293 cells. It has been previously established that ectopic expression of tagged/untagged HPV18 E2 or caspase-8 separately show nucleo-cytoplasmic distribution and diffused cytoplasmic localization respectively [42]. To study the relationship, we ectopically co-expressed GFP-E2 and catalytically inactive mCherry fused caspase-8 C360A mutant which would prevent induction of cell death upon interaction. Post 24 h of transfection, live cell images were captured by laser confocal microscope to visualize their localization. Co-expression of caspase-8 with E2 resulted in redistribution of the two proteins into co-localized cytoplasmic punctate structures (**Fig. 4.1.10B**). However, upon co-expression of E2 triple mutant and caspase-8, we did not observe any co-localization, and that caspase-8 reverted to diffuse

distribution while E2 showed a nucleo-cytoplasmic localization (**Fig. 4.1.10C**). The interaction was further confirmed by co-immunoprecipitation experiments. Immunoblot analysis of the co-immunoprecipitated proteins showed that E2 wild-type associates with procaspase-8 while the triple mutant fails to do so (**Fig. 4.1.11**). We, therefore, conclude that interaction of E2 with caspase-8 involves residues R41, W42 and F48 of HPV18 E2 TAD.

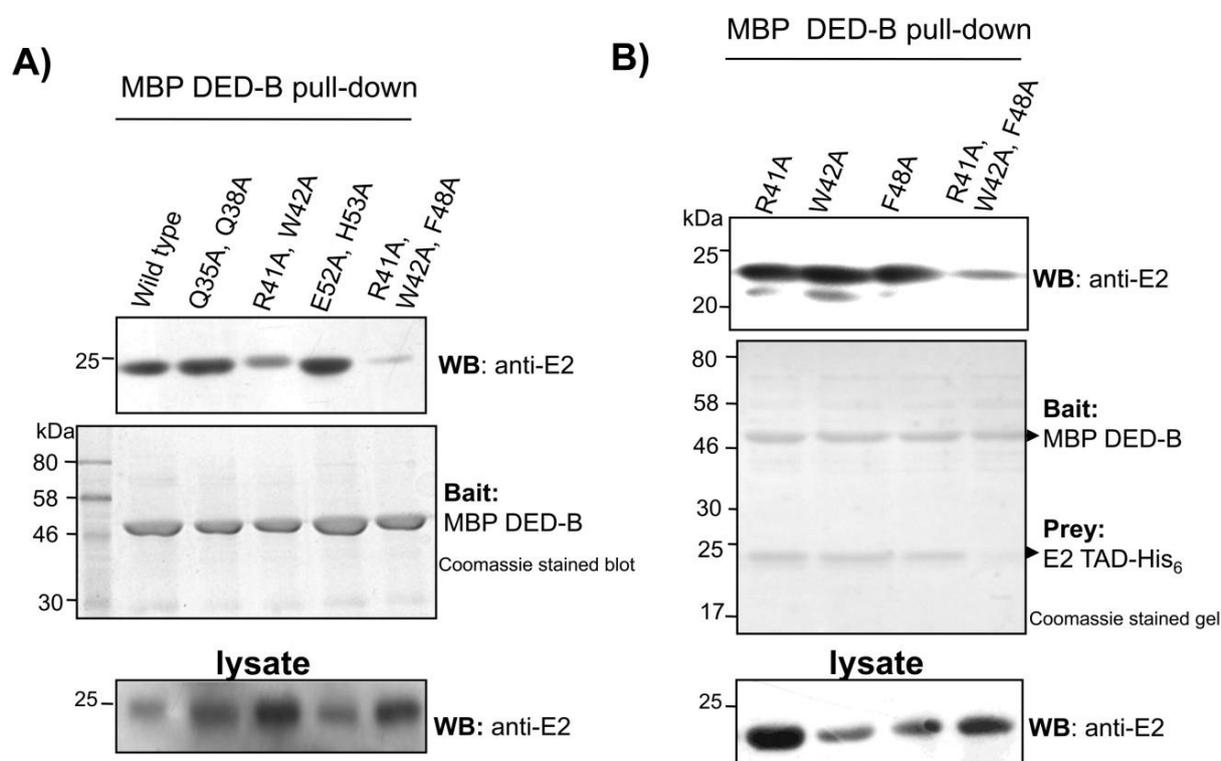


Figure 4.1.7 MBP pull-down assay for E2 wild-type and mutants with procaspase-8 DED-B. A) E2 TAD wild-type and mutants as indicated were tested for their binding to MBP-tagged DED-B (bait). B) Western blotting analysis of E2 TAD single mutants (R41A, W42A or F48A) and triple mutants (prey) with MBP DED-B as bait. 20 μ g of the total cell lysate was immunoblotted with anti-E2 to confirm the presence of E2 TAD in all the samples.

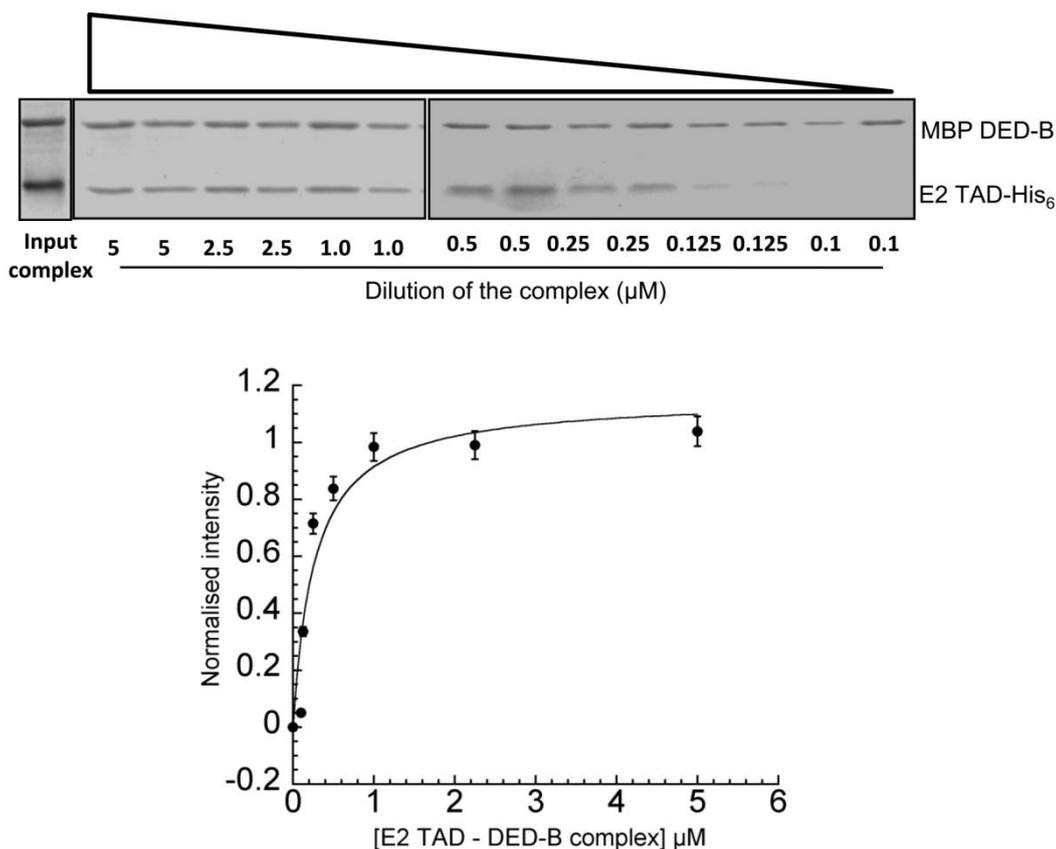


Figure 4.1.8 Dilution experiment with the isolated E2 TAD – DED-B complex. Top panel-The complex diluted to different final concentrations as indicated were analyzed on 12% SDS-PAGE (representative gel). The lower panel shows a plot derived from quantitative SDS-PAGE analysis of background corrected band intensity measured as E2 TAD to DED-B ratio for various dilutions of the complex. The measured intensity is represented as normalized intensity for the data obtained from three independent complex preparations and their dilutions. The error bars show the standard error.

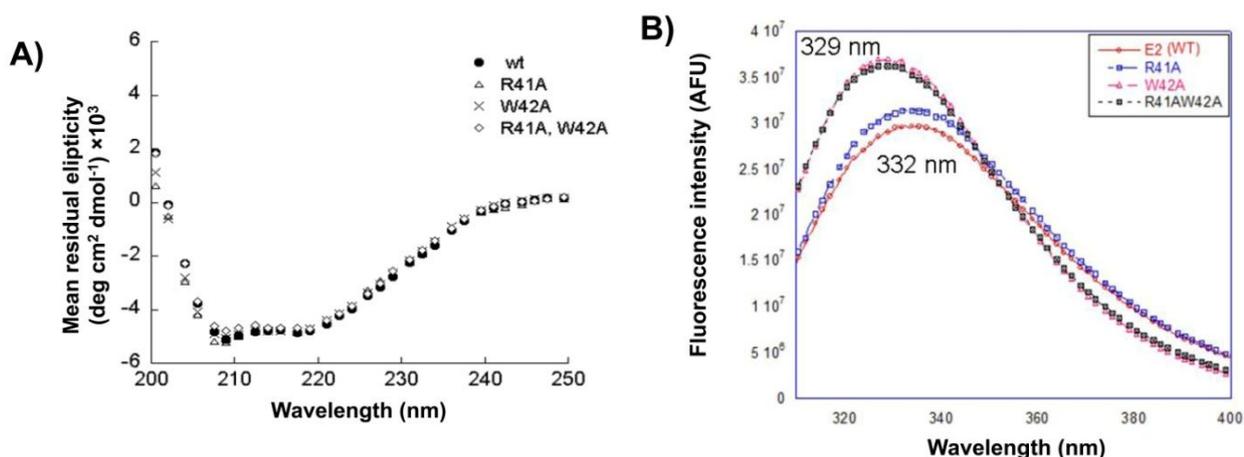


Figure 4.1.9 Secondary and tertiary structural analyses of E2 TAD wild-type and mutant proteins. A) Far ultraviolet circular dichroism spectra of E2 wild-type and mutant proteins. The measurements were made of at 25 °C for 10 μM of protein solution in buffer 10 mM NaH₂PO₄/Na₂HPO₄ pH 8.0, 100 mM NaCl. Contributions to the spectra by the buffer were subtracted using control scans. All the curves are represented as mean residual ellipticity in unit deg cm² mol⁻¹.

The CD spectra reveal that the mutants have similar secondary structure characteristics as the wild-type protein. B) Fluorescence emission scans recorded for 2 μ M of wild-type and mutant E2 TAD using an excitation of 295 nm and emission in the range of 310-400 nm. Note that the emission maximum blue shifted by 3 nm in case of the single and double mutants W42A and R41A, W42A respectively compared to wild-type and R41A mutant is most likely due to the substitution of the surface exposed W42.

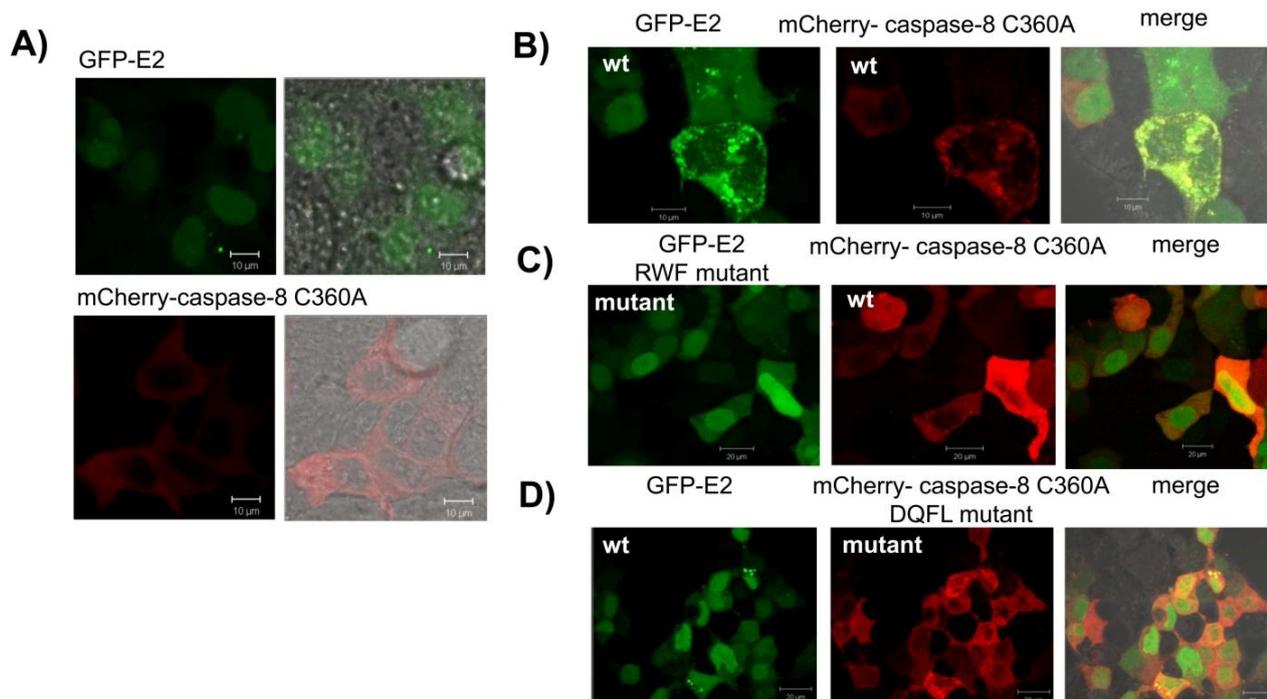


Figure 4.1.10 Expression and localization of procaspase-8 and HPV18 E2 in transfected cells. A) HEK 293 cells transfected with GFP tagged E2 (upper panel) and mCherry fused catalytically inactive mutant of procaspase-8 (lower panel). GFP-E2 and mCherry-procaspase-8 show a nucleo-cytoplasmic and diffused cytoplasmic distribution respectively. In panel B and C, cells were co-transfected with catalytically inactive wild-type procaspase-8 and wild-type (B) and R41A, W42A, F48A (RWF) triple mutant (C) GFP-E2 constructs. D) Cells co-transfected with wild-type E2 and the quadruple mutant DQFL (D158A, Q166A, F122A, L123A). Live cell images were acquired post 24 h of transfection using Zeiss LSM 510 confocal microscope, and representative images for each transfection is shown along with the scale bar (Original magnification x630 with 2X optical zoom). Note that cells expressing both wild-type proteins show redistribution of the two proteins into co-localized cytoplasmic punctate structures. However, in case of cells co-expressing wild-type and mutant proteins, caspase-8 reverts to a diffuse distribution while E2 shows nucleo-cytoplasmic localization.

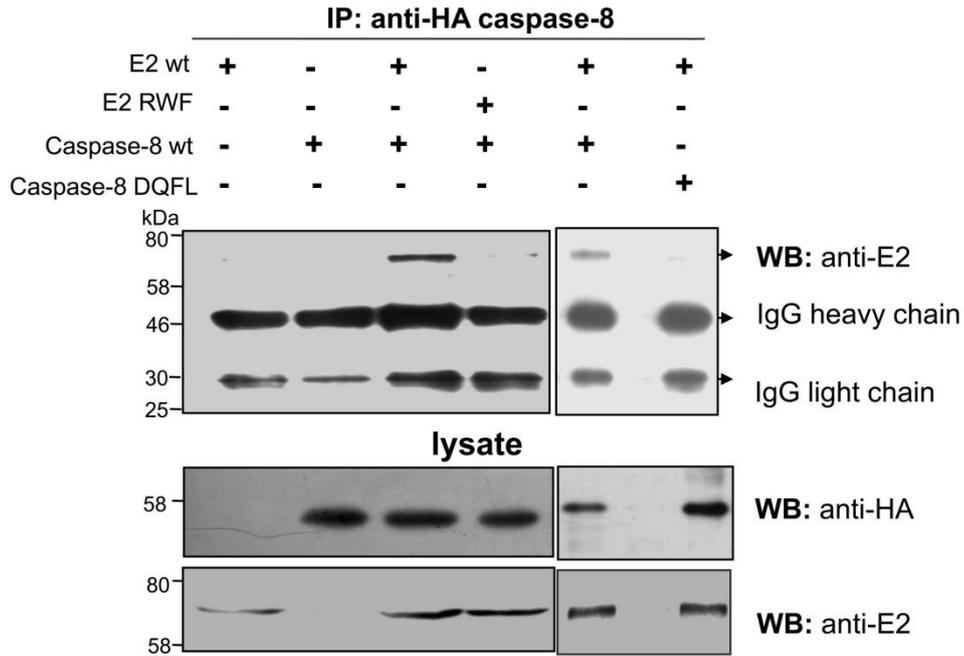


Figure 4.1.11 Immunoprecipitation of E2/procaspase-8 complex. Extracts from HEK 293 cells transfected as described were immunoprecipitated using an anti-HA antibody followed by immunoblotting using anti-E2 antibody. 20 μ g of the lysate were immunoblotted with anti-HA and anti-E2 antibodies.

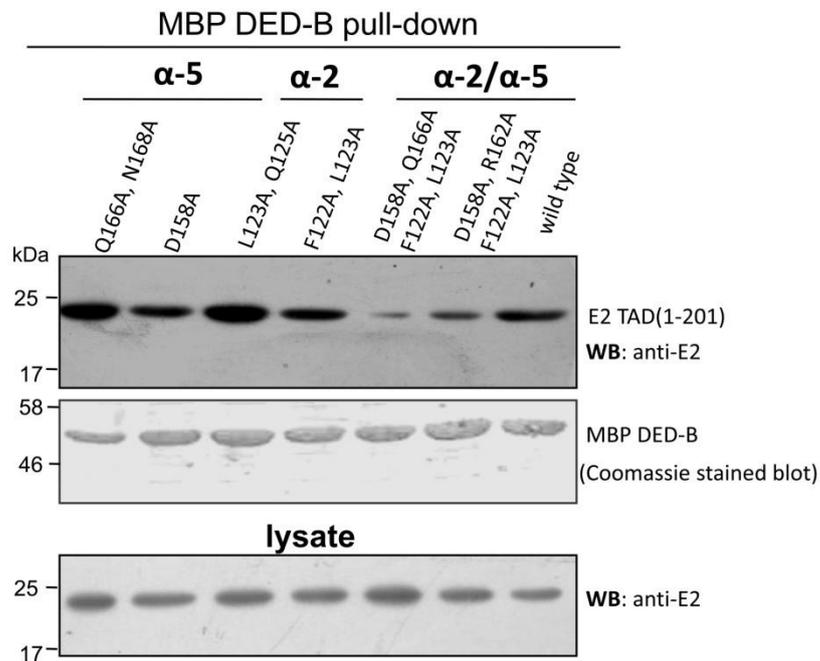


Figure 4.1.12 MBP pull-down assay for wild-type E2 TAD and procaspase-8 DED-B mutants. A) MBP-tagged procaspase-8 DED mutants (bait) as indicated were tested for their binding to E2 TAD-His₆ (bait) by probing with anti-E2. 20 μ g of the total lysate was immunoblotted with anti-E2 to confirm the presence of the E2 TAD in all the samples.

4.1.2.4 Identification of E2 specific binding site on procaspase-8

A closer look into the available DED family complex interface reveals large interaction surface [50], which often lacks the presence of defined and centred interaction sites as observed in case of regulatory complexes [148-150]. From our docking studies, we observed a wide surface area of the interface involving different helices of DED-B. Based on the docking and deletion mutagenesis studies, we identified that helices $\alpha 2$ and $\alpha 5$ of procaspase-8 DED-B might be important in mediating E2-procaspase 8 interaction.

To pin-point the key residues involved in the interaction, series of alanine substitutions were carried out individually and in combination on these two helices of DED-B. Using these mutants, we performed the pull-down assay as described earlier. We observed that substitution of residues F122, L123 (helix $\alpha 2b$) and D158, Q166 (helix $\alpha 5b$) in combination significantly affected the binding of procaspase-8 with E2 (**Fig. 4.1.12**). To further substantiate these findings, we performed a co-immunoprecipitation assay. The HA-tagged caspase-8 wild-type or mutant and GFP-E2 were co-transfected into HEK 293 and the cell lysate was immunoprecipitated using anti-HA. The immunoprecipitates were then examined by western blotting using anti-E2 antibody (**Fig. 4.1.11**). Our results of co-immunoprecipitation confirmed the role of these residues as potential participants in mediating E2 – procaspase-8 interaction. Furthermore, co-localization studies were carried out with ectopic expression of GFP-E2 and mCherry tagged procaspase-8 quadruple mutant (**Fig. 4.1.10D**). It was observed that caspase-8 did not redistribute into the typical cytoplasmic punctate structure in presence of E2, as seen in case of wild-type complex. Therefore, collectively these results indicate that binding of E2 – procaspase-8 involves interaction between helix $\alpha 2$ of E2 TAD with $\alpha 2/\alpha 5$ helices of procaspase-8 DED-B.

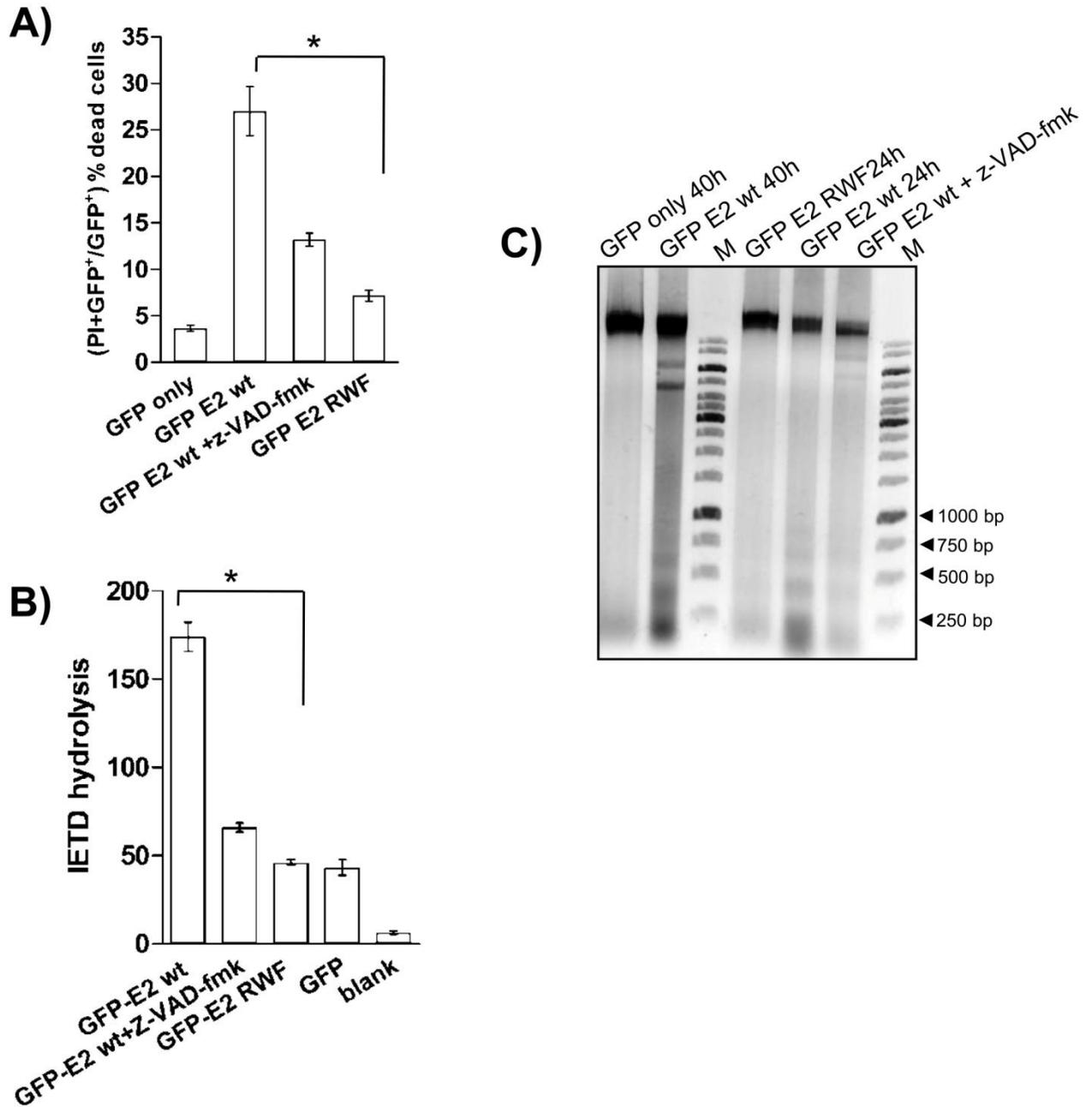


Figure 4.1.13 E2-induced cell death analysis. A) Viability of HEK 293 cells transfected with the expression plasmids as indicated, determined post 22 h of transfection by propidium iodide uptake. B) Caspase-8 activity was measured for the crude extracts isolated from the cells transfected with the plasmids as indicated, using fluorogenic Ac-IETD-AFC substrate. The bars represent the mean and the error bars the standard deviation. Statistical analysis using student's t-test, * $p < 0.05$ (A) and $p < 0.01$ (B). C) Agarose gel electrophoresis of apoptotic DNA fragments. Total DNA extracted from cells was loaded on a 2% agarose gel. 180 bp DNA ladder was observed as a typical feature for apoptosis in case of wild-type E2 expressing cells while it was significantly reduced in mutant cells (M stands for 1 kb DNA marker). For each experiment, cells expressing wild-type E2 incubated with pan-caspase inhibitor (z-VAD-fmk) was kept as control to monitor the effect on caspase activation.

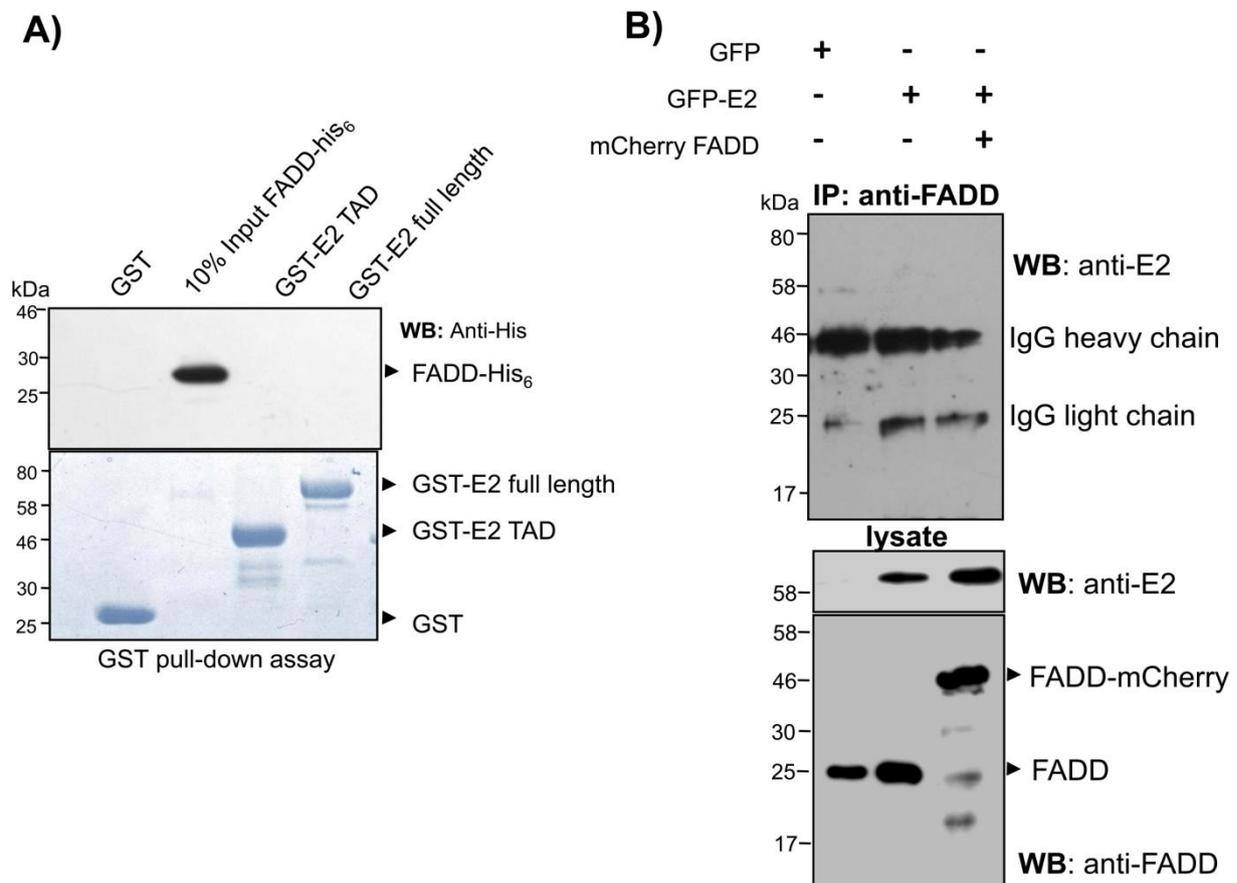


Figure 4.1.14 E2-FADD interaction analysis. A) Upper panel, western blot analysis of GST pull-down assay using HPV18 E2 full length and TAD proteins as bait and wild-type FADD-His₆ as prey. Lower panel shows the coomassie stained gel for the recombinant GST and fusion proteins. B) Immunoprecipitation of FADD. Extracts from HEK 293 cells co-transfected as described were immunoprecipitated using an anti-FADD antibody followed by immunoblotting using anti-E2. 20 µg of the lysate were immunoblotted with anti-FADD and anti-E2 antibodies.

4.1.2.5 Monitoring the effect of interface E2 mutant on cell death

It has previously been documented that E2 proteins of high risk HPVs induce apoptosis in all human cell types, including transformed as well as in primary epithelial cells [7, 48, 151]. To determine whether loss of interaction between E2 and procaspase-8 affected the pro-

apoptotic activity of E2, cell death assays were performed in presence of wild-type or mutant E2 protein in HEK 293 cells. When ectopically expressed, GFP-E2 wild-type induced cell death as measured by propidium iodide uptake (**Fig. 4.1.13A**). However, expression of GFP-E2 α -2 helix triple mutant showed a significant reduction in percent cell death. Overexpression of gfp *per se* did not show any considerable cell death. In addition, the presence of pan-caspase inhibitor Z-VAD-fmk reduced the level of E2-induced cell death as reported previously [7]. These results highlight the requirement of interaction between E2 and procaspase-8 to bring about activation of caspase-8 and hence cell death. To measure caspase-8 activation, we determined the rate of hydrolysis of caspase-8 specific fluorogenic substrate IETD-AFC as described in the *experimental section*. We found that the IETDase activity was readily detectable in extracts from GFP-E2 wild-type expressing cells but not in the mutant expressing cells (**Fig. 4.1.13B**). We further confirmed these results by performing DNA fragmentation assay (**Fig. 4.1.13C**), a marker for apoptosis. DNA ladder was observed as a typical feature for apoptosis in case of wild-type E2 expressing cells while it was significantly reduced in cells expressing mutant protein.

4.1.2.6 E2 – FADD interaction analysis

To confirm whether the interaction between HPV18 E2 and procaspase-8 is independent of adaptor protein FADD, *in vitro* GST pull-down assay was performed with recombinant FADD and wild-type full length and TAD of E2 (**Fig. 4.1.14A**). The pull-down assay showed that FADD does not interact with either the full length E2 or TAD in agreement with the earlier reports [1, 41]. We also validated this by co-immunoprecipitation of ectopically expressed GFP-E2 and mCherry fused FADD or with endogenous FADD (**Fig. 4.1.14B**). These results confirm that E2 – procaspase-8 interaction is indeed independent of the adaptor protein FADD.

Table 4.1.1 List of interactions across E2 TAD – procaspase-8 DED-B interface evaluated using the *Protein Interactions Calculator* with the default cut-offs [147].

Type of interaction	E2 TAD	DED-B	Distance (Å)
Hydrogen bonds			
Main chain- side chain	O Asp31 ($\alpha 2$)	O ^{$\epsilon 1$} Glu126 ($\alpha 2b$)	3.04
	O Ala45 ($\alpha 2$)	NH1 Arg162 ($\alpha 5b$)	3.07
	O Arg41 ($\alpha 2$)	N ^{$\epsilon 2$} Gln166 ($\alpha 5b$)	3.01
	O ^{$\epsilon 1$} Gln38 ($\alpha 2$)	N Glu126 ($\alpha 2b$)	3.06
	NH1 Arg41 ($\alpha 2$)	O Ser119 ($\alpha 2b$)	3.36
	O ^{$\epsilon 1$} Gln80 ($\alpha 3$)	O Phe122 ($\alpha 2b$)	3.00
Side chain-side chain	N ^{$\epsilon 2$} Gln35 ($\alpha 2$)	O ^{$\epsilon 2$} Gln127 ($\alpha 2b$)	3.56
	O ^{$\epsilon 2$} Glu52 ($\alpha 2$)	N ^{ζ} Lys161 ($\alpha 5b$)	2.67
	NH2 Arg41 ($\alpha 2$)	O ^{$\epsilon 1$} Gln166 ($\alpha 5b$)	2.92
	N ^{$\epsilon 1$} Trp42 ($\alpha 2$)	O ^{$\delta 1$} Glu158 ($\alpha 5b$)	3.13
	O ^{$\epsilon 1$} Gln84 ($\alpha 3$)	O ^{γ} Ser129 ($\alpha 2b$)	2.96
Salt bridges (within 4 Å)			
	Asp31 ($\alpha 2$)	Lys130 ($\alpha 2b$)	
	Arg11 ($\alpha 1$)	Asp158 ($\alpha 5b$)	
	Glu52 ($\alpha 2$)	Lys161 ($\alpha 5b$)	
Hydrophobic interactions (within 5 Å)			
	Met81 ($\alpha 3$)	Phe122 ($\alpha 2b$)	
	Trp42 ($\alpha 2$)	Leu124 ($\alpha 2b$)	
	Trp42 ($\alpha 2$)	Ile151 ($\alpha 5b$)	
	Trp42 ($\alpha 2$)	Ile159 ($\alpha 5b$)	
	Phe48 ($\alpha 2$)	Ala165 ($\alpha 5b$)	
	Ile77 ($\alpha 3$)	Phe122 ($\alpha 2b$)	

Table 4.1.2 List of binding interface residue areas [Å^2] (hydrophilic, hydrophobic, and total) calculated with the POPS server [152].

Residue	Hydrophobicity (Å^2)	Hydrophilicity (Å^2)	Total (Å^2)	% SASA ^a
Gln35	25.25	70.98	96.22	0.46
Gln38	8.47	41.66	50.13	0.24
Arg41	9.92	61.59	71.51	0.26
Trp42	76.07	11.22	87.29	0.45
Phe48	43.79	1.80	45.59	0.27
Glu52	20.55	63.82	84.37	0.42
His53	70.45	57.73	128.1	0.70

^aSolvent Accessibility Surface Area

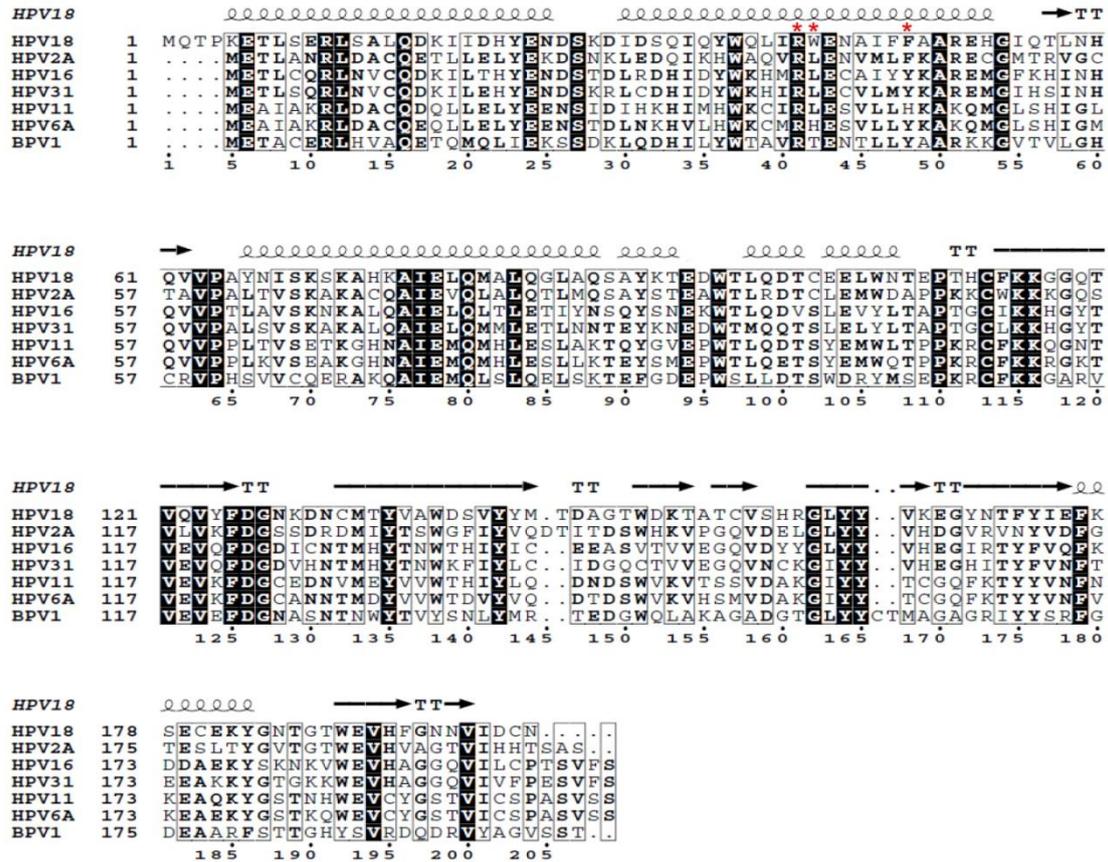


Figure 4.1.15 Sequence alignment of E2 proteins from different papillomavirus types and structure-based sequence alignment with HPV18 generated using ESPript [153]. The asterisk in red highlights the residue critical for E2 – procaspase-8 interaction. BPV – Bovine papillomavirus, HPV – Human papillomavirus. HPV16, 18, 31 are high risk while HPV2a, 6a, 11 belong to low risk types.

4.1.3 Discussion

Despite numerous reports on high risk HPV E2 induced apoptosis, little is known about the structural basis and protein-protein interactions involved in mediating this function. Earlier studies have reported that apoptosis induced by HPV18 E2 is initiated by extrinsic cell death pathway by direct interaction of its amino-terminal domain with procaspase-8 [1]. Here, we dissected the binding interface of E2 – procaspase-8 complex with an aim at understanding the driving force for this interaction and the underlying mechanism by which HPV E2 induces cell death.

Interaction analysis using series of mutants showed that the residues from $\alpha 2$ helix of E2 TAD mediate direct non-homotypic interaction with $\alpha 2/\alpha 5$ helices of procaspase-8 DED-B. The interface is largely formed by several non-bonded contacts with no single type of chemical interaction being predominant. The interaction surface involves residues that spread out over an extended range of primary amino acid sequence. This may probably explain why we found it difficult to get abrogation of interaction by simple deletion or point mutagenesis. The dilution experiment highlighted that the complex is highly stable even at lower protein concentrations, and to a variety of buffer conditions including low to high ionic strength.

Alignment of E2 protein sequences from different papillomavirus types shows that residue R41 is absolutely conserved in all the variants while W42 and F48 are conserved in terms of their hydrophobicity (**Fig. 4.1.15**). A recent study on HPV16 induced apoptosis demonstrated that E2 directly interacts with cFLIP, a DED-containing decoy protein which is a protease-deficient caspase-8 homolog [49, 130]. Furthermore, the ability of E2 to induce apoptosis is linked to their different sub-cellular localization rather than to inherent properties of the protein [42]. Low risk HPV E2 remains strictly nuclear whereas high risk HPV E2 shows a nucleo-cytoplasmic distribution. These differences arise due to the presence of dominant functional nuclear localization signal (NLS) in the hinge domain of the low risk E2 proteins while it is absent in the high risk E2. In addition, HPV18 E2 contains a nuclear export signal (NES) in its amino-terminal domain that is responsible for active export of the protein from the nucleus to the cytoplasm. A cytoplasmic NLS mutant of low risk HPV11 E2 is reported to induce apoptosis similar to that of the high risk E2 proteins [42]. Moreover, the pro-apoptotic activity of E2 is shown to be independent of all its other virus associated functions [41, 151]. Altogether, these informations raise an interesting possibility of the intrinsic potential of E2 proteins from different virus types to bind DED-containing proteins via the identified conserved binding surface, provided its localization is in cytoplasm, and

thereby promote apoptosis. However, further studies with E2 proteins from low and high risk viruses need to be carried out to test this hypothesis.

Co-expression of E2 with caspase-8 resulted in colocalization of the proteins in cytoplasmic punctate structures. These structures share some features similar to the proapoptotic 'death effector filaments' formed due to the high level expression of certain DED-containing proteins such as caspase-8 DEDs or FADD [154]. Importantly, caspase-8 activation has been proposed to depend on its oligomerization induced by death receptor/FADD ligation or increase in their local concentration upon ectopic expression [155]. Oligomerization is believed to initiate proximity induced auto-cleavage and release of active caspase-8 [133, 156, 157]. Consistent with the previous reports [8], we found that E2 binds caspase-8 without any involvement of the adaptor protein FADD. Thus, these punctate structures may represent oligomerization of caspase-8 induced upon interaction with E2, resulting into caspase activation and hence apoptosis.

Based on these evidences, we speculate that E2 influences the external cell death pathway by acting as an adaptor protein which recruits procaspase-8 and enhances its capability to oligomerize thus promoting apoptosis in a novel adaptor independent manner. In order to gain further insight into this novel mechanism of caspase activation, it would be important to understand the basis of activation for the classical FADD-mediated pathway and establish a relationship between these two mechanisms. It could be possible that FADD-mediated apoptosis can also occur simultaneously provided E2 does not compete for the same binding sites on procaspase-8 or it may be mutually exclusive. Also, the molecular basis of oligomerization, a pre-requisite for caspase-8 activation, needed investigation. In addition to this, it would be intriguing to identify how E2 adopts a death-fold behavior and act as a potent binding partner of procaspase-8. These issues have been addressed in details in the next two sub-sections of this chapter.

CHAPTER-4.2

4.2 Molecular details of DED assembly and its role in death-fold mediated caspase-8 activation

4.2.1 Introduction

Caspases, a family of cysteinyl proteases, which initiate and execute apoptosis, rely on the events that lead to its activation, albeit in different ways depending on the signaling pathway. Activation of the apical caspases is achieved through their oligomerization driven by the death-fold domain proteins to form the death-signalling complexes [133, 156, 157]. Assembly of the death complex provides the upstream caspases with a sufficient degree of oligomerization to allow proximity induced self-activation and initiation of caspase cascade. From a cell signaling perspective, these strings of events need to be tightly regulated, and the binding partners do not (and should not) associate in the absence of stimuli.

Classical apoptotic cascades follow two distinct pathways, the intrinsic pathway which originates in the mitochondria, while the extrinsic pathway triggered by ligation of cell-surface “death receptors” followed by formation of multiprotein death-inducing signaling complex (DISC). The DISC comprises oligomerized death receptors such as Fas, the adaptor protein FADD, procaspase-8, and cFLIPs. In case of Fas-mediated signaling, extracellular Fas ligand binds to its receptor leading to intracellular recruitment of FADD via homotypic interactions between the death domain (DD) of receptor and FADD. Engagement with the receptor exposes the N-terminal death effector domain (DED) of FADD, which subsequently interacts with the tandem DEDs of procaspase-8 leading to its oligomerization, proteolytic cleavage and activation [50]. Similar to the other oligomeric signaling platforms, the DISC acts as a control switch, which remains in the ‘off’ position in the absence of stimulus followed by clustering of its members to form active (‘on’ position) oligomeric platform.

In recent years, new insights into the DISC assembly and caspase activation, especially at the level of DED-containing proteins are emerging [155, 158, 159]. The caspase-8 activation has been proposed to depend on the ability to form DED-chain assembly induced by death receptor/FADD ligation and/or increase in their local concentration [155]. The ectopically expressed FADD spontaneously self-oligomerizes via its DED to form structures called ‘death effector filaments’ which are able to recruit and activate caspase-8 inducing apoptosis independent of death receptors [154]. Although full length procaspase-8 contains DED it does not self-associate, but strictly depends on the homotypic interaction with FADD to oligomerize and self-activate [67, 136]. Thus, interaction between FADD and caspase-8 through the conserved ‘DED’ and the ability to form a chain seems to be a crucial step in caspase activation and hence cell death.

Despite remarkable structural details on death-fold domain subfamily members and their complexes, no multivalent complexes involving DEDs have been identified so far. This is in part due to the fact that DEDs have a strong tendency to aggregate and exhibit poor solution behavior, making biochemical and biophysical characterization an arduous task [64, 160]. Although few biochemical studies have provided information on the binding interface of FADD – caspase-8 complex [67], there is considerable lack of understanding on the mechanism of DED assembly and procaspase-8 activation. Structural studies of the death-fold domain complexes, such as Fas/FADD DD [52, 161], RAIDD/PIDD DD [59], have shed some light on the mechanistic basis of the formation of these multimeric assemblies, which reside in their ability to form three distinct types of asymmetric interactions. Considering the similar role of other death-fold domain proteins in formation of intricate higher-order complexes, does FADD – caspase-8 DED complex employ an analogous binding mechanism remains elusive.

Here, we present an intriguing insight into the molecular basis of DED-chain formation and define the surface for the physical interaction between FADD and procaspase-8.

4.2.2 Experimental Procedures

4.2.2.1 Expression plasmids and antibodies – GFP-E2, mCherry procaspase-8 C360A, mCherry FADD (aa 1-208) and pMALc5-TEVcaspase-8 DED-AB expression plasmids are described previously in chapter-4.1. The caspase-8 prodomain (residues 1-181), DED-A (1-80), DED-B (100-181) were cloned by the conventional PCR-based technique in frame with a C-terminal mCherry or GFP tag. The specific amino acid mutations were generated with the QuikChange site-directed mutagenesis kit (Stratagene). The resulting mutations were confirmed by DNA sequencing. Antibodies used were as follows: anti-caspase-8 (clone 1C12, mouse mAb 9746, Cell Signaling Technology), anti-FADD (sc-56093, Santa Cruz Biotechnology, Inc.), APO 1–3 anti-Fas (clone CH11, 05-201, Millipore), anti-GFP (sc-9996), anti-HA (Sigma H9658).

4.2.2.2 Expression of human FADD and procaspase-8 fusion proteins – His-tagged human FADD was expressed as inclusion bodies in *E. coli* BL21 (DE3) cells, while the FADD mutants and MBP fused caspase-8 DED-AB were majorly soluble. A single colony transformed with pET29b FADD-His₆ plasmid or pMALc5-caspase-8 DED-AB was grown overnight at 37 °C in 20 ml of LB medium containing 50 µg/ml kanamycin or 100 µg/ml ampicillin, respectively. The culture was then inoculated into 1 L of fresh liquid LB medium containing the respective antibiotics and once the cells reached an A₆₀₀ of 0.6 – 0.8, expression of the protein was induced at 18 °C for an additional 16 h by adding IPTG to a final concentration of 0.5 mM. The cells were pelleted by centrifugation at 6,000 rpm at 4 °C for 10 min.

4.2.2.3 Purification of human FADD and procaspase-8 fusion proteins – The FADD wild-type purification was followed as described previously [162]. Briefly, the bacterial pellet was re-suspended in a volume of 20 ml buffer-A (20 mM sodium phosphate buffer, pH 8.0, 150 mM NaCl, 0.1% Triton-X 100, 1X protease cocktail inhibitor) and ultra sonicated for 10 min on ice. The insoluble fraction was washed with buffer-B (20 mM sodium phosphate buffer, pH 8.0, 150 mM NaCl, 1% Triton X-100) by stirring with the magnetic agitator for 30 min. After centrifugation for 20 min at 12,000 rpm and 4 °C, the pellet of inclusion bodies was re-suspended in 10 ml buffer-C (20 mM sodium phosphate buffer, pH 8.0, 150 mM NaCl, and 8 M urea) and dissolved by stirring with the magnetic agitator for 2 h at room temperature followed by centrifugation at 12,000 rpm, 20 min, 4 °C. The supernatant was added to the Ni²⁺ affinity column equilibrated with buffer-C and incubated for 1 h. The column was washed with buffer-D (20 mM sodium phosphate pH 8.0, 20 mM imidazole, 150 mM NaCl) to remove nonspecific binding proteins. The His₆-tagged human FADD was eluted with buffer-E (20 mM sodium phosphate, pH 8.0, 250 mM imidazole, 150 mM NaCl). Refolded human FADD was obtained by dialyzing the denatured human FADD in 2 L of refolding buffer (20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 10% sucrose, 1 mM EDTA) over the period of 30 h at 4 °C.

The purification of FADD mutants and MBP fused procaspase-8 DED-AB was performed in non-denaturing condition. Briefly, the cell pellet was re-suspended in buffer-A and sonicated for 10 min in ice-cold condition. After centrifugation for 30 min at 4 °C, the cell lysate was passed through pre-equilibrated Ni²⁺ (for FADD) or amylose (DED-AB) affinity resin for 1 h at 4 °C. Beads were washed three times with buffer-B and the bound proteins were later eluted using imidazole or maltose gradient, respectively. The proteins were further monitored and purified by gel filtration chromatography using Superdex 75 or

200 prep grades. The purified proteins were analyzed by 12% SDS-PAGE, and the final product was stored in multiple aliquots and frozen at -80 °C until further use.

4.2.2.4 Western blotting and co-immunoprecipitation assays – HEK 293 cells were maintained in DMEM media supplemented with 10% foetal bovine serum. Sub-confluent six-well dishes were transfected with 1 µg of the indicated constructs by Lipofectamine reagent (Life Technologies) as per the manufacturer's protocol. 20 µM of z-VAD-fmk was added when active caspase-containing constructs were used. 24 h after transfection, cells were lysed for 30 min on ice in lysis buffer containing 50 mM Tris-HCl pH 7.4, 137 mM NaCl, 1 mM EDTA, 0.5% NP-40, 1X protease cocktail inhibitor. The detergent-insoluble fraction was pelleted by centrifugation at 18,000 rpm for 10 min. The cell pellet was washed three times with lysis buffer before boiling in SDS sample buffer. Samples containing lysates and pellets from equal cell numbers were electrophoresed on 12% Tris/glycine/SDS gels and blotted onto nitrocellulose membrane using wet transfer apparatus. Blots were blocked with 3% bovine serum albumin (BSA) for 30 min, and were probed with 1:1,000 dilution of anti-GFP (sc-9996, Santa Cruz Biotechnology, Inc.) followed by 1:6,000 dilution of goat anti-mouse IgG HRPO (12-349, Millipore) with three washes after each incubation. Incubations and washes were performed with Tris buffered saline containing 0.1% Tween-20. Bands were imaged using enhanced chemiluminescence ECLprime kit (GE Healthcare). For immunoprecipitation, cells were extracted in 1% CHAPS buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 1% CHAPS, 0.1% Triton-X, 5% sucrose) supplemented with protease inhibitors for 30 min at 4 °C, followed by centrifugation for 30 min at 13,000 rpm. Lysate (1 mg) was pre-cleared overnight with Protein G Sepharose beads (GE Healthcare) followed by incubation with 2 µg of anti-HA antibody for 4 h. The immunoprecipitate was then mixed with 100 µl of Protein G Sepharose beads and incubated for 2 h at 4 °C. After several washes,

beads were re-suspended in Laemmli SDS sample buffer and heated at 95 °C for 5 min and processed for immunoblotting using antibody against FADD.

4.2.2.5 Live cell confocal imaging – Confocal imaging was performed with Zeiss LSM 510 Meta confocal laser scanning microscope equipped with 63× 1.4 NA (numerical aperture) objective. Cells were grown to 60% confluency in DMEM media supplemented with 10% foetal bovine serum on glass-bottomed dishes (Cell E&G, USA). Prior to transfection, the cells were placed in Opti-MEM media, and transfected with 1 µg of the indicated constructs using Lipofectamine reagent. Post 20 h of transfection, the cells were imaged at room temperature. GFP fluorescence was excited with a 488-nm argon laser and collected between 505 and 550 nm. Similarly, mCherry was excited with a 543-nm helium/neon laser and emission was captured between 560 and 760 nm. Z-stacks were collected at intervals of 2–4 seconds. Images acquired were further processed using LSM 510 image examiner software.

4.2.2.6 Molecular docking and simulation – The docked homodimeric structure of caspase-8 DED-AB and FADD, procaspase-8 – FADD complex was generated using the HADDOCK (High Ambiguity Driven protein-protein DOCKing) [143] and ClusPro 2.0 [163, 164] web servers. The caspase-8 DED modeled structure and FADD (structure 2GF5) were used as the starting structure for docking calculations. Active site residues were defined based on the mutagenesis data from this study and surface accessibility (>30% exposed to solvent), and passive residues were allowed to be defined automatically. A total of 1000 complex structures were calculated using rigid body docking, with the best 200 structures subjected to further refinement and cluster analysis. The lowest energy structure was chosen for further examination. In addition, molecular docking was performed with a different and advanced blind docking program, ClusPro 2.0. The ClusPro server is based on a Fast Fourier Transform

correlation approach, which makes it feasible to generate and evaluate billions of docked conformations by simple scoring functions. It is an implementation of a multistage protocol: rigid body docking, an energy based filtering, ranking the retained structures based on clustering properties, and the refinement of a limited number of structures by energy minimization. The server returns the top models based on energy and cluster size. We selected the returned models after considering the energy and the size of the cluster – preferring lower energies and larger cluster sizes. After finalizing the docked complex, it was subjected to PRCG (Polak Ribierre conjugate gradient) energy minimization with PRIME Minimization module of Schrödinger 2014 suite (Schrödinger LLC, NewYork, LLC) using water as implicit solvent, OPLS (Optimized Potential for liquid simulations) force field up to 5000 minimization steps. Subsequently, it was optimized with protein preparation wizard of Schrödinger 2014 suite (Schrödinger LLC, New york, LLC) and further subjected to molecular dynamic simulation. The binding interface residues involved in hydrophobic, hydrogen bonding, ionic and cation-pi interactions were evaluated using the PDBsum generate server [145] and Protein Interactions Calculator [147] with the default cut-offs. The homodimeric caspase-8 DED-AB or FADD structure was then serially docked using ClusPro to generate the DED-chain assembly.

4.2.2.7 CD spectroscopy – Far-UV CD spectra for wild-type and mutant FADD proteins were recorded using a JASCO J 815 spectropolarimeter with 1mm path length cell. 5 μ M of protein in buffer (10 mM NaH₂PO₄/Na₂HPO₄ pH 8.0, 100 mM NaCl) was scanned in a wavelength range of 200 – 250 nm at 20 °C. Average blank corrected data of three independent scans were considered. Molar ellipticity was calculated, and data analysis was done using DichroWeb server (<http://dichroweb.cryst.bbk.ac.uk>) [165]. For thermal denaturation, wild type and mutant proteins were unfolded in a temperature range of 10 – 90

°C and wavelength 200 – 230 nm. Fraction unfolded was calculated at the different temperatures. The experiment was performed three times independently, and an average data was considered. To calculate the melting temperature, data fitting was done according to two-state transition model described earlier in chapter-3.

4.2.2.8 Caspase-8 cleavage assay – HEK 293 cells were transfected with the plasmids as indicated in the figure. Post 12 h, the agonist anti-Fas antibody (1 µg/ml) was added for 10 h. For each experiment, cells expressing wild-type caspase-8 incubated with pan caspase inhibitor (z-VAD-fmk) was kept as a control. Cells were lysed on ice with caspase assay buffer (20 mM HEPES pH 7.7, 100 mM NaCl, 10 mM DTT, 1 mM EDTA, 0.1% CHAPS, 10% sucrose 0.5% NP40) in the absence of protease inhibitors followed by centrifugation at 13,000 rpm for 15 min at 4 °C. 20 µg of cell extracts were used for SDS-PAGE and subsequent western blotting. Caspase-8 activity was determined using synthetic tetrapeptide fluorogenic substrate Ac-IETD-AFC. Caspase activity was measured with 30 µg proteins with 200 nM of Ac-IETD-AFC in 100 µl of caspase assay buffer at 37 °C for 2 h. The reaction was performed using a fluorescence microplate reader with excitation and emission wavelengths of 405 and 510 nm respectively. The rate of IETD hydrolysis was calculated using linear regression analysis using KaleidaGraph (Synergy software).

4.2.2.8 Cell viability - HEK 293 cells were transfected with the plasmids as indicated in the figure. Post 12 h, anti-Fas antibody (1 µg/ml) was added and the cells were incubated for 10 h. The cells were then trypsinized and washed twice with ice-cold PBS. One million cells were re-suspended in 400 µl PBS containing 2 µg/ml propidium iodide (PI), and analyzed by FACS Calibr (BD Bioscience) for GFP fluorescence and PI content. Percent dead cell was quantitated by % of GFP and PI-positive cells /GFP-positive cells.

4.2.3 Results

4.2.3.1 Structural basis of inter-molecular DED/DED interaction

To determine the basis of DED assembly, we performed comprehensive mutational and computational analyses of DED-containing proteins, procaspase-8 and FADD. Till the date, there are only two reports on DED-DED interaction [166, 167], based on the crystal structure of MC159 vFLIP which contains tandem DED similar to procaspase-8, however, incapable of self-association. Structural overlay of vFLIP DED and modeled caspase-8 DED showed a root mean square deviation of 1.6 Å over 178 aligned C α atoms (**Fig. 4.2.1A**). Structural analysis highlighted that similar to vFLIP, the intra-molecular interface of caspase-8 DEDAB exhibited a characteristic (helix-helix)/(helix-helix) interaction pattern, wherein helices 2 and 5 of DED-A line up against the helices 1 and 4 of DED-B (**Fig. 4.2.1A**). In this typical arrangement, the highly conserved hydrophobic patch phenylalanine/leucine (FL) of DED-A which lies on the helix α_2 , is integral to the interface between the two domains [168]. The corresponding residues on DED-B and DED of caspase-8 and FADD are F122/L123 and F25/L26 respectively. Although the hydrophobic patch on DED-A of caspase-8 is involved in tandem DED interaction, the same hydrophobic patch on DED-B is solvent exposed (**Fig. 4.2.1B**). Based upon the conserved nature of DEDs, we hypothesized that the inter-molecular DED/DED interaction might follow similar pattern as that of intra-molecular interaction. To test our hypothesis, protein-protein docking was performed between two molecules of caspase-8 DED-AB using a blind-docking software ClusPro and directed-docking software HADDOCK. Interestingly, both the docking analysis pointed a specific orientation for DED-AB inter-molecular interaction, where α_1/α_4 of DED-A showed reciprocal interaction with α_2/α_5 of DED-B (**Fig. 4.2.2A**). The inter-molecular interface had 1520 Å² of extensively buried surface area and notably involved residues R5, Y8, L42, F45 in DED-A and S119, F122, L123, Q125, E126, R162, Q166 in DED-B as shown in **Fig. 4.2.2B** (for a detailed list

of interactions please find **Table 4.2.1**). Thus, based on these observations it seems that DED-chain assembly involves the characteristic (helix-helix) / (helix-helix) association.

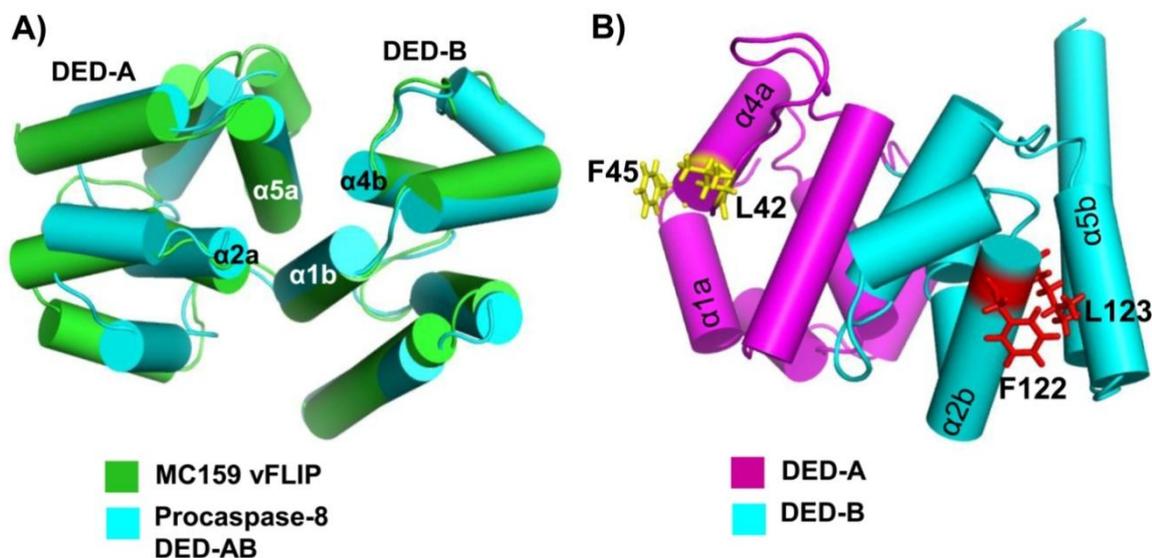


Figure 4.2.1 Molecular modeling and superposition of tandem DEDs. A) Homology model of human procaspase-8 DED-AB (residues 1-185) superimposed on MC159 vFLIP X-ray structure (PDB id. 2BBR) shown in green and cyan colours respectively. The intramolecular interface of the tandem DED involves typical (helix-helix) / (helix-helix) association, where $\alpha 2a/\alpha 5a$ of DED-A line up against $\alpha 1b/\alpha 4b$ of DED-B. The procaspase-8 DED-AB and MC159 share 23% identity; the rmsd calculated on $C\alpha$ atoms is 1.6 Å. B) Cartoon representation of procaspase-8 tandem DED showing the two surface exposed hydrophobic patches, *LXXF motif* (L42A/F45A) from DED-A and *FL motif* (F122/L123) in stick model.

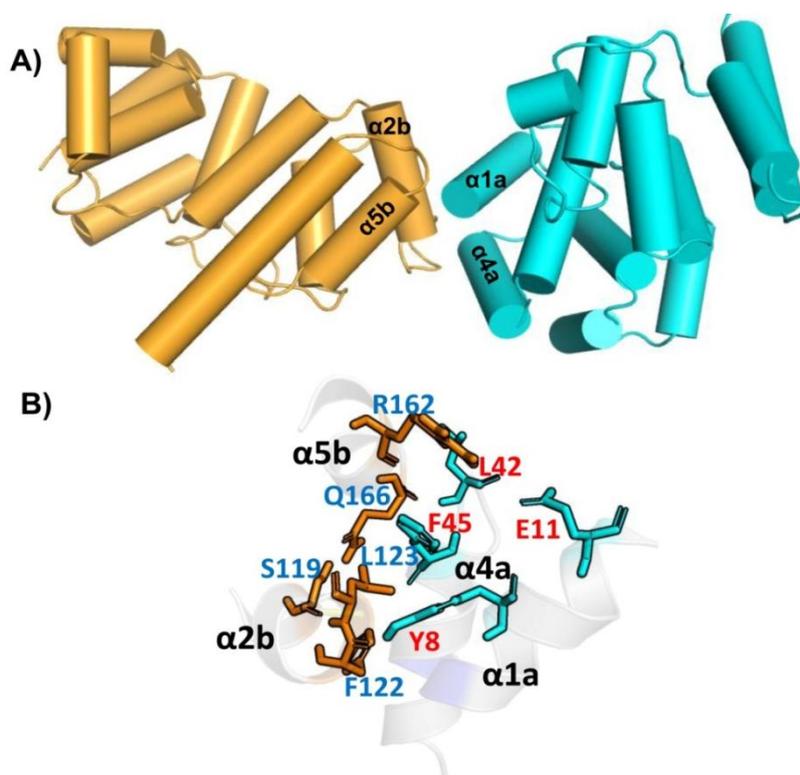


Figure 4.2.2 Inter-molecular interface of the procaspase-8 DED/DED complex. A) Docked homodimeric structure of procaspase-8 tandem DED. The inter-molecular homotypic interaction between tandem DEDs involves two surfaces, $\alpha 2/\alpha 5$ of DED-B from one tandem DED molecule (bright orange) interacting with $\alpha 1/\alpha 4$ of DED-A from the second molecule (cyan). B) Binding interface of the tandem DED intermolecular complex. A close-up view of the interface is shown in cartoon representation with selected side chains from each molecule shown in stick model. The residues from $\alpha 2b/\alpha 5b$ surface are shown in orange and $\alpha 1a/\alpha 4a$ in cyan colour.

Table 4.2.1 List of interactions across DED-A – DED-B inter-molecular interface evaluated using the *Protein Interactions Calculator* with the default cut-offs.

Type of interaction	DED-B	DED-A	Distance (Å)
Hydrogen bonds			
Main chain- side chain	O Phe122	OH Tyr8	2.71
	N Gln125	OH Tyr8	3.18
Side chain- side chain	O ^{ε1} Glu127	NH2 Arg5	3.21
	O ^γ Ser119	N ^{ε2} Gln49	2.91
	N ^{ε2} Gln166	O ^{ε1} Gln46	2.79
	O ^γ Ser122	N ^{ε2} Gln49	2.91
Main chain-main chain	O Gln166	N Leu42	3.18
Salt bridge			
	O ^{ε1} Glu127	N ^ε Arg5	2.70
Hydrophobic interactions (within 5Å)			
	Phe122	Leu48	
	Phe122	Phe3	
	Phe122	Met53	
	Leu123	Phe45	
	Leu123	Leu48	
	Leu123	Tyr8	
	Ile159	Phe45	
	Val163	Phe45	
	Ala165	Leu42	
	Ala165	Met43	
	Ile167	Leu42	
Cation-pi interaction			
	Phe3	Lys130	4.67

4.2.3.2 Molecular evidence of procaspase-8 DED assembly

It has been established that cells with overexpressed tandem DED of procaspase-8 form distinct cytoplasmic filamentous DED-chain network called 'death effector filaments' or DEF

[154, 169-171]. Therefore, to test the modeling predictions and identify the molecular basis of DED-chain formation, DED-AB of caspase-8 was fused with GFP and series of mutations were introduced on helices $\alpha 1/\alpha 4$ and $\alpha 2/\alpha 5$ of caspase-8 DED-A and DED-B respectively (**Table 4.2.2**), which were then tested for their ability to form DEF. As previously observed [154, 155], upon overexpression in HEK 293 cells, wildtype DED-AB formed fine long cytoplasmic filaments. When we examined cells expressing mutant DED-AB, strikingly, alanine substitution of residues L42/F45 (helix α -4) and F122/L123 (helix α -2) on DED-A and DED-B surface respectively abolished DED-chain filament formation and showed diffused cytoplasmic localization (**Fig. 4.2.3A**). Interestingly, we observed that neither one of these point mutation alone inhibited filament formation, although both mutations together did. Thus, it is unlikely that these mutations disrupted the structure of caspase-8 DED-AB. In general, modification of the surface residue to a less bulky side chain such as alanine is improbable to alter the overall protein conformation. To further investigate the importance of these residues in homotypic DED interaction, we performed multiple sequence alignment of the DEDs of FADD, caspase-8, cFLIP, MC159 vFLIP and PEA-15 to ascertain whether the residues are conserved (**Fig. 4.2.4**). Strikingly, in the DEDs of FADD and cFLIP, the residues corresponding to L42, F122, and L123 are preserved while the F45 is conserved in terms of its hydrophobicity. However, it is interesting to note that the DED-containing proteins vFLIP and PEA-15 which do not oligomerize showed no conservation. This suggests that the inter-molecular DED-DED association involves typical interaction between conserved binding surfaces. Hence, plausibly DED-DED recognition is mediated through the two surface exposed hydrophobic patches '*LXXF*- and *FL-motifs*' on DED-A and DED-B respectively (where X represents any amino acid).

Table 4.2.2 List of mutants characterized to identify the critical residues involved in DED/DED interaction.

Protein		Mutant	Mutation sites	Domain		DEF formation ^a
				DED-A	DED-B	
GFP DED-AB	Hydrophobic patch	D1	L42A	α 4		+
		D2	F45A	α 4		+
		D3	L42A, F45A	α 4		-
		D4	F122Y		α 2	+
		D5	L123A		α 2	+
		D6	L123A, Q125A		α 2	+
		D7	F122A, L123A		α 2	-
	Polar / charged residue	D8	Y8G	α 1		+
		D9	R5A, Y8A	α 1		+
		D10	S119A, E126A		α 2	+
		D12	(R5A, Y8A), (S119A, E126A)	α 1	α 2	-
		D13	K161A, R162A		α 5	+
		D14	R162A, Q166A, N168A		α 5	+
		D15	(S119A, E126A), (R162A, Q166A, N168A)		α 2/ α 5	-
		D16	(R5A, Y8A), (R162A, Q166A, N168A)	α 1	α 5	-
	mCherry-FADD	F1	V6A		α 1	+
F2		H9A		α 1	-	
F3		H9G		α 1	-	
F4		T21A		α 2	+	
F5		F25Y		α 2	-	
F6		L28A		α 2- α 3 loop	+	
F7		K33R		α 3	-	
F8		R34A		α 3	+	
F9		K35R		α 3	+	
F10		E37A		α 3	-	
F11		R38A		α 3	+	
F12		L43A		α 4	+	
F13		L43D		α 4	-	
F14		D44A		α 4	+	
F15		E51A		α 4	-	
F16		L62A		α 5	+	
F17		L63A		α 5	+	
F18		A68F		α 5	+	

^a (+) indicates no effect on DEF formation, (-) Loss or decrease in filament formation

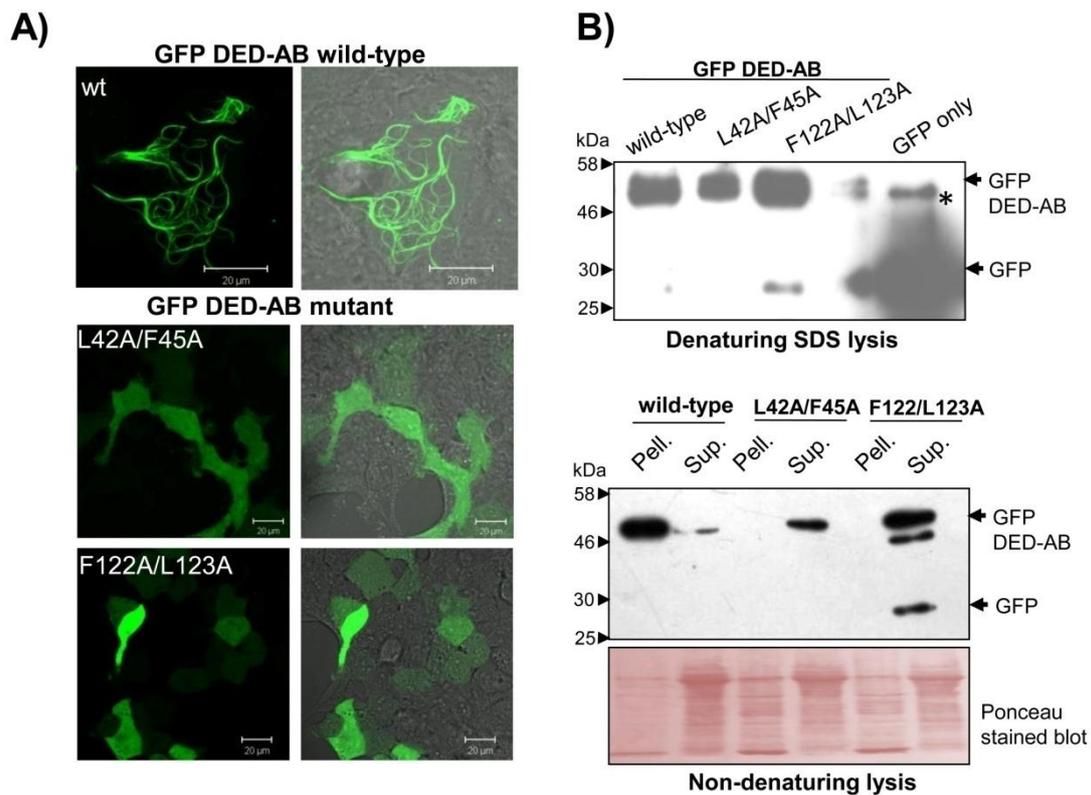


Figure 4.2.3 Mutational analysis of caspase-8 Death Effector Filament formation. A) Sub-cellular localization of procaspase-8 DED-AB wild-type or mutant proteins. HEK 293 cells were transiently transfected with 0.5 μ g of wild-type or mutant L42A/F45A, F122A/L123A caspase-8 DED-AB fused to GFP. Cells were imaged after 20 h using a Zeiss LSM 510 confocal microscope, and a representative field for each transfection is shown with the scale bar. B) Solubility of DED proteins in detergent correlates with their morphological appearance. In top panel, equal number of transfected cells lysed in denaturing SDS lysis buffer and in the lower panel, 0.5% NP40 lysate supernatants (Sup.) and pellets (Pell.) were separated on 12% SDS polyacrylamide gels and blotted with anti-GFP mAbs. The asterisk (*) represents a non-specific band.

Since the DEF display morphology similar to cytoskeletal proteins which are known to be completely insoluble in non-ionic detergents, we examined the solubility properties of DED-AB wildtype and mutant proteins. It was observed that DEF formed by wildtype DED-AB was detergent-insoluble, whereas the F122A/L123A and L42A/F45A mutant proteins were predominantly in the soluble fraction corresponding to their diffused distribution (**Fig. 4.2.3B**). To test the structural feasibility for the formation of chain assembly, we performed serial docking of DED-AB as described in *experiment section*. The resulting model suggested that assembly of the DED-chain exhibit a typical helical characteristic with a propensity to form continuous chain like structure (**Fig. 4.2.5**).

In order to test the specificity of DED-DED inter-molecular interaction, we checked the properties of DED-A and DED-B individually. Interestingly, upon overexpression both the domains alone did not form any filament. This led us to examine whether individual DEDs would affect the filament forming ability of tandem DED-AB. Indeed, co-transfection of DED-A or DED-B with tandem DED resulted in disruption of the filaments and appeared as colocalized cytoplasmic aggregates (**Fig. 4.2.6**). This possibly resulted due to the association of DED-A ($\alpha 1/\alpha 4$) or DED-B ($\alpha 2/\alpha 5$) with the reciprocal interaction sites on DED-AB (i.e ($\alpha 2/\alpha 5$) for DED-A and ($\alpha 1/\alpha 4$) for DED-B), thereby blocking or inhibiting chain elongation process. Consistent with this observation, when we co-transfected mutated L42A/F45A DED-A or F122A/L123A DED-B, the tandem DED formed an elaborate filamentous network without any colocalization with individual DEDs (**Fig. 4.2.6**). These results indicate that the assembly of DED-chain specifically involves hetero-oligomerization via the surfaces defined by helices $\alpha 1a/\alpha 4a$ and $\alpha 2b/\alpha 5b$.

Having identified that the DEDs associate tightly through a hydrophobic surface, the involvement of charged or any other residues in the assembly was needed to be investigated. To thoroughly map the interface of DED/DED complex, alanine-scanning mutagenesis was performed. Alanine-scanning mutagenesis, a method of systematic alanine substitution, has been widely used for the identification of important residues for complex formation. Computational alanine-scanning using ‘interface alanine scanning’ program of Robetta server predicted several interface residues essential for interaction (**Table 4.2.3**). In order to identify the critical residues, we generated several combinatorial mutants listed in **Table 4.2.2** and tested their ability to form DED-chain filaments. To our surprise, a significant number of polar and charged residues located on $\alpha 1a/\alpha 4a$ and $\alpha 2b/\alpha 5b$ were found to be involved in homotypic DED interaction. However, in contrast to the hydrophobic patch mutants (L42/F45 and F122/L123); only constructs with residues substituted from different helices in

combination were found defective in forming filaments. The mutants D12, D15 and D16 showed complete cytoplasmic distribution whereas all the other constructs retained their ability to form filaments. Therefore, it implies that although hydrophobic interaction provides the driving force for DED/DED interaction, additional inter-domain hydrogen bonding and van der Waals interactions further stabilize the DED complex.

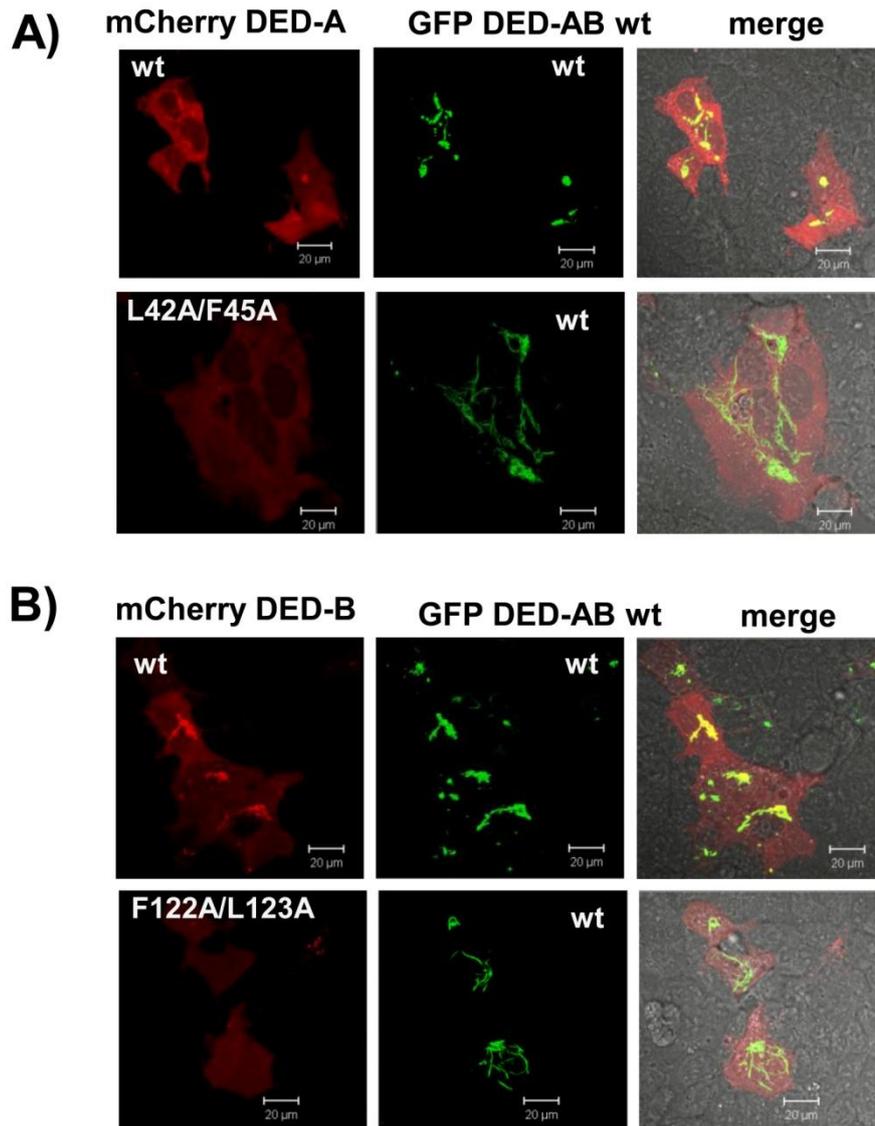


Figure 4.5.6 Subcellular localization of GFP DED-AB and individual DED domains in co-transfected cells. HEK 293 cells were transfected with GFP DED-AB and wild-type or mutant mCherry fused DED-A (A) or DED-B (B). Live cell images were acquired using Zeiss LSM 510 confocal microscope, and a representative field for each experiment is shown. Scale bar represents 20 μm . Wild-type DED-A or DED-B disrupted the DED-AB filaments and appeared as colocalized cytoplasmic speckles or aggregates (top panel for A and B), while the mutants did not (lower panel). In case of mutants, GFP DED-AB formed filamentous network and mutant DED-A or DED-B shows diffused distribution.

Table 4.2.3 Predicted changes in the binding free energy ($\Delta\Delta G_{\text{bind}}$) for all the alanine mutations across the inter-molecular DED/DED binding interface calculated using Robetta server [172, 173].

DED-A		DED-B	
Residue no.	$\Delta\Delta G_{\text{bind}}$ (kcal.mol ⁻¹)	Residue no.	$\Delta\Delta G_{\text{bind}}$ (kcal.mol ⁻¹)
Phe3	0.95	Arg118	0.17
Arg5	1.65	Ser119	2.65
Tyr8	3.04	Lys121	0.02
Gln35	0.31	Phe122	2.33
Lys39	0.18	Leu123	1.81
Leu42	1.18	Gln125	0.49
Met43	0.8	Glu126	0.73
Phe45	2.54	Ser129	0.1
Gln46	1.86	Lys130	0.47
Leu48	0.19	Arg162	-0.05
Glu49	3.67	Gln166	2.11
Arg52	0.03	Ile167	0.89
Met53	0.43	Asn168	-0.05

$\Delta\Delta G_{\text{bind}}$, predicted change in binding free energy upon alanine mutation computed using the equation: $\Delta\Delta G_{\text{bind}} = (\Delta G_{\text{complex}}^{\text{wt}} - \Delta G_{\text{partnerA}}^{\text{wt}} - \Delta G_{\text{partnerB}}^{\text{wt}}) - (\Delta G_{\text{complex}}^{\text{mut}} - G_{\text{partnerA}}^{\text{mut}} - G_{\text{partnerB}}^{\text{mut}})$. Neutral residues and hot spots are defined as residues showing a change in the binding free energy by less or more than 1 kcal/mol respectively when replaced by alanine.

As the next step, we decided to test the effect of substitution of these residues on the aggregation behavior of the recombinant caspase-8 DED-AB using size exclusion chromatography. We observed that over the concentration range of 75-100 μM , the mutation of the hydrophobic motif (L42A/F45A, F122A/L123A mutant) showed mixture of aggregated and monomeric proteins in contrast to completely aggregated wild-type protein (**Fig. 4.2.7**). Furthermore, upon substitution of additional polar residues there was increase in monomeric population; however, we could not achieve 100 percent monomers. Since the hydrophobic interaction primarily drives DED-DED interaction, the substitution of hydrophobic motif to a non-polar aliphatic amino acid ‘alanine’ probably does not impart severe effect on its association in an *in vitro* condition, where the concentration extends over micromolar to millimolar range. It would be worth trying to test the effect of polar

substitution of these residues on the aggregation behavior of the recombinant protein. These informations might thus facilitate and increase the scope of structure determination for caspase-8 DED, one of the least characterized members of the death-fold domain superfamily.

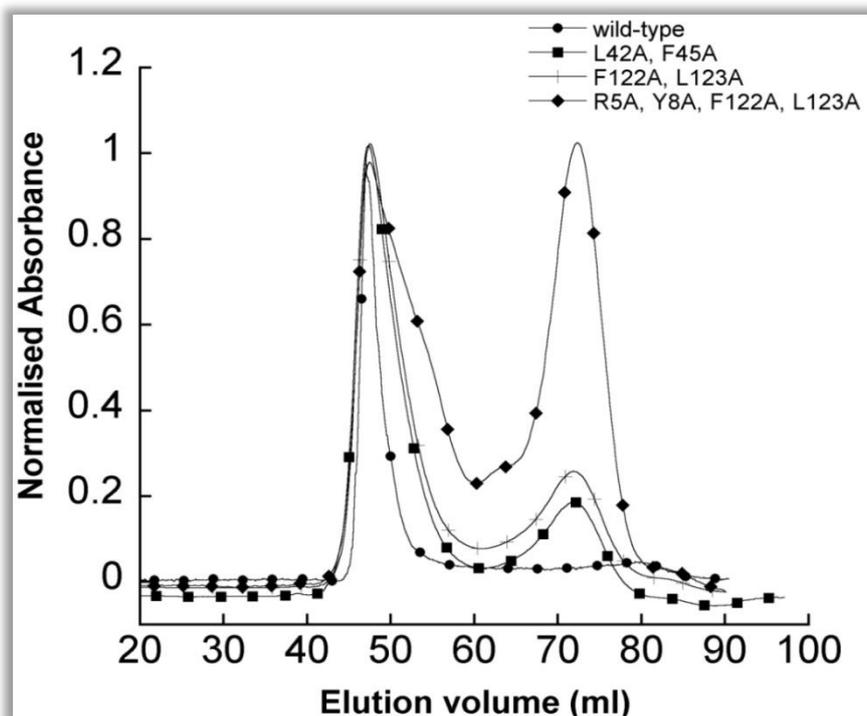


Figure 4.2.7 Oligomeric property of caspase-8 DED-AB wild-type and mutant proteins. Chromatographic profiles of MBP tagged DED-AB wild-type and mutant proteins eluting from the Superdex 200 16/600 gel filtration column. The protein eluted in two main peaks. The first peak at 44 ml corresponds to void volume of the column indicating a molecular mass of >440 kDa, whereas the second peak corresponds to a calculated molecular weight of 68 kDa, suggesting a monomer.

4.2.3.3 Molecular basis of FADD-DED self-association

Based on the high degree of homology within the DED family, we speculated that other DEDs probably also associate via a similar mechanism. To test our speculation, we characterized another DED-containing protein FADD, which self-associates through the DED, however, the exact binding interface is yet unclear. Similar to the prodomain of caspase-8, upon transient overexpression, FADD forms cytoplasmic DEFs [154]. However, unlike caspase-8 which comprises two DEDs, FADD contains only a single DED with an ability to

mediate DED/DED interaction. The self-association of FADD is required for stable interaction with activated death receptor followed by caspase activation [174], thereby further emphasizing the need for its characterization.

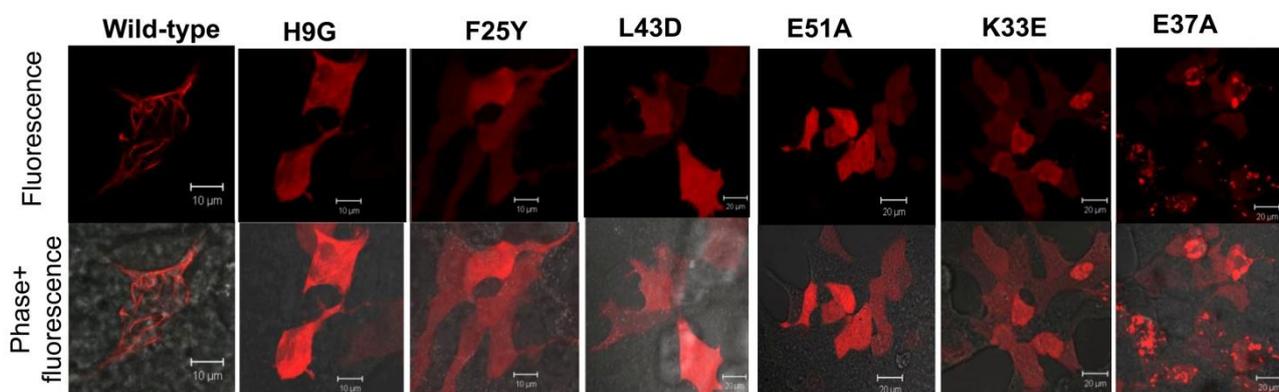


Figure 4.2.8 Mutations disrupting the FADD DED filament formation. HEK 293 cells were transiently transfected with 0.5 μg of wild-type or single mutant FADD constructs fused to mCherry. Cells were imaged after 20 h using a Zeiss LSM 510 confocal microscope, and a representative image for each transfection is shown with the scale bar.

Although a previous report highlighted the importance of residue F25 (helix α_2) in FADD DED self-association [64], we used an unbiased approach and designed 18 point mutants to explore all potential modes of interaction (**Table 4.2.2**). Based upon the available monomeric FADD (F25Y) structure (PDB id. 2GF5), we targeted the surface-exposed residues with solvent accessibility greater than 35%, to minimise any structural perturbations, and at least two residues per helix were mutated to cover the entire surface. These mutants were then tested for their ability to self-associate and form filaments. As anticipated, mCherry fused FADD wild-type showed filament formation upon overexpression in HEK 293 cells. When we tested the mutants, it was observed that point mutation of residues H9 (helix α_1), F25 (helix α_2) and L43, E51 (helix α_4) showed complete cytoplasmic distribution (**Fig. 4.2.8**). However, instead of residues from α_5 , mutants of helix α_3 , K33E and E37A demonstrated remarkable decrease in their ability to form filaments. It is interesting to note that of the residues corresponding to *LXXF* and *FL motif* in FADD DED, i.e. F25/L26 and

L43A/F46A, single substitution of F25Y or L43D is sufficient to prevent FADD DED self association, while the residues L26 and F46 were solvent inaccessible and therefore did not form part of the binding interface.

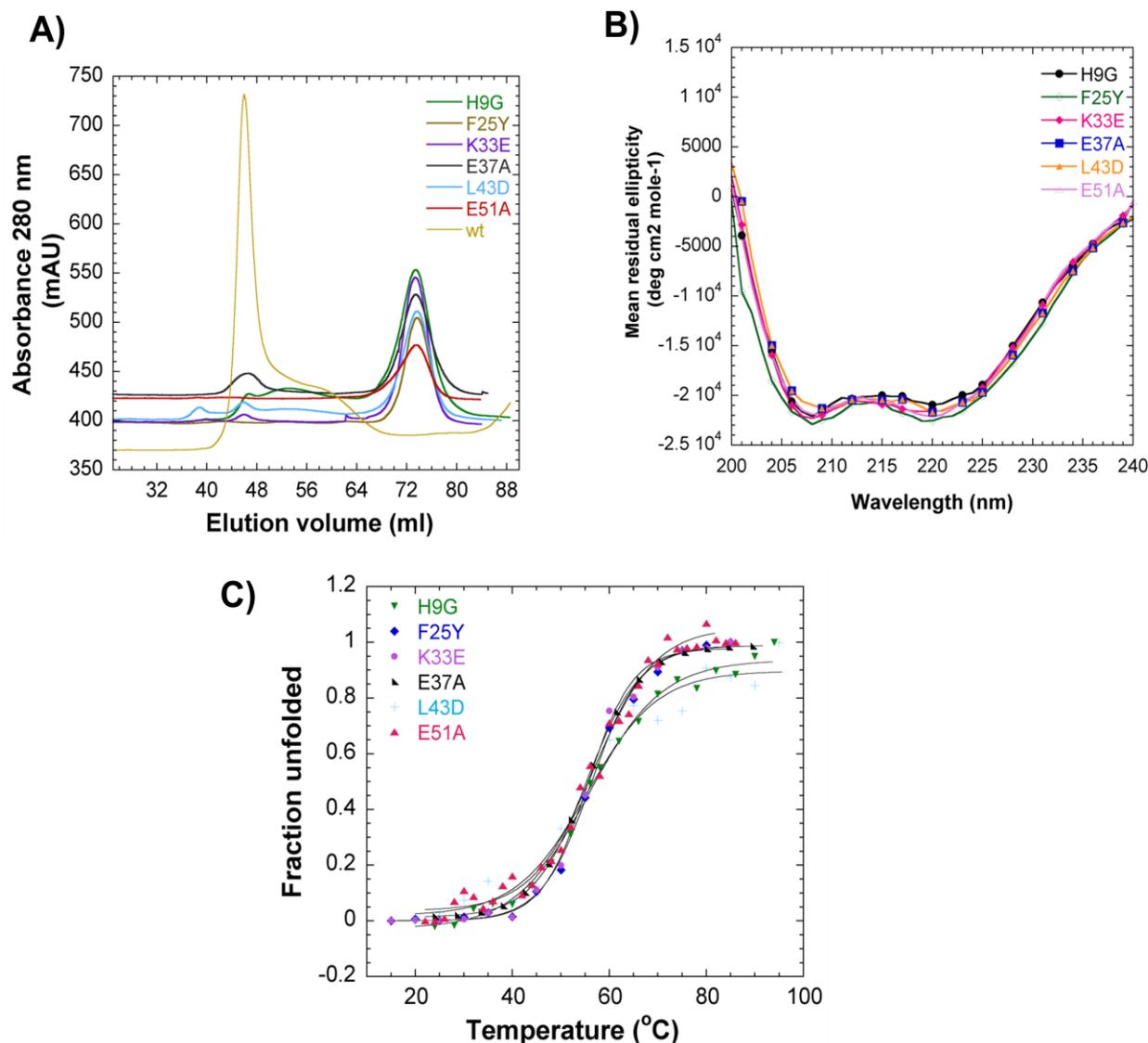


Figure 4.2.9 Characterization of FADD mutants. A) Gel filtration chromatography elution profile of FADD-His₆ wild-type and mutant proteins injected in Superdex 75 16/600 column. The wild-type FADD eluted as an aggregate, while the mutants eluted predominantly in the elution volume corresponding to a monomer. B) Far ultraviolet circular dichroism spectra of all the monomeric FADD mutants. The measurements were made at 20 °C in 10 mM NaH₂PO₄/Na₂HPO₄ pH 8.0, 100 mM NaCl. Contributions to the spectra by the buffer were subtracted using control scans. All the curves are represented as mean residual ellipticity in unit deg cm² mol⁻¹. The CD spectra reveal that all the mutants have similar helical content. C) Thermal denaturation of monomeric FADD proteins monitored by secondary structural changes with increasing temperature. Ellipticity at 222 nm was used to calculate the fraction folded plotted against the incubation temperature in °C.

Table 4.2.4 Secondary structural composition and melting temperature of FADD mutants.

FADD mutants	% Helix		T _m (°C) ^c
	By 222 nm ^a	K2D analysis ^b	
H9G	61.7	71	55.8 ± 0.9
F25Y	63.3	73	56.4 ± 0.4
K33E	62.4	72	55.6 ± 0.4
E37A	62.8	72	55.1 ± 0.3
L43D	61.9	72	55.1 ± 1.6
E51A	62.6	72	56.2 ± 0.5

^a Fractional helical content = $([\Theta]_{222} - 3000)/(-36000 - 3000)$ [175]

^b K2D analysis performed using Dichroweb server [165]

^c T_m represents mean value for three independent experiments ± S.D.

To determine whether these mutations also impact the solubility and aggregation behaviour in solution, we purified these recombinant FADD proteins. Contrary to wild-type FADD, which formed insoluble aggregates upon overexpression in bacterial system, all the above mentioned mutants showed increased solubility. In gel filtration column, over the concentration range of 50-100 μM, these mutations caused the proteins to shift to monomeric fraction (**Fig. 4.2.9A**). FADD mutants F25Y, E51A and K33E predominantly eluted as monomer, and H9G, E37A and L43D showed mixture of aggregated and monomeric species. To test any adverse effect of mutation, we assessed the structural integrity using spectroscopic probes. Structural analysis using Far-UV CD did not show any pronounced loss of secondary structure and all the mutants retained their helical characteristics (**Fig. 4.2.9B**). Analysis of the CD spectra using Dichroweb K2D analysis predicted 70 - 72% α-helix (**Table 4.2.4**), consistent with the helical content (76%) estimated for FADD F25Y (PDB id. 2GF5) structure. We also determined the relative stability of these proteins by monitoring unfolding in response to thermal denaturation. The thermal denaturation data indicate that all the FADD

mutants undergo a two-state unfolding transition (**Fig. 4.2.9C**). The melting temperature (the temperature at which half the protein molecules are unfolded) was found to be similar for all the proteins (**Table 4.2.4**).

Collectively, these studies show that the interaction surface for FADD-DED self-association is defined by helices $\alpha 1/\alpha 4$ and $\alpha 2/\alpha 3$ as against $\alpha 1a/\alpha 4a$ and $\alpha 2b/\alpha 5b$ for caspase-8. Furthermore, merely single substitution of the binding interface residues was sufficient to block FADD-DED association as opposed to caspase-8 DED that required multiple substitutions. In agreement to this, structural information for FADD was achieved with a single F25Y mutation; however, till date except for non-oligomerizing MC159 vFLIP, there is no structural data available on tandem DEDs. This suggests that possibly DED-DED interaction involving FADD is very transient or weak compared to caspase-8 DED-A/DED-B association.

To further substantiate the findings for FADD DED self-association, we generated a molecular model for FADD DED complex using HADDOCK as described in the *experimental section*. Structural analysis of the FADD DED homodimer showed that the interface involves 1062 \AA^2 buried surface area which is significantly less as compared to caspase-8 DED complex. The interaction appears to involve contribution from hydrophobic side chains, hydrogen bonding and salt bridges. Notably, residues S1, V6, H9, L43, D44 and E51 from $\alpha 1/\alpha 4$ of one molecule make contact with T21, F25, K33, R34, E37, and E51 from $\alpha 2/\alpha 3$ surface of the second molecule (**Table 4.2.5**). Overall, the observed interactions are in good agreement with the mutagenesis data. Furthermore, the docked model shows that the FADD DED homodimer (**Fig. 4.2.10A**) is oriented such that it enables the two interaction sites of each subunit to engage in further independent interactions either to build a higher order FADD oligomer (**Fig. 4.2.10B**) or bind with other interacting partners.

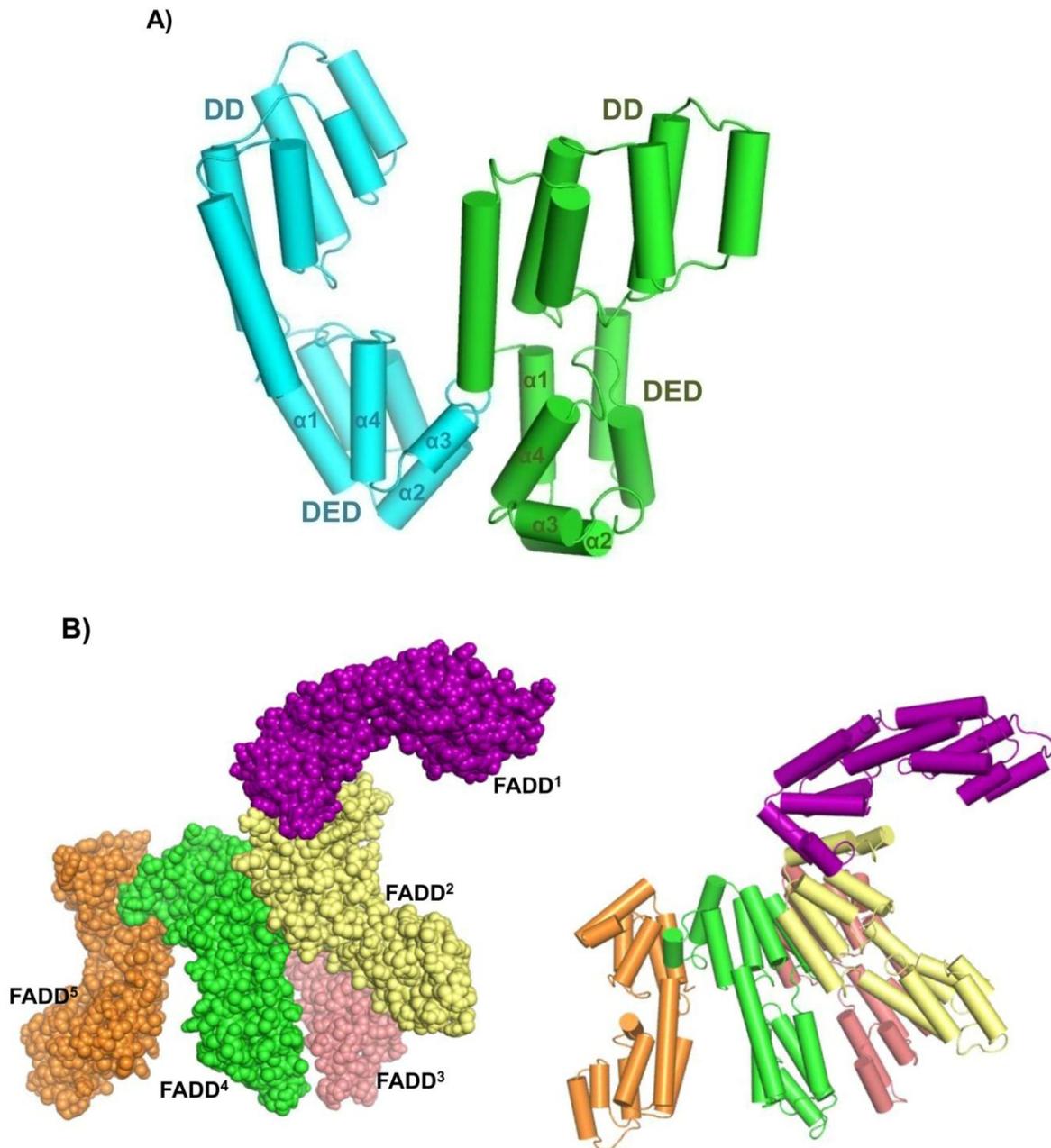


Figure 4.2.10 Structural modeling of FADD self-association. A) Cartoon representation of docked homodimeric modeled structure of FADD. The self-association of FADD involves interaction between $\alpha 2/\alpha 3$ of FADD DED with $\alpha 1/\alpha 4$ of the second FADD molecule. Note that the FADD DED homodimer are oriented in a manner, so as to enable the other interaction sites of each subunit ($\alpha 1/\alpha 4$ and $\alpha 2/\alpha 3$ in the first and second FADD molecule respectively) to engage in further independent interactions, either to build a higher order FADD oligomer or interact with other binding partners. B) Molecular assembly of FADD forming DED-chain like structure. The FADD DED assembly was generated using ClusPro by serial docking of the FADD units represented as spheres (left panel) or cylindrical helices (right panel). Each FADD molecule is shown by a different colour.

Table 4.2.5 Detailed list of interactions at FADD DED-DED interface evaluated using the *Protein Interactions Calculator* with the default cut-offs.

Type of interaction	DED (chain-A)	DED (chain-B)	Distance (Å)
Hydrogen bonds			
Side chain-main chain	O Gly29 ($\alpha 2$)	O ^{γ} Ser1 ($\alpha 1$)	2.82
	O Leu28 ($\alpha 2$)	N ^{$\delta 1$} His9 ($\alpha 1$)	3.43
Side chain-side chain	N ^{ζ} Lys33 ($\alpha 3$)	O ^{$\delta 2$} Asp44 ($\alpha 4$)	2.60
	NH1 Arg34 ($\alpha 3$)	O ^{$\epsilon 1$} Glu51 ($\alpha 4$)	3.25
Salt bridges			
	Lys33 ($\alpha 3$)	Asp44 ($\alpha 4$)	
	Glu37 ($\alpha 3$)	Asp44 ($\alpha 4$)	
	Lys33 ($\alpha 3$)	Glu51 ($\alpha 4$)	
	Arg34 ($\alpha 3$)	Glu51 ($\alpha 4$)	
Hydrophobic interactions (within 5Å)			
	Leu28 ($\alpha 2$)	Val6 ($\alpha 1$)	
	Phe25 ($\alpha 2$)	Leu43 ($\alpha 4$)	
	Leu36 ($\alpha 2$)	Leu43 ($\alpha 4$)	

4.2.3.4 Mechanism of FADD – procaspase-8 assembly

The results described above indicate that DED association involves a typical interaction between conserved binding surfaces, suggesting FADD and procaspase-8 must also engage through these patches. Protein-protein docking predicted that the interaction is feasible with either $\alpha 1/\alpha 4$ of caspase-8 DED-A binding to FADD DED $\alpha 2/\alpha 3$ or $\alpha 1/\alpha 4$ of FADD DED mediating reciprocal interaction with $\alpha 2/\alpha 5$ hydrophobic patch of caspase-8 DED-B (**Fig. 4.2.11**). Therefore, structurally there are three possibilities which could explain the initial recruitment of caspase-8 to FADD. Either procaspase-8 DED-B interacts with FADD while DED-A mediates caspase-8 chain formation or it could be vice a versa. Alternatively, procaspase-8 is recruited to FADD simultaneously by DED-A and DED-B with chain formation initiated by both DED-B and DED-A respectively. When ectopically expressed, FADD spontaneously self-oligomerizes and provides an oligomeric platform that is able to recruit and activate caspase-8 inducing apoptosis independent of death receptors [154, 176]. Therefore, this provided us an excellent way to examine the recruitment of procaspase-8 to FADD and initiation of DED-chain assembly.

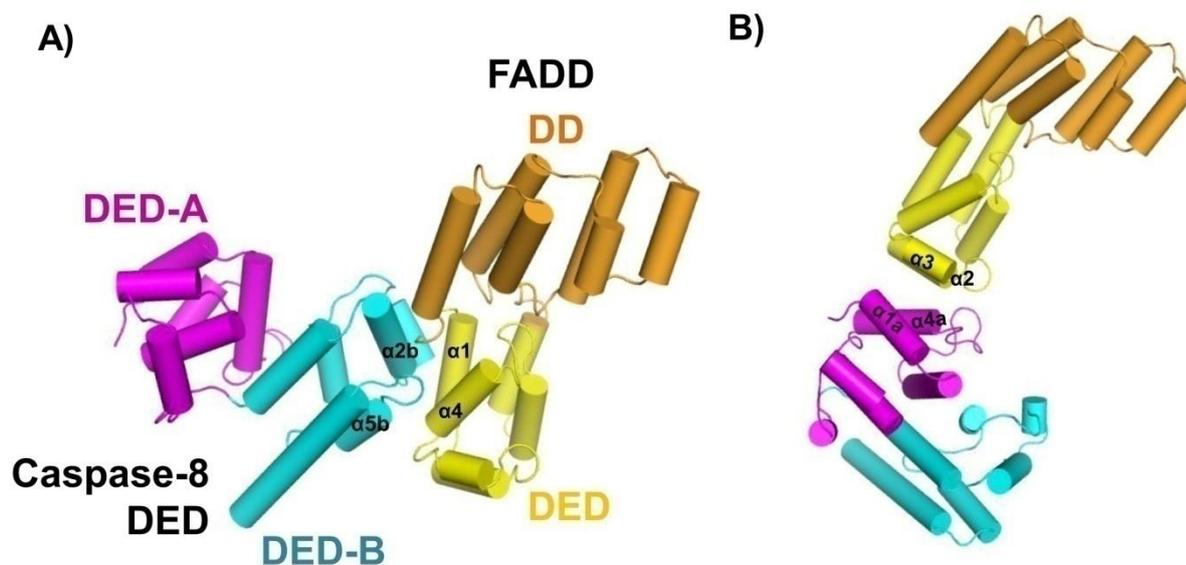


Figure 4.2.11 Two possible orientations of the FADD/procaspase-8 interaction. A docked model of FADD/procaspase-8 DED complex constructed with the program HADDOCK. The resultant models reveal that possibly $\alpha 2/\alpha 5$ helices of procaspase-8 DED-B make contact with $\alpha 1/\alpha 4$ of FADD DED (A) or the surface defined by helices $\alpha 1/\alpha 4$ of DED-A interact with $\alpha 2/\alpha 3$ of FADD DED (B).

To assess recruitment of procaspase-8 to FADD, we co-expressed wild-type FADD and catalytically inactive procaspase-8 C360A full length wild-type, *LXXF* or *FL motif* mutants. Notably, full length procaspase-8, which although contains DED, shows diffused cytoplasmic localization [177]. When cells co expressing mCherry-tagged FADD and GFP-tagged caspase-8 full length was examined using confocal microscope, we observed both the proteins colocalized in the form of short cytoplasmic filaments (**Fig. 4.2.12A**). In the presence of FADD, the normally diffuse distribution pattern of wild-type caspase-8 C360A-GFP re-localizes to the DEF. However, with respect to the mutants, the co-expression of *FL-motif* mutated caspase-8 C360A-GFP and FADD resulted in formation of colocalized cytoplasmic speckles or aggregates (**Fig. 4.2.12B**), instead of short filaments seen in case of wild-type proteins. Remarkably, the *LXXF-motif* mutant did not show any co-localization with FADD and we observed caspase-8 redistribution into cytoplasmic diffused state with FADD forming long filaments (**Fig. 4.2.12C**). To verify caspase-8 is essentially recruited to the oligomerized FADD platform, we coexpressed monomeric FADD mutants F25Y or E51A

and wild-type caspase-8 C360A-GFP. As anticipated, we did not see any relocalization of procaspase-8 and both the proteins showed a diffused distribution (**Fig. 4.2.12D and E**).

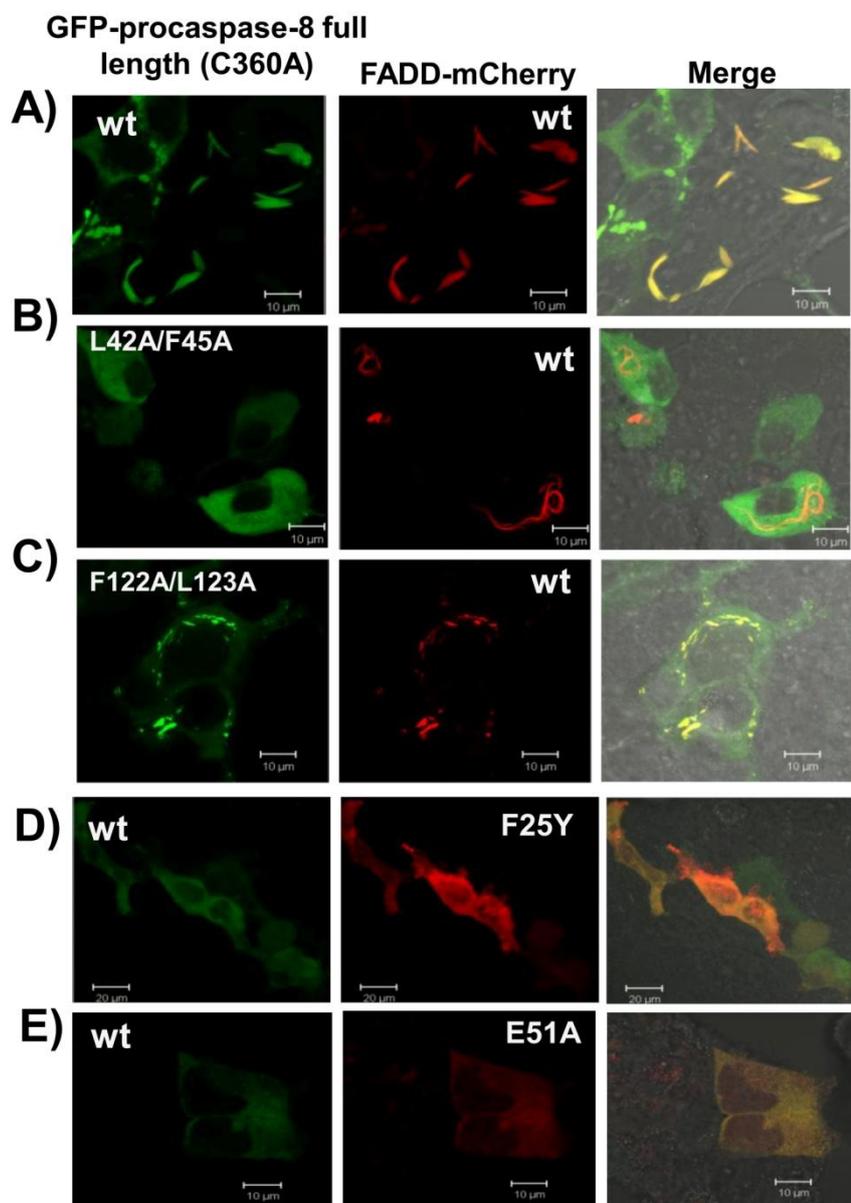


Figure 4.2.12 Localization of procaspase-8 and FADD in co-transfected cells. A-C) HEK 293 cells were transfected with mCherry fused FADD wild-type and GFP-tagged catalytic inactive procaspase-8 wild-type (A), L42A/F45A mutant (B) and F122A/L123A mutant (C). In panel D and E, cells were transfected with catalytically inactive wild-type procaspase-8 and F25Y (D) and E51A (E) FADD constructs. Live cell images were acquired using Zeiss LSM 510 confocal microscope, and a representative image for each transfection is shown along with the scale bar (Original magnification x630 with 2X optical zoom). Note that cells expressing both wild-type proteins show the presence of short cytoplasmic filaments and the two proteins show 100% colocalization as seen in the merged image. However, the L42A/F45A mutant do not show any colocalization with wild-type FADD, while F122A/L123A expressing cells show the presence of discrete colocalized cytoplasmic speckles or aggregates. The non-oligomerizing FADD mutants did not show any association with procaspase-8.

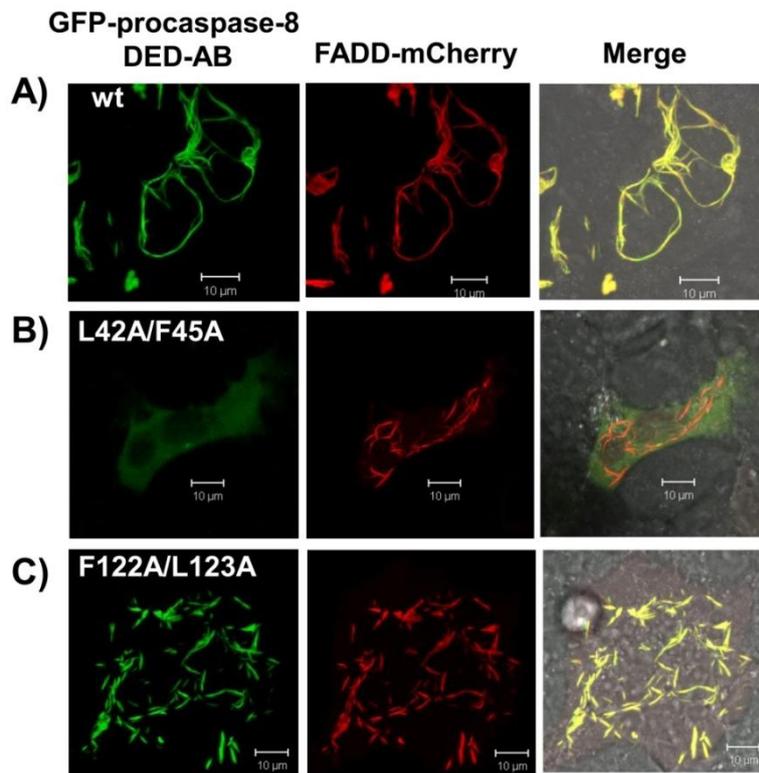


Figure 4.2.13 Localization of procaspase-8 DED-AB and FADD in co-transfected cells. A-C) HEK 293 cells were transfected with mCherry fused FADD wild-type and GFP tagged procaspase-8 DED-AB wild-type(A), L42A/F45A mutant (B) and F122A/L123A mutant (C). Live cell images were acquired using Zeiss LSM 510 confocal microscope, and a representative image for each transfection is shown along with the scale bar (Original magnification x630 with 2X optical zoom).

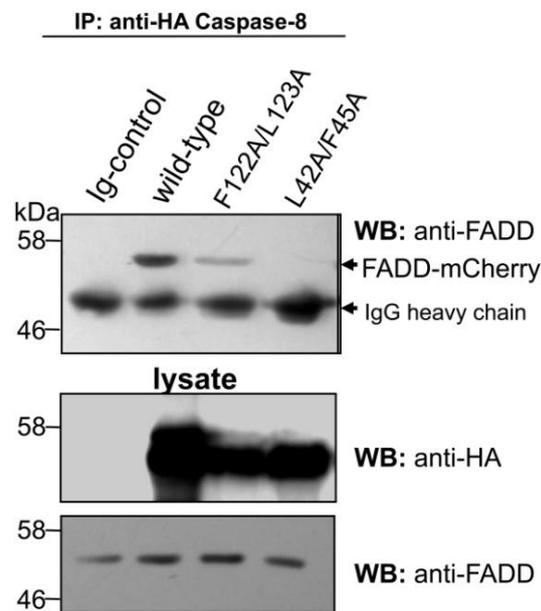


Figure 4.2.14 Immunoprecipitation of FADD/procaspase-8 complex. Extracts from HEK 293 cells co-transfected as described were immunoprecipitated using an anti-HA antibody followed by immunoblotting using an anti-FADD antibody. In lane-1 cells were transfected with vectors pcDNA3.0 and pmCherry-FADD; in lane 2-4 with pmCherry-FADD and HA-tagged procaspase-8 wild-type (lane-2), F122A/L123A mutant (lane-3), and L42A/F45A mutant (lane-4). 20 µg of the lysate were immunoblotted with anti-HA and anti-FADD antibodies.

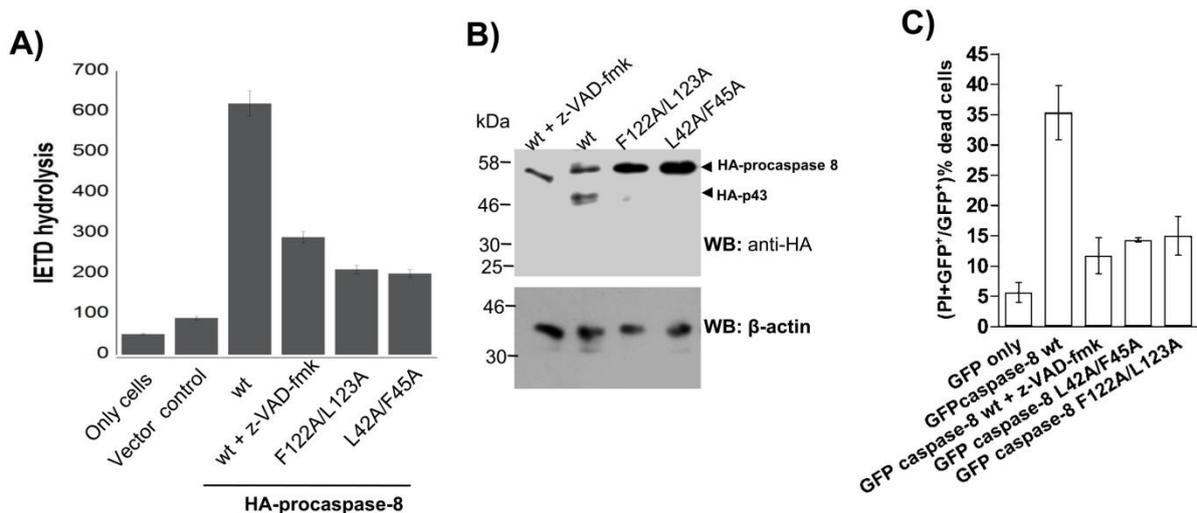


Figure 4.2.15 Fas-induced cell death analysis in presence of mutant procaspase-8. A) Caspase-8 activity was measured for the crude extracts isolated from the cells transfected with the plasmids as indicated, using fluorogenic Ac-IETD-AFC substrate. B) 20 μ g of cell extract immunoblotted with anti-HA and anti- β -actin antibodies. β -actin serves as a loading control. C) Viability of HEK 293 cells transfected with the expression plasmids as indicated, determined post 22 h of transfection by propidium iodide uptake. The bars represent the mean and the error bars the standard deviation.

The interaction of caspase-8 with FADD involves the prodomain and is reported to be independent of its catalytic domain [154, 178, 179]. Therefore, to gain further insight into DED assembly, we overexpressed mCherry-FADD in presence of GFP tagged-DED-AB wild-type, L42/F45 or F122A/L123A mutant proteins. When we examined the cells overexpressing the wild-type proteins, we observed 100% colocalization of DED-AB and FADD with the formation of long filaments. However, in case of the *FL-motif* mutant, it did show 100% co-localization with FADD but appeared as shorter filaments whereas the *LXXF-motif* mutant as anticipated did not show any co-localization (**Fig. 4.2.13**). The interaction was further validated by co immunoprecipitation assay which showed that F122A/L123A reduced but did not totally prevent caspase-8 recruitment to FADD, whereas the L42/F45A did not show any binding (**Fig. 4.2.14**).

To test the effect these mutations on procaspase-8 activation, we measured the enzymatic activity of active caspase-8 wild-type and mutant using fluorogenic IETD substrate upon treatment with the agonist anti-Fas antibody. Surprisingly, both the mutants (including F122A/L123A which still interacted with FADD) showed significant reduction in

the rate of IETD hydrolysis (Fig. 4.2.15A). Unlike the wild-type protein, the auto-proteolysed fragment of caspase-8 was absent in both these mutants (Fig. 4.2.15B). Furthermore, we observed significant reduction in the percent cell death in the mutant caspase-8 expressing cells compared to the wild-type (Fig. 4.2.15C).

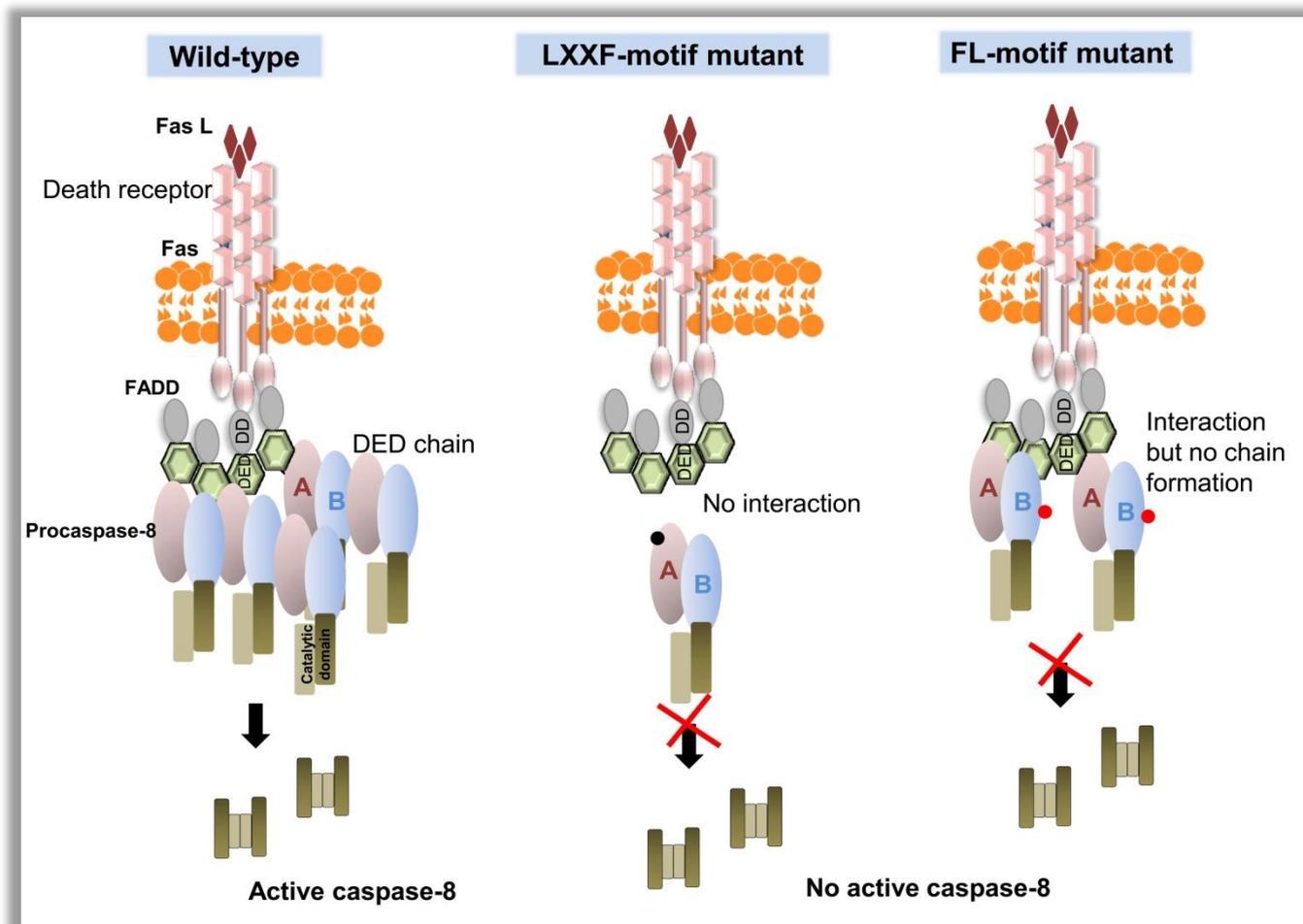


Figure 4.2.16 Proposed mechanism for death-fold mediated procaspase-8 activation. In case of the wild-type procaspase-8, upon stimulation by the appropriate death ligand, death receptors recruit FADD via their death domains (DDs) to form the oligomeric FADD platform. FADD in turn recruits the initiator caspases such as procaspase-8 through a DED-mediated interaction. FADD initially recruit procaspase-8 via direct interaction with *LXXF motif* of DED-A, followed by sequential interaction mediated by the DED-B *FL-motif* to form a chain, which in turn facilitates procaspase-8 dimerization leading to full activation to release active caspase-8 to promote cell death. In presence of *FL-motif* mutated procaspase-8, mutation of hydrophobic patch on DED-B does not affect the initial recruitment of procaspase-8 to FADD; however, it plausibly stalls the subsequent DED-chain formation and hence caspase-8 activation. However, mutation on DED-A completely blocks the initial binding of caspase-8 to FADD, hence DED assembly and activation of caspase-8.

Collectively, based on all the above experiments we propose that mutation of hydrophobic patch on DED-B does not affect the initial recruitment of procaspase-8 to FADD; however, it hampers the subsequent DED-chain formation, while mutation on DED-A completely blocks the initial binding of caspase-8 to FADD, hence DED assembly and activation of caspase-8 (**Fig. 4.2.16**). Therefore, most likely procaspase-8 is recruited to FADD solely via $\alpha 1/\alpha 4$ of DED-A followed by sequential interaction mediated by $\alpha 2/\alpha 5$ of DED-B to form the caspase-8 activating chains crucial for its activation and hence cell death.

4.2.4 Discussion

Assembly of DISC is the key initiating event of the extrinsic apoptotic pathway, enabling activation of caspase cascade and hence cell death [179, 180]. The DISC formation is largely dependent on protein interaction domains of death-fold superfamily which includes the death domain and death effector domain [160]. The Fas-mediated DISC assembly involves highly specific interactions between the homologous domains of Fas receptor/FADD (DD-DD) and FADD/prodomain of procaspase-8/10 (DED-DED) [50, 181]. Despite knowledge of the DISC components, the molecular details and their mode of assembly especially at the level of DED-containing proteins have remained elusive. This is in part due to lack of comprehensive structural information on multivalent complexes of these proteins. Recently a new paradigm in the DISC assembly has been unravelled which further added to the complexity of this mechanism [155, 158]. It was shown that procaspase-8 is recruited not only through interaction with FADD but also by interacting with itself, resulting in formation of DED-chain which in turn facilitates its activation.

In this present work, we report for the first time the precise mode and the binding interface of DED/DED interaction promoting DED-chain formation, a prerequisite for death-fold mediated procaspase-8 activation. The binding surfaces required for self-association and

recruitment of caspase-8 to FADD was identified using *in silico*, extensive mutagenesis screening, biochemical and *ex vivo* studies. It is interesting to note that despite the strong structural homology in the death-fold superfamily which is characterized by the hall mark six α -helical bundled structure, each domain of this superfamily is constrained to interact with the proteins of its own subfamily. The stringent selectivity is proposed to occur due to the difference in non-covalent binding modes combined with the complementarity in their surfaces [160]. Although much is known about the CARD-CARD [70, 182, 183], PYD-PYD [184-186], DD-DD interactions [161, 187-190], the DED-DED interaction has remained elusive. Here, we provide detailed evidence for the specificity mediating the homophilic DED-DED interaction. One of the most important findings was the unexpected revelation that procaspase-8 DED-DED interaction involves association between helices $\alpha 1a/\alpha 4a$ and $\alpha 2b/\alpha 5b$ while the adaptor protein FADD DED self associates through the surfaces defined by helices $\alpha 1/\alpha 4$ and $\alpha 2/\alpha 3$.

Construction of a model for homodimeric tandem DED complex revealed that the domains pack tightly against each other through centrally localized hydrophobic residues. Notably, our *ex vivo* and *in vitro* data suggest that two conserved hydrophobic patches (*LXXF* and *FL-motif*) on the opposite surfaces are essential for homotypic inter-molecular DED interaction, and are possibly the major contributors of the interaction. Mutational analyses of varied protein-protein interfaces have shown that buried hydrophobic contacts lying at the center of any interface are responsible for the majority of interfacial binding energy and represent the hot-spot in the interaction [149, 150, 191]. The hydrophobic patches identified in DED self association are therefore likely to be the hot-spots or the functional epitopes responsible for DED-DED interaction. Furthermore, each of the DED domains, i.e. DED-A and DED-B, on an individual level behaves differently. Interestingly, both of these single domains did not show any self-association. This raises the possibility of individual DED

domains contributing unique structural modules toward self-association and their roles being non-interchangeable.

The homodimer caspase-8 DED interface fairly resembles an asymmetric type I interaction, which is one of the three distinct interaction types found in death-fold superfamily [192]. The type I interaction was originally described in Apaf1/procaspase-9 CARD complex, wherein helices $\alpha 2/\alpha 3$ of Apaf1 interact with $\alpha 1/\alpha 4$ of procaspase-9 CARD [182], and subsequently was found in almost all the other oligomeric signalling complexes (reviewed in [160, 168]). The model we built for oligomerization via the type I interface present in homodimer DED complex demonstrated helical arrangement with the propensity to form chain-like structure, which probably explains the ability of DED to form filaments *ex vivo*. In contrast to caspase-8 DED, the FADD DED self association was found to closely resemble the typical type I interaction pattern. Based on our data and the previously studied Apaf1-procaspase-9 CARD complex, it seems that the adaptor protein FADD preferentially engages in the homotypic interaction with the binding partners (another molecule of FADD or procaspase-8) through the surface defined by helices $\alpha 2/\alpha 3$ and $\alpha 1/\alpha 4$. Thus, although structurally similar, the members of this sub family have evolved distinct binding modes that construe their roles in the apoptotic signaling pathways. However, they share certain hot spots on the binding surface critical for mediating the interaction.

FADD DED-mediated self-association is reported to be essential for competent signaling of the DISC [67, 174]. It is attributed to perform two main functions, self-association and interaction with tandem DED proteins such as caspase-8 and MC159 vFLIP which would promote or inhibit cell death respectively. Our data demonstrate that single mutations such as F25Y or E51A that disrupt FADD self-association indirectly also affect its interaction with procaspase-8, thus highlighting that these two functions are not mutually exclusive. Protein oligomerization is frequently observed in the assembly of complexes mediated by members of

the death-fold superfamily including the Apoptosome [193], PIDDosome [59], Myddosome [58] complexes. For instance, apoptosome is built using heptameric Apaf-1, which has a much higher affinity for procaspase-9 than an isolated Apaf-1 CARD [194]. The ability of FADD to self-associate has been previously speculated to create the structural framework necessary for procaspase-8 recruitment and activation [67]. Overall, it implies that FADD has a central role in the regulation of death-receptor mediated signalling and must be under tight check in the cell. In agreement to this assumption, our studies unequivocally demonstrate that FADD/FADD interaction is very weak, which may possibly explain as to why although constitutively present in the cell they do not spontaneously self-associate to recruit and activate caspase-8 in the absence of stimuli (such as death receptor/FADD ligation or increase in local concentration upon ectopic expression). The oligomerization would probably be necessary to increase the binding avidity and/or form the correct arrangement of FADD via conformational change to facilitate recruitment of procaspase-8.

Although FADD acts as the key player in regulation of the extrinsic apoptotic pathway, fate of the cell is largely dependent on the activation of zymogenic caspases. Until recently, most studies concentrated on the role of cFLIP in regulation of caspase activation and its involvement in life or death decision [195, 196]. However with the discovery of the DED-chain [94, 158], another important constraint has emerged for DISC-mediated apoptosis. Furthermore, it has been previously reported that both the two DED domains of procaspase-8/procaspase-10 are required for successful formation of the DISC [197-200]. Therefore, it was critical to understand the basis of DED assembly and its role in caspase-8 activation. Since DED-DED interaction involved typical interaction between conserved binding surfaces, we proposed structure-based three possible models to explain the initial recruitment of caspase-8 to self-associated FADD platform. With the experimental evidences, we are in support of the model whereby procaspase-8 is recruited to oligomerized FADD platform exclusively through

DED-A which initiates subsequent formation of caspase activating DED-chain to induce proximity induced cleavage and hence cell death. Another interesting aspect of our model is that it may now explain the importance and the requirement for both the domains in DISC assembly. In accordance with our observations, several mutations in procaspase-8 DED have been identified in colon and gastric carcinoma which although do not affect its binding with FADD but interfere with procaspase-8 activation hence promoting inappropriate cell survival [199-202]. Thus, in context of the normal cell physiology, caspase-8 activation is tightly regulated at various levels of the DISC assembly so as to prevent inadvertent cell death as well as inappropriate cell survival.

CHAPTER-4.3

4.3 Biochemical and biophysical characterization of E2 transactivation domain

Background:

In chapter-4.1, we showed that E2 via $\alpha 2$ helix of transactivation domain directly binds to the DED-containing prodomain of caspase-8 mediating procaspase-8 activation and hence cell death. While as discussed in the previous chapter-4.2, caspase-8 activation in the classical cell death pathway occurs through homotypic interaction between the death-fold domain proteins. Therefore, we were intrigued as to how E2 in spite of being a non death-fold protein act as a potent binding partner of caspase-8. To understand this we carried out biophysical and biochemical characterization of E2 transactivation domain.

4.3.1 Introduction

E2 proteins characterized from numerous human and animal serotypes have three distinct modules; an amino-terminal transactivation domain (TAD) of ~200 amino acids followed by a flexible, proline-rich hinge region and a carboxyl-terminal DNA binding domain (DBD) of ~100 amino acids. The TAD is a protein interaction module that binds to the viral helicase E1 [203] and to various host proteins such as topoisomerase I [4], Brd4 [204] and procaspase-8 [1] to name a few. Several crystal structures of the E2 TAD, either alone (free form) or in complex with the C-terminal helicase domain of E1, or with a Brd4 peptide, have been reported [141, 142, 205-208]. These structures revealed that E2 TAD is L-shaped comprising two sub-domains, N1 and N2, separated by two consecutive single helical turns. The DBD of E2 is reported to dimerize (K_d in nanomolar range), and binds tightly to viral DNA sites with consensus sequence of ACCGN₄CGGT, where N represents any base pair [209, 210]. Although the main dimerization interface is located within the C-terminal DBD, the arrangement of the N-terminal TAD within crystal lattices of some of the structures mentioned above suggest that this domain can also homodimerize, however, with dissociation

constant in the micromolar range [142, 207]. For E2 from HPV16, the dimerization of TAD was found to involve residues from helices $\alpha 2$ and $\alpha 3$ of N1 domain and residues 142–144 from N2. Detailed examination of the dimer interface revealed hydrogen bonding and ionic interaction between side chains of R37 with E80 of second subunit to be primarily responsible for stabilization of the dimer [211]. The TAD–TAD interaction was shown to occur between E2 dimers (*forming inter-dimers*); serving as a key to the formation of DNA loops [211, 212]. This in turn leads to transcription activation by bringing together distally bound transcriptional factors. However, with respect to HPV18 E2 TAD, the possibility of TAD dimerization is still elusive. The available crystal structure has 65 N-terminal residues (two-thirds of N1 domain) missing which lacks key amino acids reported for dimerization [213]. Sequence of E2 TAD is highly similar among different virus types with ~ 33 of the total 201 residues completely or partially conserved. Furthermore, the conservation or conservative substitution of residues at dimer interface (**Fig. 4.1.15**) suggests that E2 TAD dimerization might be the intrinsic property of all known papillomaviruses. Consistent with this, apart from human papillomaviruses, E2 TAD dimerization is also reported for bovine papillomaviruses [208, 214].

Based on all these information and our present findings, we speculate that the surface defined by helices $\alpha 2$ and $\alpha 3$ of E2 TAD enables it to act as a pseudo death-fold protein, with potential to homodimerize and interact with death-fold domain of caspases via the typical (helix-helix)/(helix-helix) association. We therefore characterized complete HPV18 E2 TAD so as to delineate its capability of homodimerizing *in-vitro* using molecular modeling, mutagenesis and bio-physicochemical approaches.

4.3.2 Experimental Procedures

4.3.2.1 Proteins expression and purification – E2 TAD wild-type and mutant (R41A, W42A, and R41A/W42A) expression vectors have been described in chapter-4.1. For protein expression, the E2 TAD plasmids were co-transformed with pMALc5 empty vector in *E.coli* BL21 (DE3) cells. Cells were grown in the medium supplemented with 2 µg/ml of arabinose (to induce ara-operon driven expression of E2 TAD) at 37 °C till the optical density ($A_{600\text{ nm}}$) reaches 0.5. Cells were further induced with 0.4 mM IPTG to induce expression of MBP and further cultured at 18 °C for 20 h. Next, the bacteria were harvested by centrifugation, re-suspended in lysis buffer (20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ pH 7.8, 200 mM NaCl, 0.1 mM DTT, 1% glycerol, 10 mM imidazole, and 1X protease cocktail inhibitor) and subjected to sonication. After centrifugation at 16,000 rpm for 30 min, the lysate was passed through pre-equilibrated Ni-IDA resin for 1 h at 4 °C. The column was washed three times with binding buffer (20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ pH 7.8, 200 mM NaCl, 20 mM β -mercaptoethanol, 0.1% TritonX-100, and 10 mM imidazole) and the bound proteins were later eluted in the binding buffer using imidazole gradient (25 - 500 mM). The proteins were visualized on 12% SDS-PAGE; relatively pure proteins were pooled and concentrated to 2 ml volume using 10 kDa centricon (Millipore). The protein was further purified and characterized for the oligomeric property by size exclusion chromatography using Superdex 75 pre-grade or analytical grade columns respectively. Integrity and purity of protein was again checked by 12% SDS-PAGE and the identity was confirmed with western blotting using anti-E2 antibody (sc-26939, Santa Cruz Biotechnology, Inc.).

4.3.2.2 Fluorescence and CD spectroscopy – Urea denaturation was performed in 20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ buffer pH 7.8, 100 mM NaCl, 20 mM β -mercaptoethanol for the protein concentrations described in the Figure 4.3.4. For the equilibrium unfolding studies, protein

samples were incubated for 12 h to equilibrate at the appropriate final concentrations at 25 °C prior to acquisition. Fluorescence spectra measurements were recorded on Fluorolog-3 spectrofluorometer (Horiba Scientific). The samples were excited at 280 or 295 nm; the slit width was equal to 4 nm during the excitation and 5 nm for the emission. The signal was acquired for 0.3 s integration time, and the increment of wavelength was equal to 0.5 nm. The fluorescence signal for the protein was blank corrected by subtraction of the buffer signal for the solvent alone. For renaturation experiments, the protein was incubated in 8 M urea-containing buffer. Post incubation for 5 h at 25 °C, the protein was allowed to refold by diluting it into phosphate buffer such that the final urea concentration was as indicated in the figure legends. The samples were incubated at 25 °C for 12 h before measurements. Far UV CD spectra of urea denatured protein samples were acquired using JASCO J 815 spectropolarimeter from 250 nm to 210 nm owing to increase noise at shorter wavelengths. Spectrum was the average of 3 scans at a 50 nm/min scan speed, and the buffer baselines were subtracted.

4.3.2.3 Analysis of the equilibrium unfolding profile – E2 TAD wild-type followed a three-state folding pathway with the formation of dimeric intermediate given as follows: $N_2 \rightleftharpoons I_2 \rightleftharpoons 2U$. The data was normalized and fitted in the equations 13-18 described in *Materials and Methods* section to obtain the parameters $\Delta G_1^{\text{H}_2\text{O}}$, $\Delta G_2^{\text{H}_2\text{O}}$, m_1 and m_2 .

4.2.2.4 Oligomeric property analysis – Glutraldehyde cross-linking and dynamic light scattering (DLS) experiments were performed by following the protocol as described under *Materials and Methods*.

4.3.2.5 Molecular docking of monomeric HPV18 E2 TAD to generate E2 TAD homodimer – HPV16 E2 protein TAD domain (PDB id. 1DTO) crystal structure solved at high resolution of 1.90 Å, having C2 biological symmetry, and 47.4% sequence similarity with HPV18 E2 TAD, turned into a base for biological dimeric symmetrical assembly. To obtain the dimer, the modeled monomeric form of HPV18 E2 TAD was superimposed on HPV16 E2 TAD protein using PyMol Version 1.5.0.4 (Schrödinger, LLC). In order to verify correct assembly of dimeric interface of modeled structure, molecular docking of the monomer was performed with ClusPro 2.0 [163], without defining the active site residues. ClusPro 2.0 performs rigid body docking of protein complexes based on PIPER (FFT based protein docking program with pair potentials). In ClusPro 2.0 rigid body docking, 70,000 rotations of ligand molecule with respect to receptor molecule within the range of 9 Å were performed to generate cluster of poses. Furthermore, the best docked structure generated from clusters of poses was structurally aligned to the dimer obtained from superimposition to finalize the correct assembly of the dimeric interface. The final structure was subjected to PRCG (Polak Ribierre conjugate gradient) energy minimization with PRIME Minimization module of Schrödinger 2014 suite (Schrödinger LLC, New York, LLC) using water as implicit solvent, OPLS (Optimized Potential for liquid simulations) force field and minimization steps up to 5000 steps. Subsequently E2 TAD dimer was prepared and optimized with protein preparation wizard of Schrödinger 2014 suite (Schrödinger LLC, New York, LLC) and further subjected to molecular dynamic (MD) simulation.

4.3.2.6 MD Simulation of modeled E2 TAD dimer – Molecular dynamic studies of optimized E2 dimer was carried out with Gromacs v 4.5.6 molecular dynamics package using OPLS-AA (Optimized Potential for liquid simulations-all atom) force field to evaluate the dynamics and strength of interaction within the dimer interface assembly. The dimer complex

was then solvated with SPC216 (Simple Point Charge water) water solvent model into a cubic box of 1.0 nm from the molecule to the edge of the box. In addition PME (Particle Mesh Ewald) electrostatics was used for long range electrostatic interactions and PBC (Periodic Boundary Condition) was applied in all the directions. The system was further neutralized with mono atomic ions to achieve total zero charged system. The tolerance was set up to 1000 kcal/mol and 50,000 energy minimization steps was carried out using Steepest Descent and subsequently Conjugate Gradient minimization method. For short range electrostatic, coulomb cut-off was set to 1.0 nm with Fourier spacing of 0.16 nm. All bond angles were constrained with LINCS (Linear Constraints Solver) holonomic algorithm, though geometry of water molecules was constrained with SETTLE algorithm. Parinello-Rahman barostat was used to set the pressure (1 bar) of the system, while V-rescale thermostat was applied to modulate the temperature (310 K) of the system. Moreover Position restrains algorithm was applied for both NVT (canonical ensemble) and NPT (isothermal-isobaric ensemble) MD of 100 ps. The position restrained system was well equilibrated in accordance with both the NVT and NPT MD of 100 ps evidenced by the results. The system was later run for 50,000 ps (50 ns) MD production. The coordinate frames of every 2 ps of the 50 ns MDS trajectory were saved for further analysis. The entire computation was performed on BRAF-HPC facility with 6x Intel Xeon running on 2.67 GHz processor on Cent-OS linux-based operating system. The simulated models were verified and validated using structural analysis and verification server (SAVES) which include PROCHECK [215], ProSA-web [216], Verify-3D [217] and ERRAT verification algorithms [218].

4.3.2.7 Binding free energy calculations – E2 TAD – TAD and E2 TAD – DED binary protein complex binding energies were calculated with the help of MM-GBSA (Molecular

Mechanics-Generalized Born model Solvent Accessibility) module provided by Schrödinger Package 9.35. MM-GBSA is an efficient tool for calculation of binding energy based on difference of Gibbs free energy of protein-protein complex and complex free protein molecule (receptor and ligand). The average interaction energies of receptor and ligand are usually obtained by performing calculations on an ensemble of uncorrelated snapshots collected from an equilibrated MD simulation. The final equation for the solvated binding free energy $\Delta G_{\text{bind, solv}}$ can be calculated as: $\Delta G_{\text{bind, solv}} = E_{\text{complex (minimized)}} - E_{\text{ligand (minimized)}} - E_{\text{receptor (minimized)}}$

4.3.2 Results and Discussion

4.3.2.1 In silico analysis of E2 TAD dimerization

The crystal structure of N-terminal TAD of HPV16 E2 protein revealed a dimerization surface constituting residues from helices $\alpha 2$ and $\alpha 3$. To assess whether HPV18 E2 TAD could homodimerize, a structural model was generated as described in the *experimental section*. The overall quality factor of the model was assessed by ERRAT analysis which was found to be 82%, implying good accuracy of the predicted structure. The stereochemistry of the refined model (PROCHECK analysis) revealed that 100% of residues were situated in the most favorable region and additional allowed regions, whereas none of the residues fell in the generously allowed or disallowed region of the Ramachandran plot (**Fig. 4.3.1A**). ProSA-web evaluation provided a compatible Z score value of -7.9 (**Fig. 4.3.1B**), which is well within the range of native conformations of crystal structures. The overall residue energies of the 3D model were largely found to be negative (**Fig. 4.3.1C**). These observations indicated that the model quality was reliable and could be considered for further analysis.

Superposition of HPV16 and 18 E2 TAD dimers showed an RMSD of less than 0.8 Å (**Fig. 4.3.2A**). The dimerization interface of HPV18 E2 involved residues Q38, R41, W42,

E52, Q80, M81, Q84, Q88, K92 from the helices $\alpha 2$ and $\alpha 3$ (**Fig. 4.3.2B**). Notably, residues R41, W42, Q84 and Q88 seem to play a central role in this interaction (**Fig. 4.3.2C**). Furthermore, it is interesting to note that the TAD – TAD interface shared few residues that formed part of the procaspase-8 binding surface (reported in chapter 4.1). Therefore, we decided to explore the wild-type and $\alpha 2$ helix mutants (R41A, W42A, and R41A/W42A) to test for their ability to homodimerize.

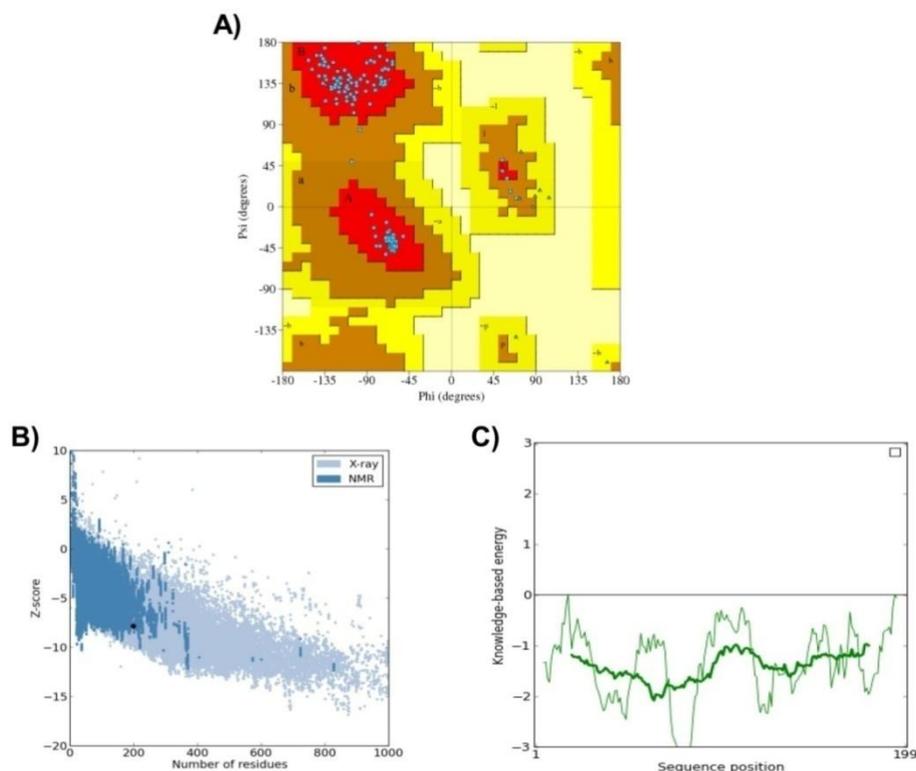
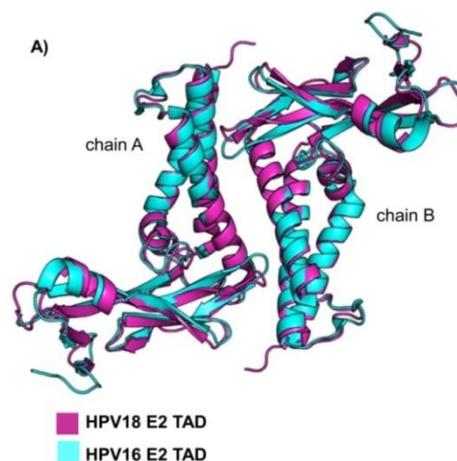


Figure 4.3.1 Validation of the E2 TAD dimer. A) Ramachandran plot of the model (red, dark yellow, and light yellow regions represent the most favored, allowed, and generously allowed regions). (B) ProSA-web Z-scores of all protein chains in PDB determined by X-ray crystallography (light blue) and nuclear magnetic resonance spectroscopy with respect to their length. The Z-score of E2 was present in the range represented in black dot. (C) Energy plot for the predicted E2 model of HPV18.



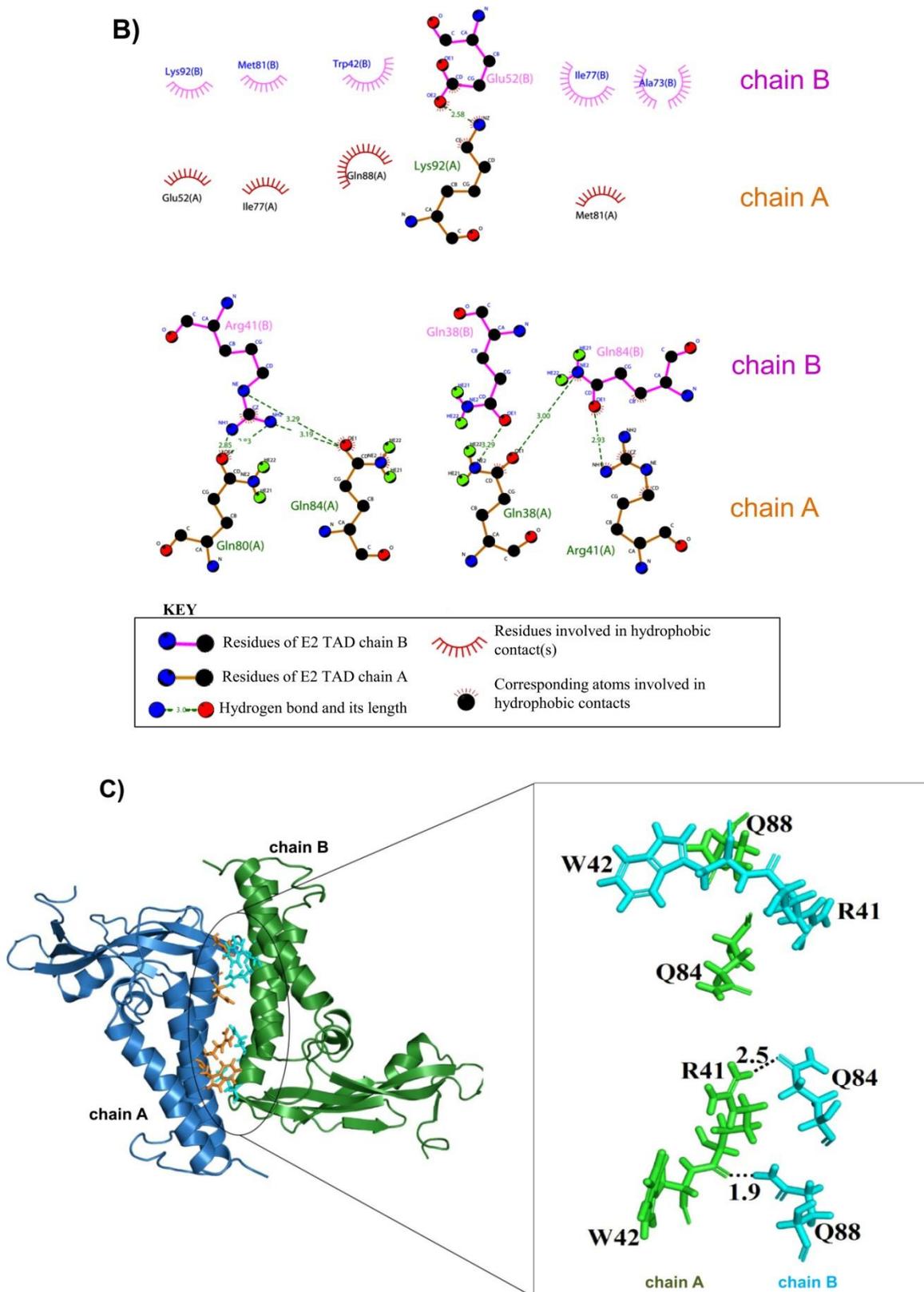


Figure 4.3.2 Homodimeric model of HPV18 E2 TAD. A) Superposition of HPV18 and 16 E2 TAD. Modeled HPV18 E2 TAD homodimer (residues 1-199) superimposed on HPV16 E2 TAD structure (PDB id. 1DTON) shown in magenta and cyan colours respectively. B) Dimplot of the HPV18 E2 TAD dimer interface. C) A close-up view of the interface with selected side chains from each molecule shown in stick model. The residues from chain A surface are shown in green and chain B in cyan colour.

4.3.2.2 *In vitro* characterization of wild-type and mutant E2 TAD

As the first step, wild-type and mutant E2 TAD were expressed in *E.coli*, purified by immobilized metal-ion affinity chromatography followed by size-exclusion chromatography. The gel filtration profile presented in **Fig. 4.3.3A** show that the wild-type and mutant proteins exist predominantly as a monomer, however, except for the double mutant, a small peak corresponding to the dimer was also observed. To confirm the presence of the dimeric species, 20 µg of wild-type and mutant proteins were cross-linked with glutaraldehyde and run on SDS-PAGE. It was observed that wild-type and single mutants show mixture of monomeric and dimeric species, however, the concentration of the former being higher; whereas the double mutant exists only in the monomeric form (**Fig. 4.3.3B**). To further substantiate these findings, we determined the heterogeneity of the purified proteins using DLS (**Fig. 4.3.3C**). We observed that the average hydrodynamic radii of E2 wild-type and single mutants were larger compared to the double mutant. Based upon the radii, the calculated molecular weight of the proteins is given in **Table 4.3.1**. For wild-type and the single mutants, the calculated molecular weight closely resembled the theoretical molecular weight of a dimeric E2 TAD, whereas the double mutant showed a homogenous monomeric population.

Collectively, these data suggest that E2 TAD can coexist as monomer and dimer in solution, however, with the equilibrium more favorable towards the monomeric population. Furthermore, both the residues R41 and W42 might be critical for homodimerization which is in contrast to the HPV16 E2 TAD, where R37A alone resulted in destabilization of the dimer [219]. Based upon the structure of HPV16 E2 TAD dimer (PDB id. 1DTO) and our modeled structure, we assume that the possible difference could be due to the type of interactions at the dimer interface (please refer **Table 4.3.2** and **4.3.3**). In case of HPV16 E2 TAD, R37 by itself makes extensive hydrogen bonding and ionic interactions with E80 to stabilize the

symmetrical dimer interface, while in case of HPV18 E2 TAD, the glutamic acid is replaced with glutamine which although makes hydrogen contact with R41 but does not participate to form a stabilizing salt bridge. The hydrogen bonding between W42 and Q88 might therefore be probably required to provide favorable energy to form the dimer. However, further mutagenesis studies would be required to test this assumption and to obtain clear view on the role of these residues in dimerization.

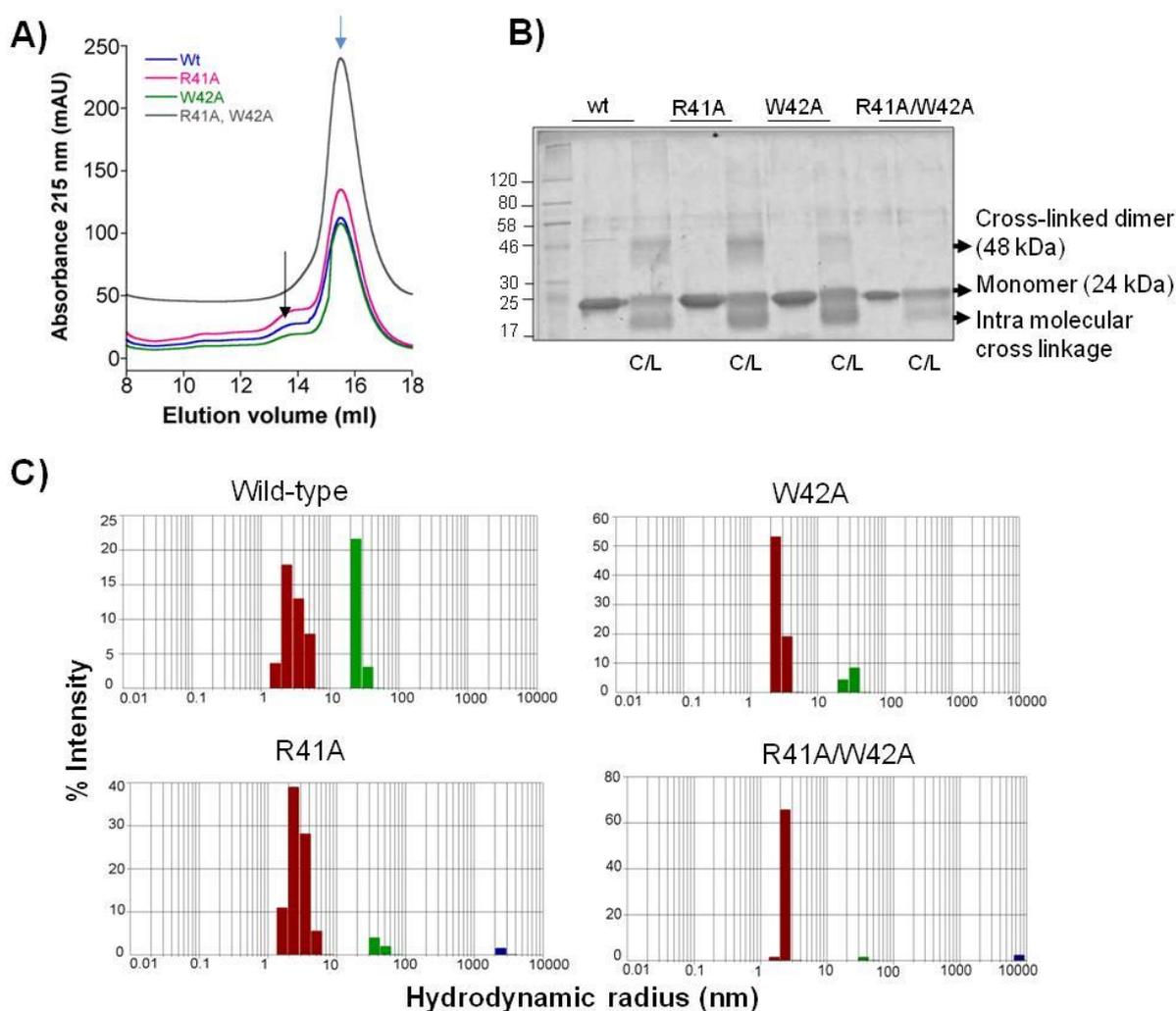


Figure 4.3.3 Determination of the oligomeric state of E2 TAD wild-type and mutant proteins. A) Size exclusion chromatography elution profile of the wild-type and mutant (R41A, W42A, and R41A/W42A) E2 TAD proteins using a Superdex 200 10/300 (analytical grade) column. A major peak (blue arrow) corresponding to the calculated molecular weight of a monomer (24.5 kDa) and a minor peak (black arrow), except for the double mutant, related to the dimeric species (49 kDa). B) Glutaraldehyde cross-linking reaction product resolved on SDS polyacrylamide gel. Cross-linked dimeric species is visible in case of wild-type and single mutant proteins. C/L represents glutaraldehyde treated Cross-Linked sample. Note that the intensity of the cross-linked dimer is lesser

for the W42 mutant compared to wild-type and R41A mutant. A band corresponding to size lower than 24 kDa in the cross-linked lane is due to the intra molecular cross linkage of the protein resulting in increased mobility. C) Dynamic light scattering of mutant and wild-type E2 TAD proteins. The vertical co-ordinate axis corresponds to the relative percent intensity and the horizontal axis corresponds to the hydrodynamic radius (on a logarithmic scale). The histogram in red for the wild-type and single mutants shows the average size distribution of the particle of an average molecular mass of approx. 50 kDa, while a molecular mass of approx. 25 kDa for the double mutant. The bar in green corresponds to the size distribution of an aggregate and in blue for dust particle.

Table 4.3.1 Dynamic light scattering analysis of E2 wild-type and mutant proteins.

E2 TAD	Hydrodynamic radius ^a	% Polydispersity ^b	Calculated molecular weight (kDa)	Theoretical molecular weight (expasy protparam) (kDa)	
				monomer	dimer
Wild-type	2.98	34.9	55.4	24.2	48.4
R41A	2.93	31.3	52.7	24.1	48.2
W42A	2.90	18.4	51.1	24.1	48.2
R41A,W42A	2.42	4.5	26.2	24.0	48.0

^aMean hydrodynamic radius derived from the measured translational diffusion coefficient using the Stokes-Einstein equation

^bPolydispersity divided by the hydrodynamic radius

Table 4.3.2 Detailed list of interactions at HPV18 E2 TAD-TAD interface evaluated using the *Protein Interactions Calculator* with the default cut-offs.

Type of interaction	TAD (chain-A)	TAD (chain-B)	Distance (Å)
Hydrogen bonds			
Side chain-side chain	NH1 Arg41	O ^{ε1} Gln84	2.14
	N ^{ε2} Trp42	O ^{ε1} Gln88	3.37
	O ^{ε1} Gln88	N ^{ε2} Trp42	3.37
	O ^{ε1} Gln84	NH1Arg41	2.14
Salt bridges			
	Lys70	Asp146	
	Asp146	Lys70	
Hydrophobic interactions (within 5Å)			
	Phe48	Ala147	
	Tyr66	Ala147	
	Ile77	Met81	
	Met81	Ile77	
	Ala147	Phe48	
	Ala147	Tyr66	

Table 4.2.3 Detailed list of interactions at HPV16 E2 TAD-TAD interface evaluated using the *Protein Interactions Calculator* with the default cut-offs.

Type of interaction	TAD (chain-A)	TAD (chain-B)	Distance (Å)
Hydrogen bonds			
Main chain-side chain	NH2Arg37	O Leu77	2.97
Side chain-side chain	NH2Arg37	O ^{ε1} Glu80	2.61
	N ^{ε2} Gln76	O ^{ε1} Glu80	2.88
	O ^{ε1} Glu80	N ^{ε2} Gln76	2.88
	O ^{ε1} Glu80	NH2 Arg37	2.61
Salt bridges			
	Lys34	Glu80	
	Arg37	Glu80	
	Lys66	Glu141	
	Lys66	Glu142	
	Glu80	Lys34	
	Glu80	Arg37	
	Glu141	Lys66	
	Glu142	Lys66	
Hydrophobic interactions (within 5Å)			
	Ala41	Ala143	
	Tyr44	Ala143	
	Ile73	Leu77	
	Leu77	Ile73	
	Leu123	Leu48	
	Ala143	Ala41	

4.3.2.3 Unfolding studies

Equilibrium unfolding is an excellent tool to understand the driving force of the interaction and whether dimerization is a folding process. Solvent-denaturation represents a valuable alternative to conventional methods, such as size-exclusion chromatography and DLS, for measuring dissociation constants for homodimeric proteins and provides additional information on the energetically linked unfolding reactions [220]. We therefore studied the dependence of the unfolding of E2 TAD induced by urea. In these experiments, we monitored changes in secondary structure by circular dichroism at 228 nm and changes in tertiary structure by average emission fluorescence, following excitation at 280 nm or 295 nm. The

data show a small pre-transition for the first step of unfolding which occurs between 0 and ~2.5 M urea, which is followed by a second cooperative transition between ~3 M and 7 M urea. All the three spectroscopic signals showed similar trends in the unfolding data (Fig. 4.3.4). We also observed that the folding transitions are reversible (Fig. 4.3.4D). Therefore, these results suggest that E2 TAD follows a three-state folding pathway.

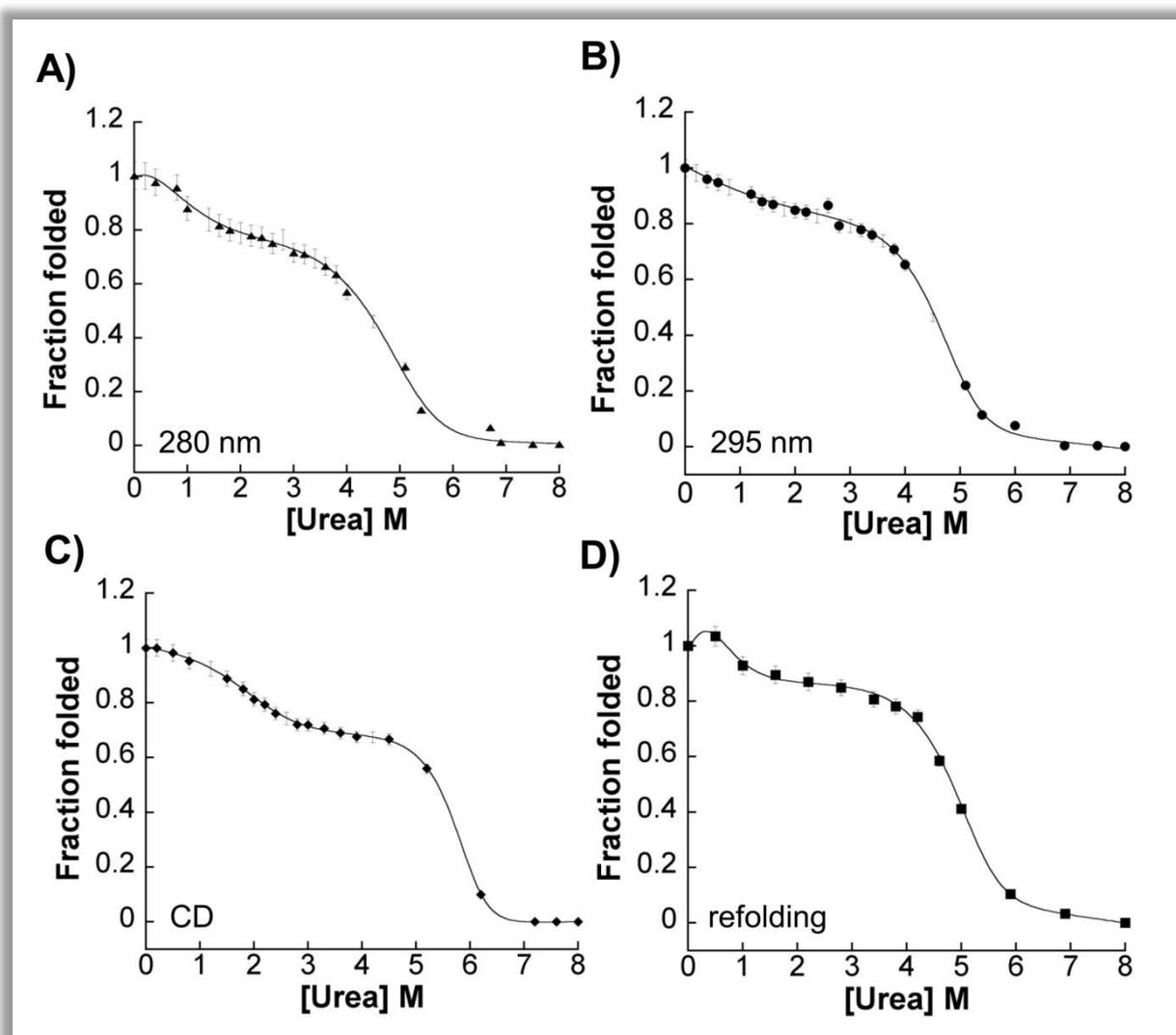


Figure 4.3.4 Equilibrium unfolding of E2 TAD wild-type. Unfolding was measured by fluorescence emission with excitation at either 280 nm (A), or at 295 nm (B) and by CD at 228 nm (C). For the fluorescence and the CD measurements the protein concentration was 1.5 μ M and 5 μ M respectively. D) Refolding of the protein measured by fluorescence emission with excitation at 280 nm to confirm the reversibility. Error bars show standard error for at least three unfolding curves. The data were fit (solid lines) using three-state folding model as described under *Materials and Methods* using Igor Pro 6.03A.

Because we observed that E2 TAD homodimerizes at higher concentrations, we went ahead to examine the effect of protein concentration on equilibrium unfolding. As shown in the **Fig. 4.3.5.**, there was a shift in the midpoint of the second transition as the concentration of the protein increased beyond 1.5 μM . However the first transition was concentration independent, and the C_m value (concentration of urea at which 50% of the protein is unfolded) was constant at 1.6 M urea suggesting that E2 TAD follows a three-state folding pathway with the formation of a dimeric intermediate. Based on this information, the global fitting of the data to a three-state folding model, discussed in chapter-3, was carried out to obtain the conformational free energies ($\Delta G^{\text{H}_2\text{O}}$) and the cooperativity indices, or m -values for each step of unfolding (**Table 4.3.4**). The results demonstrate that wild-type E2 TAD is quite stable; with total conformational free energy of $18.8 \text{ kcal.mol}^{-1}$. By using these values for $\Delta G_1^{\text{H}_2\text{O}}$, $\Delta G_2^{\text{H}_2\text{O}}$, m_1 and m_2 , we calculated the equilibrium distribution of the species for all the protein concentrations (1.5 μM , 2.5 μM , and 3 μM). As shown in the **Fig. 4.3.6**, between 0 M and 3.5 M urea, there is a decrease in the population of native dimer and a concomitant increase in the population of the dimeric intermediate, I_2 . The mid-point of the transition is 1.6 M urea, consistent with the experimental results shown in Figure 4.3.5. The relative concentration of the dimeric species is dependent on the protein concentration and reaches a maximum at ~ 3.1 M urea. The second transition, I_2 to U, has midpoint 4.3 M, 4.7 M and 4.9 M for the concentration of the proteins indicated in the figure legend. By observing this second-order transition both as a function of denaturant and protein concentration, we could determine the dissociation constant by using the equations 1 and 17 and the $\Delta G_2^{\text{H}_2\text{O}}$ and m_2 values. The calculated equilibrium constant, K_2 at 4.5 M urea was found to be $\sim 1.8 \mu\text{M}$. To further ascertain the equilibrium dissociation constant value, it would be necessary to perform analytical ultracentrifugation, which we could not do due to unavailability of the facility. Nonetheless, in the literature [221], dissociation constant for

HPV16 E2 TAD is reported to be K_d , $8.1 \pm 4.2 \times 10^{-6}$ M, which in close agreement to our data.

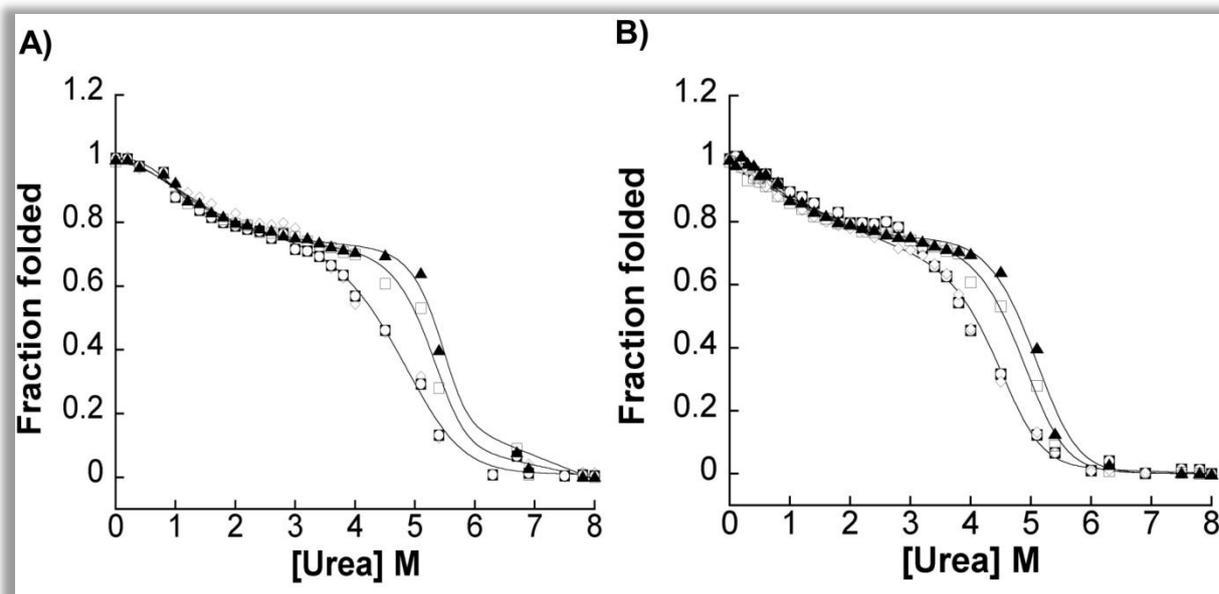


Figure 4.3.5 Concentration dependence of the equilibrium unfolding of E2 TAD wild-type. Unfolding monitored by fluorescence emission with excitation at 280 nm (A) and 295 nm B), at protein concentrations of 1 μ M (\blacksquare), 1.5 μ M (\diamond), 2.5 μ M (\square), 3 μ M (\blacktriangle). For clarity, error bars are not shown for all the data sets.

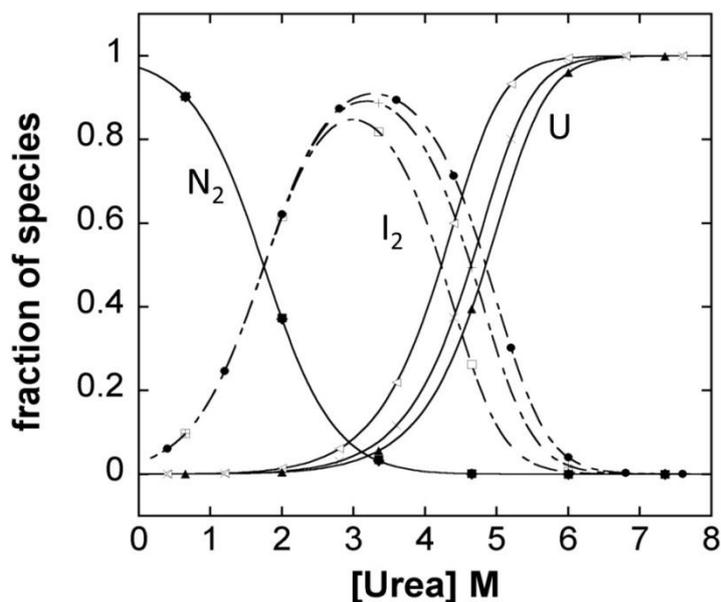


Figure 4.3.6 Fraction of species as a function of urea concentration in the unfolding process. The fraction of native, dimeric intermediate and unfolded protein was calculated as a function of urea concentration. The protein concentrations were 1.5 μ M, 2.5 μ M and 3 μ M. 'N₂' refers to the native protein 'I₂' is the dimeric intermediate and 'U' refers to the unfolded species.

Table 4.2.4 Mean parameters for the unfolding of E2 TAD wild-type, as monitored by its fluorescence intensity and calculating the average emission wavelength.

Parameter	Mean \pm S.E.
m_1 (kcal.mol ⁻¹ .M ⁻¹)	0.98 \pm 0.07
$\Delta G_1^{\text{H}_2\text{O}}$ (kcal.mol ⁻¹)	1.82 \pm 0.02
m_2 (kcal.mol ⁻¹ .M ⁻¹)	1.97 \pm 0.06
$\Delta G_2^{\text{H}_2\text{O}}$ (kcal.mol ⁻¹)	16.9 \pm 0.1

The parameters were obtained by fitting Equation 13-17 to the experimental data. Each entry corresponds to the mean \pm S.E. (standard error) of three independent experiments at three different concentrations of the protein: 1.5, 2.5, and 3 μ M.

Table 4.2.4 Energy components contributing to the binding of the E2 TAD – E2 TAD homodimer and E2 TAD – DED heterodimer complexes in terms of MM-GBSA module provided by Schrödinger Packages 9.35. $\Delta G_{\text{bind}} = E_{\text{complex}}$ (minimized) - E_{ligand} (minimized) - E_{receptor} (minimized). All the values are in kcalmol⁻¹.

E2 TAD – E2 TAD complex

Binding Energy components	Complex	Receptor (E2 TAD)	Ligand (E2 TAD)	ΔG
H-bond	-269.3	-113.9	-149.66	-5.7
Electrostatic	-13732.1	-7354.8	-6510.13	132.9
Covalent	2258.8	1008.3	1230.23	20.3
Pi-pi packing	-38.1	-23.6	-14.32	-0.2
Lipophilic	-2457.4	-1141.7	-1280.63	-35.1
Solvation energy-GB	-2743.5	-1275.7	-1355.64	-112.2
van der Waals	-1747.3	-846.8	-800.25	-100.3
Self-contact correction	-239.2	-159.6	-80.74	1.1
Total Binding Energy	-18968.8	-9907.7	-8961.1	-99.2

E2 TAD – DED complex

Binding Energy components	Complex	Receptor (E2 TAD)	Ligand (DED)	ΔG
H-bond	-168.7	-68.3	-97.3	-3.0
Electrostatic	-13217.6	-6465.5	-6879.5	127.4
Covalent	2159.7	1267.8	880.6	11.3
Pi-pi packing	-43.1	-7.5	-23.3	-12.2
Lipophilic	-2148.3	-1002.8	-1107.1	-38.3
Solvation energy-GB	-2942.4	-1490.8	-1349.7	-101.9
vdW	-691.4	179.3	-822.9	-47.8
Self-contact correction	-251.6	-117.5	-142.6	8.5
Total Binding Energy	-17303.4	-7705.6	-9541.9	-55.9

4.3.2.4 E2 TAD homo vs. E2 TAD – DED heterodimerization

Earlier in chapter-4.1, we demonstrated that interaction of procaspase-8 with E2 involves the interface defined by helices $\alpha 2/\alpha 3$ of E2 TAD, suggesting that E2 TAD homo- and E2 TAD-DED hetero-dimerization share the same surface. We therefore decided to compare the relative binding affinities of E2 TAD homodimer and the E2 TAD – DED heterodimer complexes. To estimate the relative binding affinities of these complexes, we performed binding free-energy analysis using MM-GBSA approach. This method has been used in a number of recent studies to investigate protein–protein interactions, and the approach is generally well correlated with experimental results [222-224]. The absolute values calculated are not necessarily in agreement with experimental binding affinities. However, the ranking based on the calculated binding energies (MM-GBSA ΔG_{Bind}) can be expected to agree reasonably well with ranking based on experimental binding affinity [225]. To determine the binding interactions between these complexes, the conformational data that were generated from the molecular dynamics simulations were used to perform post-trajectory implicit solvent binding free-energy calculations. The free-energy is expressed as the difference of the free energy of the complex with that of the receptor and ligand, averaged over a number of snapshots collected from the trajectory. The binding free energies calculated for each of the contributing terms are shown in **Table 4.3.5**. The inspection of free-energy components of the complex revealed that the packing and van der Waals contribution of the solvation free energy is more favorable for E2-DED complex as compared to the E2 TAD homodimer. The calculated relative binding free energies for E2 TAD homodimeric ($\Delta G_{\text{Bind}} = -55.9 \text{ kcal.mol}^{-1}$) and E2 TAD – DED complexes ($\Delta G_{\text{Bind}} = -99.2 \text{ kcal.mol}^{-1}$) clearly demonstrate favorable and higher binding affinity for E2 – DED complex, which very well corroborate with our experimental data as discussed earlier. However, further *in vitro* experiments such as

isothermal titration calorimetry are desirable to accurately determine the binding affinities of these complexes.

Collectively, results from the present study suggest that E2 TAD forms a weak dimer via the surface defined by helices $\alpha 2/\alpha 3$, and the dimeric species cannot be significantly populated unless the protein is in micromolar concentration. Strikingly, the E2 – procaspase-8 DED complex is highly stable even at lower protein concentrations and in buffer conditions ranging from low to high ionic strength (reported in chapter-4.1). Furthermore, in chapter-4.2, we have established that the classical death-fold interaction involves (helix/helix) – (helix/helix) association between the death-fold domains. Therefore, based on all these results we propose that surface defined by helices $\alpha 2/\alpha 3$ of E2 TAD enables it with the ability to act as a pseudo death-fold adaptor protein having potential to homodimerize (lower affinity) and interact with typical death-fold domain of caspases (higher affinity), to induce caspase-8 oligomerization followed by its activation and cell death.

CHAPTER-5

Serine protease HtrA2

5.1 *Introduction*

HtrA proteins belong to a unique family of multidomain serine proteases (S1, chymotrypsin family) that are conserved from prokaryotes to humans. They are well-known for their complex structural organization that is reflected in their involvement in numerous cellular processes such as protein quality control, unfolded protein response, cell growth and apoptosis. This multitasking ability associates them with a number of pathophysiological conditions such as cancer, arthritis, and neurodegenerative disorders hence making them therapeutically important [2, 74-76]. Interestingly, despite diversity in functions and low sequence identity, their overall structural integrity is evolutionarily maintained that comprises a serine protease domain and one or more C-terminal PDZ domains arranged in a pyramidal oligomeric assembly ranging from trimer to 24-mer structure. [9]. Although there is overall conservation in the protease structure, significant diversity at the N-terminus such as signal sequences, transmembrane regions and/or additional domains construe their cellular localization and functionality. These variations along with intricate conformational dynamics lead to subtle differences and hence provide basis for their specificity and distinct behavior [226].

Human HtrA2, one of the key members of this family has evolved into a mitochondrial proapoptotic molecule with ability to induce apoptosis via multiple pathways [10]. It is synthesized as a 458 amino acid precursor protein which includes a short N-terminal region, serine protease and one C-terminal PDZ domain. The N-terminus carries a transmembrane anchor and a mitochondrial localization signal, which is processed following mitochondrial transport to remove first 133 amino acids. This cleavage exposes an IAP (Inhibitor of Apoptosis Protein) recognizing tetrapeptide motif (AVPS) unique to this member of HtrA family. During apoptosis, mature HtrA2 enters the cytoplasm where it interacts with BIR (Baculovirus IAP Repeat) domains of IAPs such as XIAP (X-linked Inhibitor of Apoptosis

Protein) and relieves their inhibition on active caspases [12, 95] thereby promoting caspase-dependent apoptosis [88]. Although HtrA2 was primarily identified as an IAP-binding proapoptotic protein, the structural organization and domain plasticity has enabled it to mediate apoptosis through other caspase dependent and independent pathways via its serine protease activity. Its caspase-independent proapoptotic property is manifested by its ability to proteolytically cleave important cellular proteins such as cytoskeleton-associated actin, α -, β -tubulin, vimentin, proteins related to translational machinery [10] and few antiapoptotic proteins [13, 227]. This distinct property has made it an interesting target especially where caspase activation is prevented either due to inhibition of caspase activity or genetic inactivation of a key caspase.

The available crystallographic data on mature unbound form of HtrA2 provides a broad overview of overall structural organization of the inactive protease [14]. The mature trimeric protease has a pyramidal architecture where the core serine protease domains are gated by C-terminal PDZ domains at the base and short N-terminal regions on the top thus creating an extremely buried active-site milieu. The structure-based model suggests PDZ is stereochemically limiting the accessibility of substrate in an otherwise proteolytically competent active-site pocket, thus hinting higher activity for the Δ PDZ variant [14]. However, recent reports from our group demonstrated significant decrease in protease activity of this variant [118, 119], thus highlighting the importance of intricate PDZ-protease domain crosstalk and conformational plasticity in defining HtrA2 activity and specificity. Interestingly, although PDZ acts as a regulatory domain in all the members of HtrA family, uniqueness of HtrA2 is manifested by its ability to bind subset of proteins, such as IAPs, through its N-terminus and subsequently cleave them [3]. This phenomenon unequivocally demonstrates existence of multiple or complex mechanisms of HtrA2 activation involving PDZ as well as other regions of the protein.

Here, with a goal to delineate the structural correlates of HtrA2 activation as well as to develop a universal model for its mechanism of action, we carried out comprehensive binding and enzyme kinetic analyses of HtrA2-XIAP interaction using full length protein, its minimal binding region and the corresponding peptides. Our studies clearly demonstrate the role of XIAP as an allosteric activator of HtrA2 thus establishing the importance of N-terminus in regulation of protease activity. The observations suggest how structural plasticity mediated by its substrate binding process and inter-subunit networking drive HtrA2 function.

5.2 Experimental Procedures

5.2.1 Peptides – All the peptides were commercially synthesized (USV Ltd. Mumbai, India) with >95% purity. The primary sequences of the peptides were HtrA2-PDZ_{opt} (NH₂-GQYYFV-COOH) [121], BIR2_{pep} residues (NH₂-²⁰⁴GGKLKNWEPCDRAWSEHRRHF²²⁴-COOH), BIR3_{pep} residues (NH₂-³⁰⁵GGLTDWKPSQEDPWEQHAQWY³²⁴-COOH) and control nonspecific peptide (NH₂-KNNPNNAHQN-COOH).

5.2.2 Recombinant protein production and purification – Mature (Δ 133) HtrA2 full length with C-terminal His₆-tag in pET-20b (Addgene, Cambridge, MA, USA) and PDZ lacking variant HtrA2 Δ PDZ were expressed and purified as described previously [118]. The canonical PDZ-peptide groove mutant (G230A) was generated using site-directed mutagenesis (Stratagene, Austin, TX, USA). Full length XIAP clone in pGEX-4T with GST-tag was obtained from Addgene. XIAP-BIR2, BIR3 domains and mutants were sub-cloned into modified pMALc5-TEV vector (New England Biolabs, Ipswich, MA, USA) that provides an N-terminal maltose binding protein (MBP) tag. The BIR domain and IAP binding motif (IBM) groove mutant tested were XIAP-BIR2: 124–240, E219R/H223V; XIAP-BIR3: 241–356, Q319R/W323V. The IBM groove mutations were also introduced in GST-XIAP full

length as XIAP-E219R/H223V (BIR2 domain mutant), XIAP-Q319R/W323V (BIR3 domain mutant), XIAP-E219R/ H223V, Q319R/W323V (BIR2 and BIR3 domain mutant). Recombinant proteins were expressed in *E. coli* strain Rosetta (DE3) (Novagen, Billerica, MA, USA). Cells were grown at 37 °C till the OD₆₀₀ reaches 0.6 and was then induced with 0.4 mM IPTG. Cells were further cultured at 18 °C for 20 h post induction. Proteins with GST-tag were purified by affinity chromatography using GST-sepharose resin (Novagen, MA, USA) while MBP-tag proteins were purified using amylose resin (New England Biolabs, Ipswich, MA, USA) in buffer 20 mM Na₂HPO₄/NaH₂PO₄, pH 7.8 containing 100 mM NaCl, 25 mM β-mercaptoethanol (β-ME), 50 μM Zn(Ac)₂. The MBP-tag was cleaved using TEV protease [228] and was further subjected to size exclusion chromatography. All the fractions with >95% purity as estimated by SDS-PAGE were stored in aliquots at -80 °C.

5.2.3 Enzyme assay – The protease activity of wild-type and its variants were determined using substrate FITC (fluorescein isothiocyanate) labelled β-casein (Sigma Chemicals, St. Louis, MO, USA) as described previously [118]. The fluorescent substrate cleavage was measured by incubating respective concentration of enzymes in presence or absence of XIAP (proteins or peptides) with increasing concentrations (0 – 25 μM) of β-casein at 37 °C in cleavage buffer (20 mM Na₂HPO₄/NaH₂PO₄, pH 7.8 containing 100 mM NaCl, 0.1 mM DTT). FITC-fluorescence was monitored in a multi-well plate reader (Berthold Technologies, TN, USA) using excitation wavelength of 485 nm and emission at 545 nm. Reaction rates (v_0) were calculated using linear regression analysis. XIAP and BIR domain protein activation curves were fitted to the following equation: $\text{velocity} = \text{basal} + \text{max} / (1 + [K_{\text{act}} / (\text{peptide})]^n)$, where ‘ K_{act} ’ is half maximal activation constant and ‘ n ’ is the Hill constant [229]. The steady-state kinetic parameters were calculated from the reaction rates by fitting the data to Hill form of Michaelis-Menten equation, $\text{velocity} = V_{\text{max}} / (1 + (K_{0.5} / [\text{substrate}])^n)$,

where ' V_{\max} ' is the maximum velocity and $K_{0.5}$ is substrate concentration at half maximal velocity using KaleidaGraph (Synergy software). All the experiments were done independently in triplicate and the mean \pm S.E.M. values are shown in the plots.

5.2.4 Surface Plasmon Resonance (SPR) – Real time kinetic analyses of HtrA2 and XIAP-BIR interaction were performed using SPR (Biacore 3000, GE Healthcare Bio-Sciences, Uppsala, Sweden). Carboxymethylated sensor chips (type CM5) were first activated with a 1:1 mixture of 0.2 M N-ethyl-N'-(3-dimethylaminopropyl) carbodiimide and 0.05 M N-hydroxysuccinimide in water. The ligand, HtrA2 (50 μ g/ml in 10 mM sodium acetate, pH 4.5) was immobilized on the activated CM5 sensor chip by amine coupling (Biacore) and the unreacted sites were blocked with 1.0 M ethanolamine-HCl, pH 8.5. Control flow cells were activated and blocked in the absence of HtrA2 protein. The analytes, BIR proteins (BIR2 and 3) and corresponding to IBM groove peptides of these two domains were then injected as a function of increasing concentration in running buffer (10 mM HEPES pH7.5, 200 mM NaCl) under continuous flow of 20 μ l/min at 25°C. The chip was regenerated after each cycle by washing it with 2 M NaCl followed by running buffer. The sensorgrams for specific interactions were obtained by subtracting the reference unimmobilized flow cell response from that of the sample and were evaluated using the BIA evaluation software package (4.1 versions, GE Healthcare, Sweden). Data were analyzed by fitting to a 1:1 Langmuir binding model of both the association and dissociation phases. The apparent equilibrium dissociation constants (K_D) were determined from the ratio of the dissociation and association rate constants (k_d/k_a).

5.2.5 Isothermal Titration Calorimetry (ITC) – ITC measurements were performed at 25 °C using MicroCal ITC200 (GE Healthcare Bio-Sciences, Sweden). All the solutions were

degassed before titration. The calorimetry cell contained 200 μ l of 33 μ M wild-type or G230A mutant HtrA2 in 20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ buffer, pH 7.8 containing 100 mM NaCl. The titration was initiated with one 0.4 μ l injection followed by 20 injections of 1.5 μ l each of PDZ_{opt} peptide reconstituted in the same buffer. To correct for heat of dilution, peptide injections into a cell containing sample buffer was carried out under identical conditions and subtracted from the raw data prior to analysis. The data were analyzed using MicroCal ORIGIN software with the integrated heat peaks fitted to a one site-binding model.

5.2.6 Pull down experiment – For binding assay, MBP-tag BIR domain and GST-tag XIAP wild-type and mutant proteins were incubated on amylose and GST-sepharose resins respectively for 1 h at 4 °C. Beads were washed three times with binding buffer (20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$, pH 7.8 containing 200 mM NaCl, 1 mM DTT, 0.1% TritonX-100). The washed beads were further incubated with equivalent amount of wild-type HtrA2 for 3 h at 4 °C. The unbound HtrA2 was removed by washing the beads thrice with the binding buffer. The bound proteins were separated on 12% SDS-PAGE and were probed with anti-His antibody (Abcam, Cambridge, MA, USA) against HtrA2.

5.2.7 Protein Modeling and Docking – Crystal structure of inactive unbound HtrA2 (PDB entry-1lcy) was obtained from Protein Data Bank. The protein has missing N-terminal residues (AVPSP), flexible loop (¹⁴⁹ARDLGLPQT¹⁵⁷) and linker region (²¹¹RGEKKNSSSGISGSQ²²⁵). The loop and linker region were modeled and refined using Prime 3.0 [120] (Schrödinger, LLC, New York, 2012). Missing N-terminal residues (AVPSP) were modeled using 'CrossLink proteins' module of BioLuminate 1.2 (Schrödinger, LLC, New York, 2013). The Crosslink Proteins module allows cross-linking of prepositioned proteins by connecting chain termini with peptide linkers. HtrA2 trimer model without PDZ

domain (HtrA2 Δ^{PDZ}) residues 1-211 was generated using Prime 3.0 'Homo-multimer program' which allows generation of homo-multimer from a monomeric protein. Docking of HtrA2 trimer with BIR3 IBM groove peptide (GGLTDWKPSQDPWEQHAQWY) was performed using 'protein-protein docking' program [230] of BioLuminate 1.2 (Schrödinger, LLC, New York, 2013). Protein-protein docking calculations were done by specifying 70,000 ligand rotations to probe, which is approximately to sample every 5° in the space of Euler angles. Protein-protein docking of HtrA2 trimer with BIR3 produced N-terminal docked poses. Three poses showed the peptide docked at N terminal tetrapeptide AVPS region and the docked poses have no steric clashes. BIR3 peptide residues W323, W310, A321, and L307 showed interactions with each chain in the trimer.

5.2.7 Molecular Dynamics Simulation and Analysis – MD Simulation was performed and analyzed using Desmond 3.4 (Schrödinger, LLC, New York, 2013) software package as described previously [118, 231]. Briefly, protein structures were solvated with SPC orthorhombic water box with a 10 Å buffer space. The system was neutralized by replacing water molecules with sodium and chloride counter ions. Nose–Hoover thermostats and Martyna–Tobias–Klein method were used to maintain constant simulation temperature and to control pressure respectively. The entire system was equilibrated using the default protocol provided in Desmond. Two rounds of steepest descent minimization were performed followed by series of four molecular dynamics simulations. The default equilibration was further followed by 5000 ps NPT (constant number of particles, pressure and temperature) simulation to equilibrate the system. A total 20 ns NPT production simulation was then run and coordinates were saved every 2 ps of time interval. Simulation data was analyzed with the help of analytical tools in the Desmond package. For MD Simulation quality analysis, potential energy of the protein and total energy of entire system was calculated. The lowest

potential energy conformations were used for structural analysis of peptide bound and unbound structures. The trajectories of peptide bound complex and unbound form were then compared based on their overall calculated domain wise RMSD and RMSF values. These are graphically represented using GraphPad Prism 5.0 (GraphPad Software, San Diego, CA, USA).

5.3 RESULTS

5.3.1 *Interaction analysis and elucidating the role of IAPs in activation of HtrA2*

5.3.1.1 Determination of the proteolytic activity of HtrA2 in presence of IAPs - Apart from deactivating IAPs [95, 122], the complex structural organization and domain plasticity of HtrA2 has enabled it to promote apoptosis through various caspase-mediated and independent pathways via its serine protease activity [2, 10]. To assess any role of IAPs on HtrA2 proteolytic activity, we determined its initial velocity (v_0) as a function of increasing XIAP concentration, using fixed amount of enzyme and substrate β -casein, a well established model substrate of HtrA proteases [120, 232]. We observed that the rate of proteolysis increased in the concentration dependent manner, and the half maximal activation constant (K_{act}) was determined to be 10 ± 0.7 nM of XIAP (**Fig. 5.1A**). The full length XIAP is a homodimer consisting of three tandem BIR domains, followed by single UBA (Ubiquitin Associated) and RING domains [97]. BIR2 and/or BIR3 domains have been reported to be essential for HtrA2-XIAP interaction [95]. Therefore, to further explore XIAP mediated HtrA2 activation, we tested the protease activity in presence of either BIR2 or BIR3 domains. Interestingly, we observed that in presence of BIR2, the cleavage rate was similar to full length XIAP (**Fig. 5.1B**); however, the former achieves so at the expense of four-fold greater protein concentration (58 ± 10 nM). In the case of BIR3, the rate was slightly less compared to the full length and also exhibited high K_{act} concentration (200 ± 25 nM) (**Fig. 5.1C**). These

results suggest that most likely BIR2 domain acts as the key player in XIAP mediated HtrA2 activation. However, increase in half maximal activation constant of BIR domains compared to full length suggests homodimeric XIAP with tandem BIRs may be important in enhancing its avidity towards HtrA2 through two-site interaction driven mechanism.

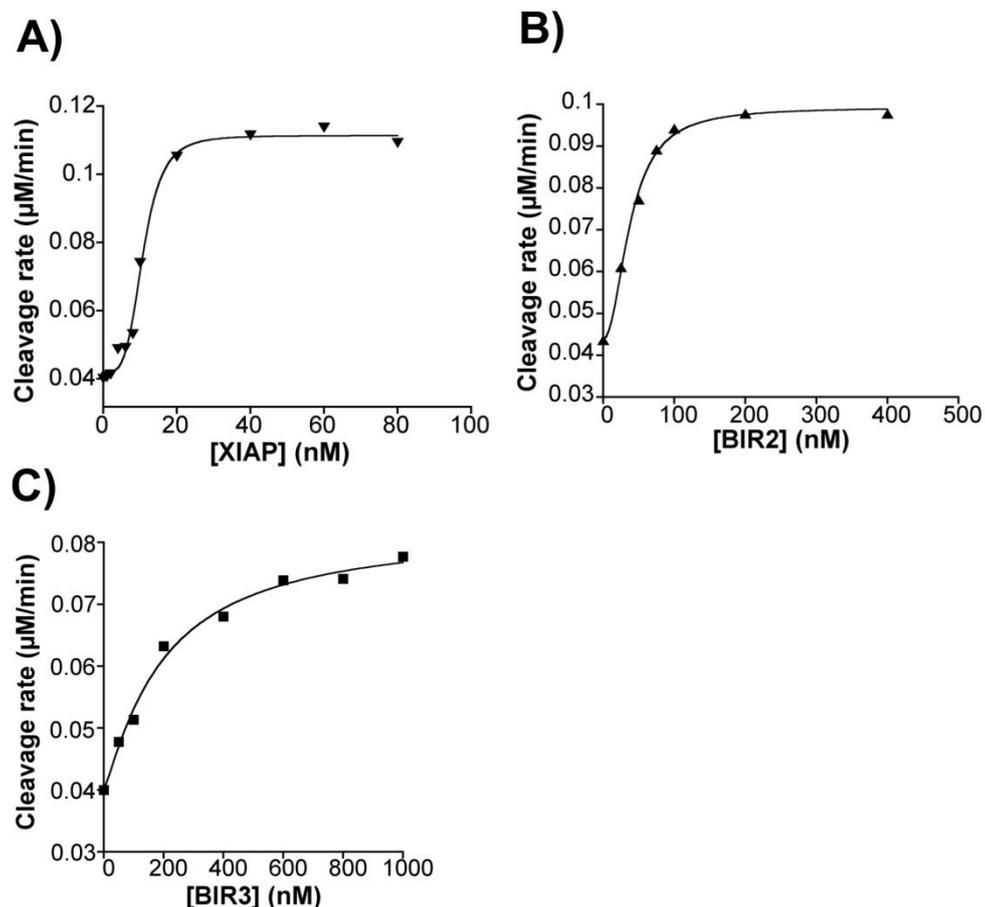


Figure 5.1 XIAP mediated activation of HtrA2. A) XIAP activation of β -casein cleavage by wild-type HtrA2. B) Activation of HtrA2 by XIAP BIR2, C) XIAP BIR3 domain. In all the panels, concentration of substrate β -casein was 5 μ M and HtrA2 was 200 nM. The solid lines are fits to the equation: $velocity = basal + \frac{max}{(1 + [K_{act}/(peptide)]^n)}$, where n is the Hill constant and K_{act} is half maximal activation constant. Plots are mean of three independent data sets.

5.3.1.2 Disruption of the XIAP–HtrA2 interaction prevents protease activation

HtrA2 interacts with XIAP via N-terminal Reaper like motif (AVPS) similar to Smac/DIABLO [88]. To validate the importance of XIAP-HtrA2 interaction in activation of the protease, we created an AVPS tetrapeptide motif mutant ($\Delta 1$, V2, P3A, S4, termed here as A1 Δ P3A). As anticipated, this variant did not interact with XIAP as observed by GST-pull

down assay (**Fig. 5.2A**). In addition, the proteolytic activity of A1 Δ P3A was found comparable to wild-type suggesting that the mutation did not affect the overall protease conformation. Strikingly, in presence of IAP proteins, there was no increase in the proteolytic activity of HtrA2 (**Fig. 5.2D**). This result suggests that interaction of XIAP with the amino-terminal tetrapeptide motif (AVPS) of HtrA2 results in the activation of the proteolytic activity of the enzyme.

To confirm the role of XIAP-HtrA2 interaction in activation of the protease, we decided to target the IAP-binding motif (IBM) groove of XIAP that is responsible for its interaction with HtrA2 [88]. Since Glu, His and Gln, Trp pairs of IBM groove residues confer selectivity to BIR2 and BIR3 respectively for interaction with caspase-3 and -7 tetrapeptide motif [96], we assumed that interaction XIAP and HtrA2 tetrapeptide motif possibly involves the same pair of residues. We therefore targeted these pairs aiming to explore the importance of these residues in interaction and its effect on HtrA2 activation. For this, we generated individual BIR domain mutants, BIR2 (E219R, H223V) and BIR3 (Q319R, W323V) as well as combinatorial mutants for full length XIAP (DM, B2M, B3M), and pull-down assay was performed. Pull-down studies with the IBM groove mutants of BIR domains showed complete abrogation of interaction whereas, it was partially retained in case of full length XIAP with only single IBM groove mutation (**Fig. 5.2C**). To explore the effect of loss of interaction on protease activity, we determined the cleavage rates in presence of these different mutant XIAP proteins. It was observed that XIAP DM completely failed to enhance the proteolytic activity probably due to its inability to interact with HtrA2 (**Fig. 5.2D**). Altogether, these results demonstrate the importance of IBM groove mediated HtrA2 interaction in activation of the protease.

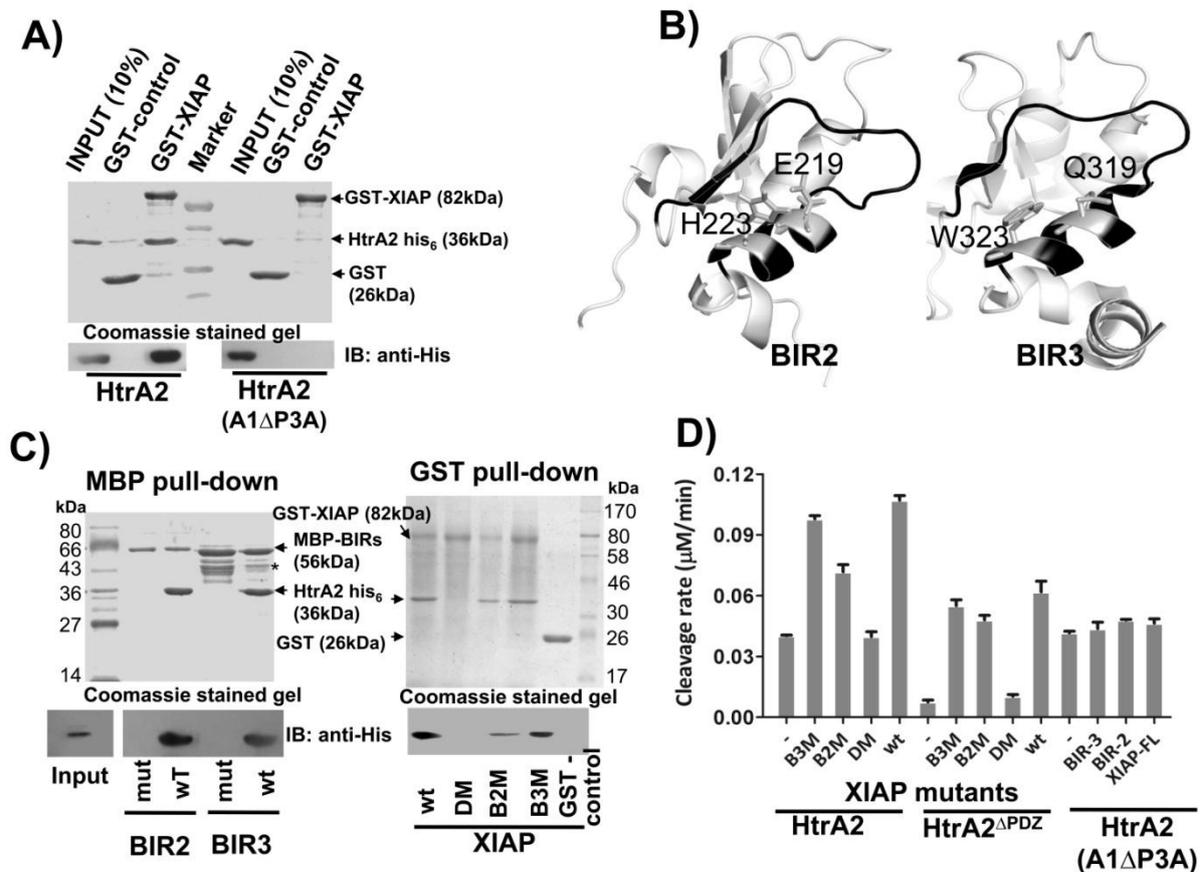


Figure 5.2 Mutational analysis of XIAP-HtrA2 interaction. A) GST-pull down assay of wild-type HtrA2 and AVPS mutant (A1ΔP3A) with XIAP. The A1ΔP3A mutant (prey) was defective in binding to GST-XIAP (bait) as compared to wild-type. GST-control did not show any binding to HtrA2. B) Structure of BIR2 (PDB id. 1c9q) compared with BIR3 (PDB id. 1f9x) highlighting IAP binding motif (IBM) groove in black. Residues for IBM groove mutations E219, H223 of BIR2 and Q319, W323 of BIR3 are shown as sticks. C) Pull down assay of IBM groove mutants using full length XIAP, BIR2 and BIR3 domains. In left panel, MBP-pull down was carried out with MBP-BIR2 and 3 wild-type and IBM groove mutants BIR2 (E219R, H223V) and BIR3 (Q319R, W323V). A representative coomassie stained SDS PAGE and corresponding immunoblotting with anti-His for detection of his-tagged HtrA2 shows that HtrA2 failed to associate with mutants (* denotes MBP contamination). In right panel, GST-pull down was performed with XIAP and its mutants. For XIAP with both BIR domains mutated (DM) showed complete loss of interaction, while, the single domain mutants XIAP-BIR2 (E219R, H223V) (B2M) and XIAP-BIR3 (Q319R, W323V) (B3M) retained binding with HtrA2. D) Basal cleavage rate of β-casein by HtrA2 and HtrA2^{ΔPDZ} in presence of XIAP and its mutants. The DM failed to increase the enzymatic activity of HtrA2 while B3M and B2M enhanced the activity similar to wild-type BIR2 and BIR3 respectively. The A1ΔP3A mutant incapable of binding to XIAP did not show any increment in proteolytic activity in presence of IAP.

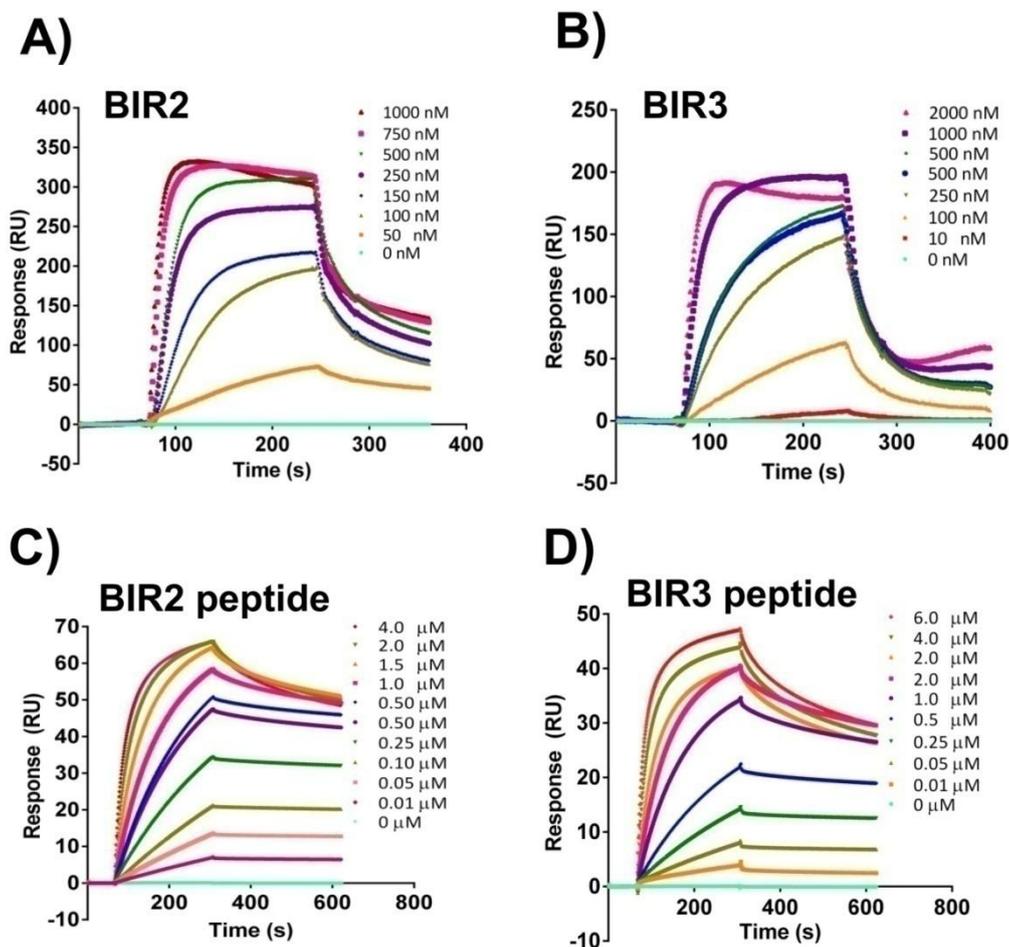


Figure 5.3 SPR binding analyses of HtrA2 with BIR domain proteins and IBM groove peptides. HtrA2 was immobilized on CM5 sensor chip as described under *experimental procedures*. Indicated concentration of BIR proteins and peptides (analyte) were injected over immobilized HtrA2 at a flow rate of 20 $\mu\text{l}/\text{min}$ at 25 $^{\circ}\text{C}$. The sensorgram shown is for binding kinetics of BIR2 (A), BIR3 (B), BIR2_{pep} (C), and BIR3_{pep} (D). The specific binding response was obtained by subtracting the background signal over a reference surface without immobilized protein. Data was evaluated using BIAevaluation software program and the constants were calculated by fitting the data using 1:1 Langmuir binding model.

5.3.1.3 Interaction analysis

To determine the kinetic and affinity constants we performed quantitative interaction analyses of HtrA2 with XIAP-BIR domains and the IBM groove peptides using Surface Plasmon Resonance (SPR). It was observed that BIR proteins and peptides readily bound to immobilized HtrA2 (**Fig. 5.3**). Kinetic analysis of the binding data for BIR2 and 3

demonstrated apparent equilibrium dissociation constants (K_D) of 54 ± 8 nM and 254 ± 18 nM, respectively (Table 5.1). A decrease in K_D observed for BIR2 was mainly due to higher association combined with moderately lower dissociation rates as compared to BIR3 (Fig. 5.3A and B). We further correlated the binding properties for BIR proteins with those of IBM groove peptides (Fig. 5.3C and D), which showed comparable apparent K_D values (Table 5.1). Thus, HtrA2 interacts with both BIR2 and BIR3 domains of XIAP but shows higher affinity for the former. In addition, these analyses establish that IBM groove is critical and sufficient in mediating XIAP-HtrA2 interaction.

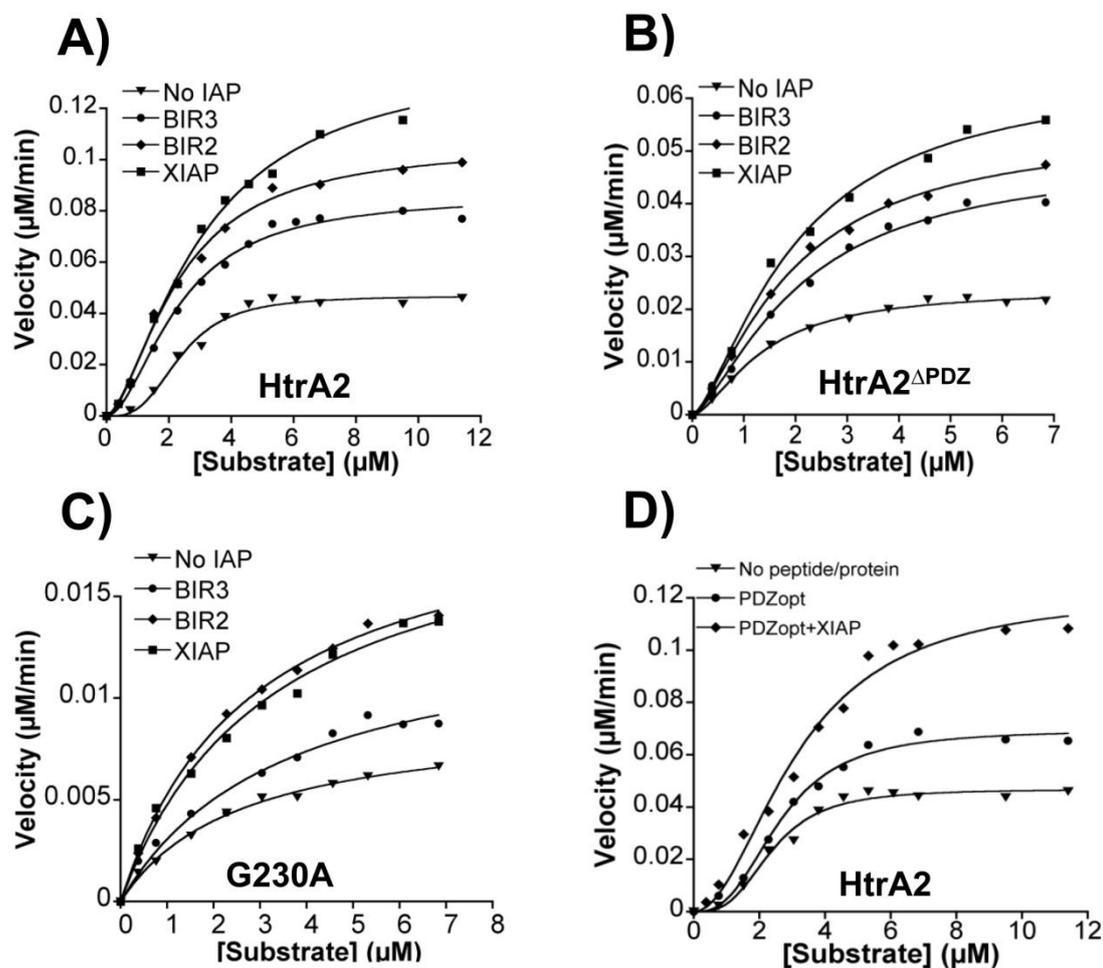


Figure 5.4 Activity of wild-type HtrA2, its variant and mutant in presence or absence of activator. The steady-state kinetics of β -casein cleavage by wild-type HtrA2 (200 nM) (A), PDZ lacking variant HtrA2 Δ PDZ (1000 nM) (B) and PDZ-peptide groove mutant G230A (200 nM) (C) in presence of saturating concentration of XIAP (20 nM), BIR2 (100 nM) and BIR3 (1000 nM) proteins. D) Plot representing activity of wild-type HtrA2 in presence of either 5 μM PDZ interacting peptide (PDZ_{opt}) and combination of PDZ_{opt} and full length XIAP (20 nM). The solid lines are nonlinear

least squares fit of the data to the Hill form of the Michaelis-Menten equation: $\text{velocity} = V_{\text{max}}/(1+(K_m/[\text{substrate}])^n)$. For clarity, error bars are not shown for all the data sets.

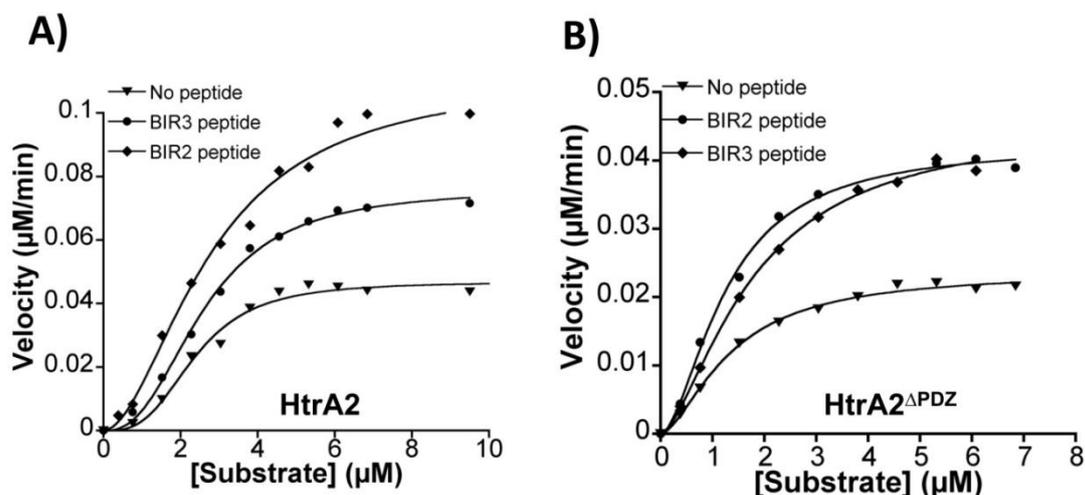


Figure 5.5 Determination of steady state kinetics for β -casein cleavage by HtrA2 and HtrA2^{APDZ} in presence of IBM peptides. Graph of β -casein cleavage in presence of IBM peptides, 100 nM of BIR-2_{pep} (NH₂-GGKLNWEPDRAWSEHRRHF-COOH) and 1000 nM BIR-3_{pep} (NH₂-GGLTDWKPSDPEQWQHAKWY-COOH) by A) HtrA2 (200 nM) B) HtrA2^{APDZ} (1000 nM).

Table 5.1 Surface plasmon resonance analysis of the interaction between HtrA2 and XIAP-BIR proteins and peptides. Kinetics of the interaction between HtrA2 and XIAP-BIR proteins or peptides was evaluated using Biacore 3000 instrument. The ligand HtrA2 was immobilized on a surface that was challenged with different concentrations of BIR proteins/peptides and the rate constants were determined by applying 1:1 Langmuir model to the data. Equilibrium dissociation constant (K_D) was calculated from the ratio of dissociation and association rate constants (k_d/k_a).

Ligand	Analyte	Association rate constant k_a (1/Ms)	Dissociation rate constant k_d (1/s)	Equilibrium dissociation constant K_D (M)
HtrA2	BIR2	1.6×10^5	8.5×10^{-3}	5.5×10^{-8}
HtrA2	BIR3	5.7×10^4	1.5×10^{-2}	2.5×10^{-7}
HtrA2	BIR2 _{pep}	1.3×10^4	2.0×10^{-4}	1.5×10^{-8}
HtrA2	BIR3 _{pep}	7.9×10^3	1.5×10^{-3}	2.0×10^{-7}

5.3.2 Determining specificity and catalytic efficiency of HtrA2

5.3.2.1 Steady-state kinetic parameters for HtrA2

To determine the steady-state kinetic parameters, we measured initial rates of substrate cleavage for different β -casein concentrations with or without saturating XIAP proteins or peptides. We observed that HtrA2 cleavage of β -casein follows a sigmoidal curve with Hill constant of 2.8 ± 0.2 (**Fig. 5.4A**). The substrate β -casein has a carboxyl-terminus PDZ-binding consensus (GPFPIIV) and has specific cleavage sites for serine protease thus acts as a C-terminal self activating substrate for HtrA2 [118]. Therefore, the positive cooperativity upon substrate cleavage suggests that binding of one substrate (β -casein) molecule to one subunit of HtrA2 trimer favors consequent binding of others. Interestingly, upon pre-incubation of HtrA2 with XIAP, the catalytic efficiency ($k_{cat}/K_{0.5}$) increased about three-fold and the Hill constant was reduced to 1.5 ± 0.1 , thus indicating greater stabilization of the relaxed (active) state of protease upon IAP binding [229]. However, the substrate concentration at half maximal velocity ($K_{0.5}$) remained unchanged, even in presence of effector (XIAP), implying existence of a similar binding pocket. Similar results were obtained in presence of saturating concentration of either BIR domain proteins or peptides, as shown in **Table 5.2**.

Here, in presence of XIAP proteins or peptides, there is an increase in V_{max} without any significant alteration in K_m of HtrA2 which indicates that it likely follows ‘*V system*’ of allosteric modulation [233]. In this model, substrate has similar affinity for both relaxed (active) and tensed (inactive) states of enzyme; hence, there is no influence of effector on substrate binding (and vice versa). It majorly depends on whether effector has maximum affinity for active state (acting as an activator) or for the inactive state (here, an inhibitor) [234]. This differential affinity of the effector towards enzyme along with relaxed state

stabilization shifts the equilibrium towards active state [229, 235]. Thus, XIAP possibly acts as an activator to allosterically modulate HtrA2 to form a catalytically active enzyme, which then cleaves β -casein.

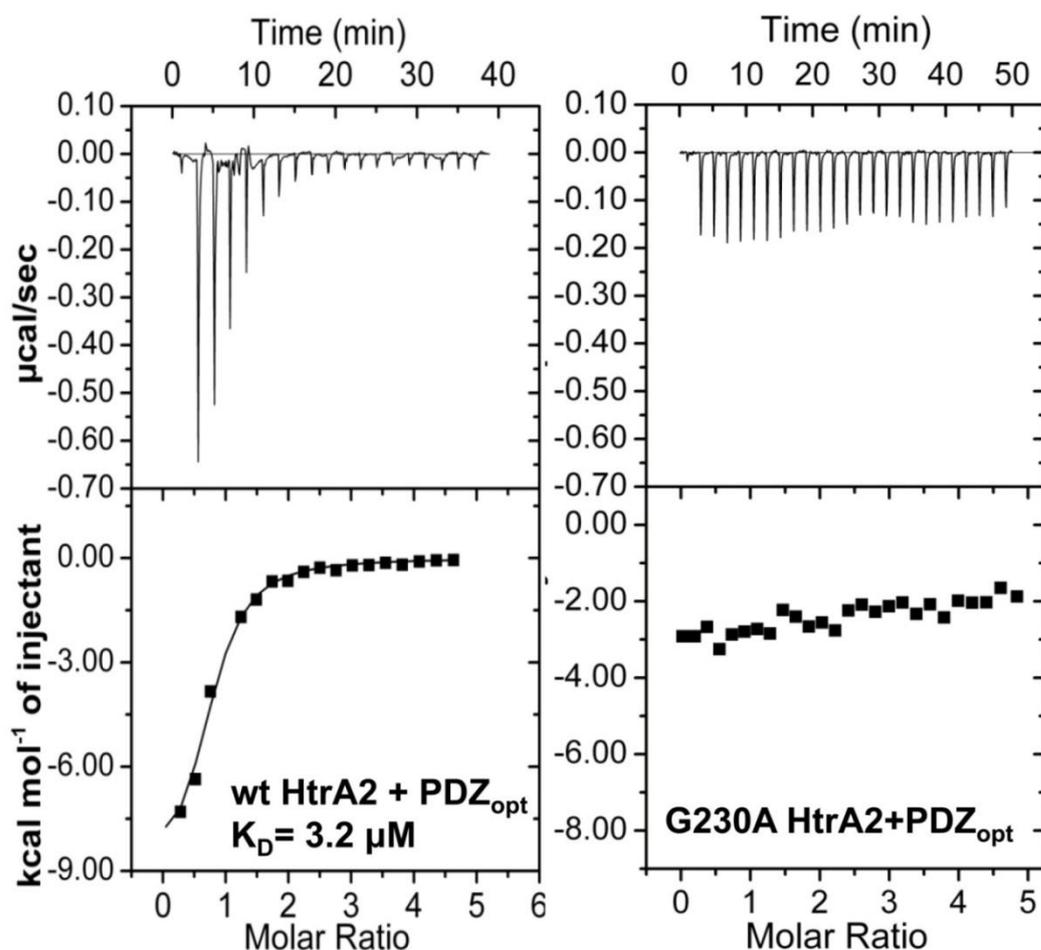


Figure 5.6 Interaction analyses of PDZ_{opt} with HtrA2 wild-type and G230A mutant measured using isothermal titration calorimetry. Binding thermograms for PDZ_{opt} peptide (H₂N-GQYYFV-COOH) (0.5 mM) with HtrA2 wild-type (0.03 mM) (left panel) and PDZ-peptide groove mutant, G230A (0.033 mM) (right panel). Data was recorded at 298 K. In each titration, the upper panel shows raw data, whereas the lower panel corresponds to corrected integrated binding isotherm together with fitted binding curve. The dissociation constant was calculated to be 3.2 µM for wild-type whereas no significant heat change was observed for PDZ-peptide groove mutant.

Table 5.2 Steady-state enzymatic parameters for HtrA2 and its variants. The initial rates for substrate cleavage in presence or absence of XIAP proteins or peptides were measured and fitted to Hill form of Michaelis-Menten equation to determine the steady-state kinetic parameters (**Fig. 5.4 and 5.5**). The values are mean \pm S.E.M. and generated from data points for at least three independent experiments

Enzyme	Activator (protein/peptide)	$K_{0.5}$ (μM)	Hill constant	Maximum velocity V_{max} (Ms^{-1}) $\times 10^{-9}$	Catalytic efficiency $k_{\text{cat}}/K_{0.5}$ ($\text{M}^{-1}\text{s}^{-1}$)
wild-type	none	2.3 ± 0.1	2.8 ± 0.2	0.84 ± 0.1	1.6×10^3
wild-type	BIR3	2.3 ± 0.2	1.8 ± 0.1	1.4 ± 0.1	3.1×10^3
wild-type	BIR2	2.3 ± 0.2	1.6 ± 0.1	1.8 ± 0.2	3.8×10^3
wild-type	XIAP FL	2.7 ± 0.1	1.5 ± 0.1	2.3 ± 0.1	4.7×10^3
G230A	none	3.2 ± 0.2	n.a.	0.17 ± 0.06	1.6×10^2
G230A	BIR3	3.5 ± 0.2	n.a.	0.28 ± 0.01	2.6×10^2
G230A	BIR2	3.2 ± 0.1	n.a.	0.38 ± 0.09	4.4×10^2
G230A	XIAP FL	3.0 ± 0.1	n.a.	0.51 ± 0.13	5.1×10^2
HtrA2 ^{ΔPDZ}	none	1.4 ± 0.04	1.7 ± 0.1	0.39 ± 0.12	2.9×10^2
HtrA2 ^{ΔPDZ}	BIR3	1.7 ± 0.1	1.5 ± 0.04	0.82 ± 0.16	5.1×10^2
HtrA2 ^{ΔPDZ}	BIR2	1.7 ± 0.1	1.4 ± 0.1	0.91 ± 0.11	6.2×10^2
HtrA2 ^{ΔPDZ}	XIAP FL	1.8 ± 0.1	1.5 ± 0.1	1.08 ± 0.12	8.0×10^2
Peptide-mediated activation of HtrA2					
wild-type	BIR3 _{pep}	2.5 ± 0.2	2.2 ± 0.2	1.3 ± 0.1	2.4×10^3
wild-type	BIR2 _{pep}	2.8 ± 0.3	1.9 ± 0.1	1.9 ± 0.1	3.3×10^3
wild-type	PDZ _{opt}	2.6 ± 0.3	2.5 ± 0.2	1.2 ± 0.1	2.2×10^3
wild-type	PDZ _{opt} + XIAP	2.9 ± 0.3	1.9 ± 0.2	2.2 ± 0.2	4.0×10^3
HtrA2 ^{ΔPDZ}	BIR3 _{pep}	1.8 ± 0.1	1.5 ± 0.1	0.75 ± 0.06	5.4×10^2
HtrA2 ^{ΔPDZ}	BIR2 _{pep}	1.3 ± 0.2	1.2 ± 0.1	0.76 ± 0.15	6.1×10^2

(n.a. - not applicable)

5.3.2.2 Comparison of the catalytic efficiency upon N- and C-termini induced activation

To identify whether the mechanism of HtrA2 activation via N- and C-termini effector binding is additive, we carried out enzyme cleavage assay in presence of earlier reported optimal HtrA2 PDZ domain binding peptide (NH₂-GQYYFV-COOH), termed as PDZ_{opt} [121]. It

was observed that pre-incubation of HtrA2 with PDZ_{opt} peptide increased the cleavage rate of substrate (**Fig. 5.4D**). In addition, presence of saturating concentration of XIAP exhibited further ~two-fold increase in turnover number and catalytic efficiency (**Table 5.2**), demonstrating synergy between XIAP and PDZ_{opt} induced activation.

To understand whether N-terminal XIAP-mediated activation requires any direct involvement of C-terminal PDZ domain, we first determined the enzymatic parameters for a PDZ lacking variant (HtrA2^{ΔPDZ}). HtrA2^{ΔPDZ} cleaved β-casein with steady-state kinetic parameters of $K_{0.5} = 1.3 \pm 0.05 \mu\text{M}$, $V_{\text{max}} = 0.39 \times 10^{-9} \text{ Ms}^{-1}$ and Hill constant = 1.71 ± 0.1 (**Table 5.2**). The $k_{\text{cat}}/K_{0.5}$ of this variant was 5.5-fold less than intact trimeric HtrA2 and the apparent K_m and co-operativity were also reduced. Since β-casein is a C-terminal binding substrate, deletion of PDZ domain possibly affects initial substrate binding process, allosteric modulation and cleavage, thereby leading to decrease in co-operativity and catalytic efficiency. However, interestingly, addition of XIAP (either proteins or peptides) to HtrA2^{ΔPDZ} enhanced β-casein cleavage rate to a fold increase similar to that observed for full length HtrA2 (~2.8-fold) with concomitant decrease in cooperativity (**Fig. 5.4B and 5.5B**). This suggests that N-terminal ligand is capable of independently inducing conformational changes in protease domain to form an active conformer.

To further clarify N-terminal XIAP-mediated HtrA2 activation and the role of C-terminal PDZ domain, we tested the enzymatic activity of a PDZ-binding groove mutant. PDZ domains have a canonical binding site comprising highly conserved ‘G-Φ-G-Φ motif’, where Φ denotes hydrophobic residues. The first Gly residue is highly variable among PDZ domains while the third Gly is found to be conserved. The second and fourth residues are mostly hydrophobic (Val, Ile, Leu, or Phe) [117]. This recognition sequence is substituted as

YIGV in HtrA2. To generate the PDZ-binding groove mutant, we targeted the Gly residue of the motif, denoted here as G230A, to prevent substrate interaction with PDZ domain. To test whether G230A mutation disrupts interaction of HtrA2 with PDZ binding substrates/peptides, interaction analysis was carried out by isothermal titration calorimetry using PDZ_{opt} peptide. It was found that wild-type HtrA2 interacts with the peptide; however the G230A mutant did not, thus confirming the importance of YIGV in the initial substrate binding (**Fig. 5.6**). We then determined the enzymatic properties of this mutant with or without saturating concentration of IAPs. As anticipated, due to loss in initial substrate binding, a significant decrease in the steady-state kinetic parameters was observed for G230A compared to wild-type (**Table 5.2**). To our surprise the mutant did not exhibit any cooperativity and the plot was found to be hyperbolic (**Fig. 5.4C**). This Michaelis-Menten behavior of the enzyme can be attributed to the lack of substrate-induced stabilization of active-site as well as to the steric block imparted by PDZ domain to the specificity pocket as a result of a non-functional ‘YIGV’ binding groove. This could have lead to loss in cooperativity and decreased protease activity. However, increase in V_{\max} by 2.4-fold upon addition of IAPs (**Table 5.2**) re-iterates the role of N-terminus in regulation of HtrA2 activity in a PDZ independent manner.

Table 5.3 Comparison of distances between atoms of the catalytic triad in the BIR3_{pep}-bound and unbound form of HtrA2 (in Å)

Protein Complex	N ^{ε2} (His) – O ^γ (Ser)		N ^{δ1} (His) – O ^{δ1} (Asp)	
	Bound	Unbound	Bound	Unbound
HtrA2 ^{ΔPDZ} –BIR3 _{pep}	3.2	5.1	2.6	2.9

Å: Angstrom

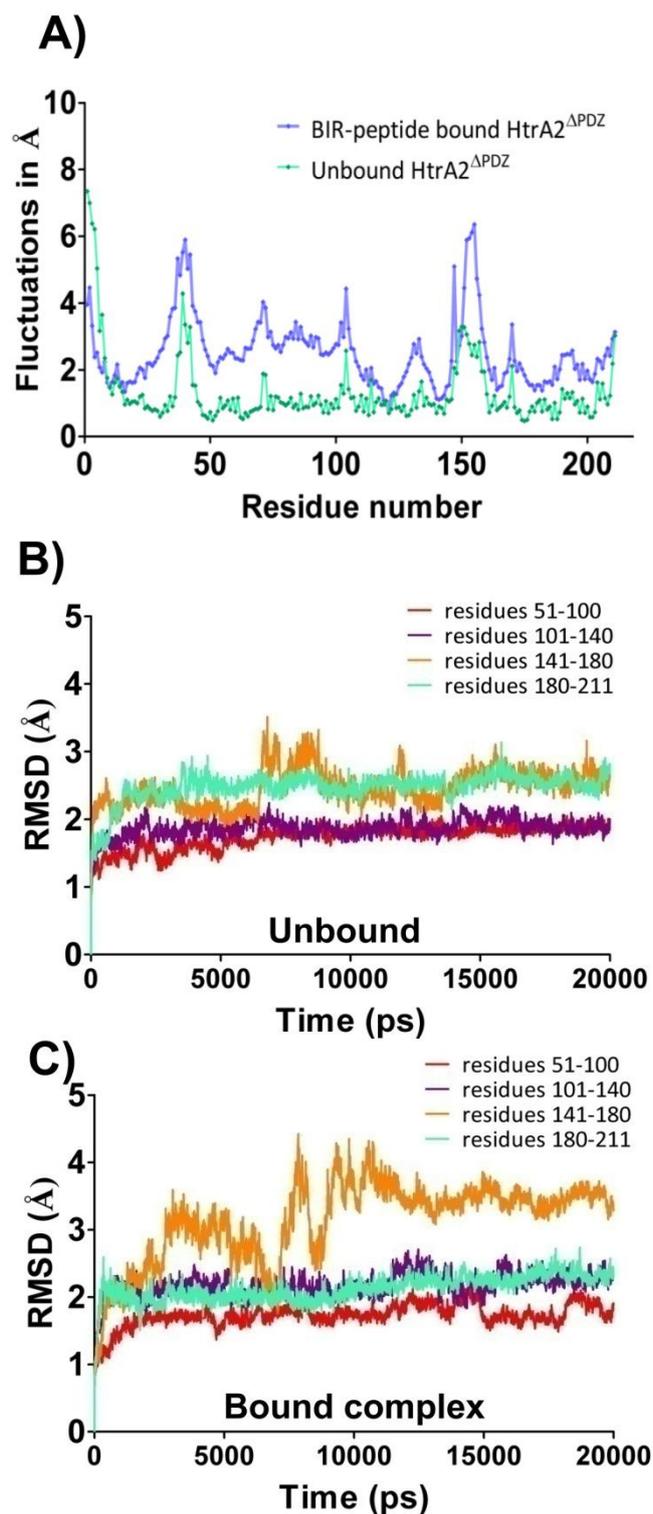


Figure 5.7 Graphical representation of movements in peptide bound and unbound form of trimeric HtrA2^{ΔPDZ}. A) RMSF plot of 20 ns MD simulation trajectory for BIR3 peptide-bound and unbound HtrA2^{ΔPDZ}. B) Domain wise RMSD plots for peptide unbound and C) bound HtrA2^{ΔPDZ}. The residues 140-180 comprising loops LD, L3, and L1 display significant deviations upon peptide binding.

5.3.3 Delineating the global conformational changes following IAP binding

5.3.2.1 Monitoring conformational changes using *in silico* approach

With the aim of understanding the structural rearrangements and molecular motions that govern HtrA2 activity, we adopted an *in silico* approach. Molecular dynamic (MD) simulation is a powerful tool for predicting the conformational plasticity and intricate molecular motions of large enzymes and their complexes [236]. To gain insight into the structural rearrangements that occur in a nanosecond time scale upon activator binding, MD analyses of BIR3 peptide-bound HtrA2 complex and unbound form were performed. Since BIR3-Smac mimetic complex has already been reported in literature [237, 238], which shows interaction of a similar molecule (Smac) with IBM groove, we carried out our MD simulation studies with BIR3 peptide. A trimeric HtrA2^{APDZ} (residues 1-211) was generated to perform docking and MD simulation, resulting in the BIR3 peptide-bound complex (BIR3_{pep}-HtrA2^{APDZ}). Simultaneously, peptide-unbound HtrA2^{APDZ} was subjected to MD simulation for comparison with the bound form. We observed a significant difference in conformation of regulatory loops of the protease upon IAP binding. RMSF (root mean square fluctuation) plot of the trajectories for BIR3_{pep}-bound complex showed higher relative fluctuations in loop LA (residues 37-41), LD (residues 126-140) and L3 (residues 142-162) (**Fig. 5.7A**). Domain wise RMSD (root mean square deviation) plots demonstrated major changes in sensor loop L3 (**Fig. 5.7B and C**). However, loop L2 (190-196) which harbors the specificity pocket remained unchanged in the bound or unliganded states suggesting presence of a well-formed substrate binding pocket. This observation corroborates very well with substrate binding affinity (apparent K_m) which was not appreciably different in presence or absence of the activator. Structural superposition of the peptide-bound and unbound forms shows an average RMSD of 1.21Å. The difference could mainly be attributed to structural deviation near the activation domain loops and some variations at the N-terminus, (residues 1-9) as well (**Fig.**

5.8A and B). Further characterization of the structural changes was performed to understand the mechanistic details of activation.

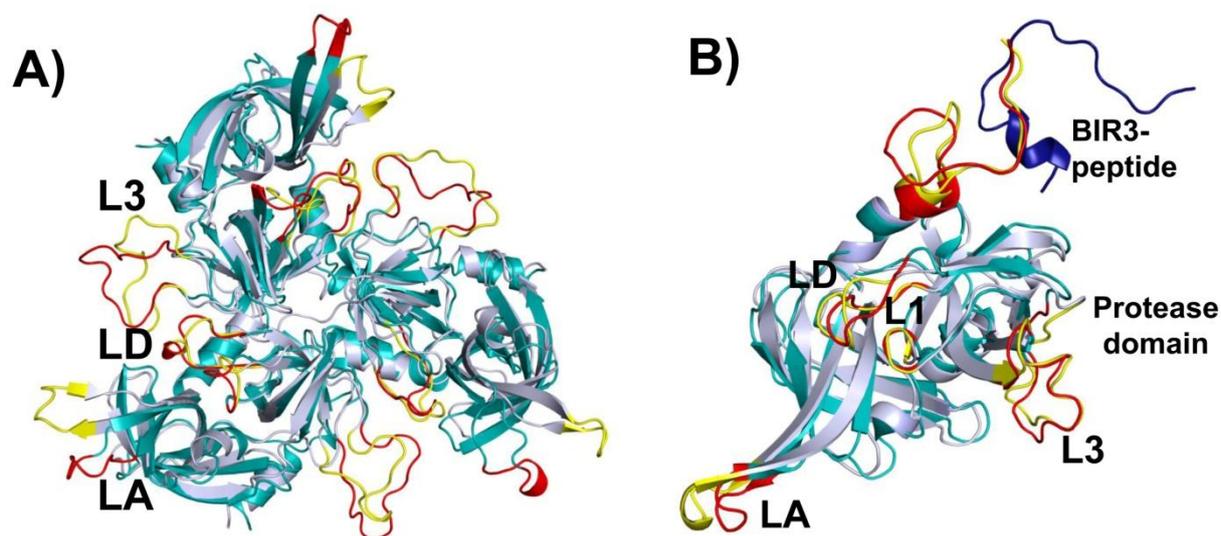


Figure 5.8 Structural analysis of peptide bound and unbound HtrA2. A) Structural alignment of BIR peptide bound-HtrA2^{APDZ} (cyan) and unbound (light blue) in cartoon representation. The movement in loops LA, LD and L3 are highlighted in red (peptide bound) and in yellow for unbound structure. B) Structural overlay highlighting single subunit of BIR3 peptide (dark blue) complex (cyan) and unbound (light blue) with the re-oriented regions shown in red (bound) and yellow (unbound).

5.3.2.2 Structural evaluation of peptide bound and unbound HtrA2

In serine proteases, formation of catalytic triad in an arrangement close enough for electron transfer from Asp to Ser through His is a prerequisite for an active enzyme [239]. For mature HtrA2, the catalytic triad that includes His65, Asp95 and Ser173 is malformed in the available unbound crystal structure [14]. Structural comparison of the unbound protease with BIR3_{pep}-bound complex shows relative movements in the active-site triad residues (**Fig. 5.9A**). It has been well established that the distance between the N^ε atom of His and the O^γ atom of Ser in active proteases is < 3.5 Å whereas in the inactive form it is >3.5 Å [9]. The atomic distances between nitrogen (ε) atom of His 65 and oxygen (γ) of Ser173 decreases

from 5.1 Å to 3.2 Å in bound form and that between nitrogen (δ) of His65 and oxygen (δ) of Asp95 for peptide-bound complex is 0.3 Å shorter when compared to the unbound structure (**Table 5.3**).

Apart from catalytic triad, another key factor for proteolysis by serine proteases is stabilization of negative charge on carbonyl oxygen of tetrahedral alkoxide intermediate (oxyanion hole) [240], which resides in loop L1 (169-173) of HtrA2. Structural alignment of the oxyanion hole residues for HtrA2 (BIR3_{pep}-bound and unbound form) in comparison with free and outer membrane porins (OMP)-bound *E.coli* DegS (closest structural ortholog of HtrA2) is shown in **Fig. 5.9A**. In BIR-bound form, the nitrogen atom of Gly171 (-2 position) acts as a hydrogen donor to form proper oxyanion hole whereas in unbound form, it is oriented in the opposite direction creating a malformed structure. Based upon structural overlay, the re-orientation of loop L1 seems to occur due to its increased proximity towards another mechanistically important loop LD via hydrophobic interaction. It thereby shifts the phenyl ring of F170 so as to accommodate LD loop and form a competent oxyanion hole as shown in **Fig. 5.9B**. Hence, the orientation of loops that accommodate oxyanion hole and catalytic triad residues might render the protease active or inactive in the presence or absence of ligand, respectively.

In HtrA group of serine proteases, intricate inter-molecular interaction among different subunits of oligomer leads to proper active-site formation [226]. From structural alignment of BIR3_{pep}-HtrA2^{ΔPDZ} with the unbound form, we observed significant deviations in loop L3 and LD upon peptide binding. This movement in particular seems to allow loop L3 of one subunit to move towards LD of neighboring unit. Specifically, in the bound complex, P148 and A149 of L3 loop of one sub-unit form non-bonded contacts with P130 and F131 of LD loop from adjacent subunit, which otherwise is separated in the unbound structure (**Fig. 5.9B**). Thus, all these loop movements might coordinate to bring about the observed concerted

conformational rearrangements in and around the active-site leading to protease activation upon IAP binding.

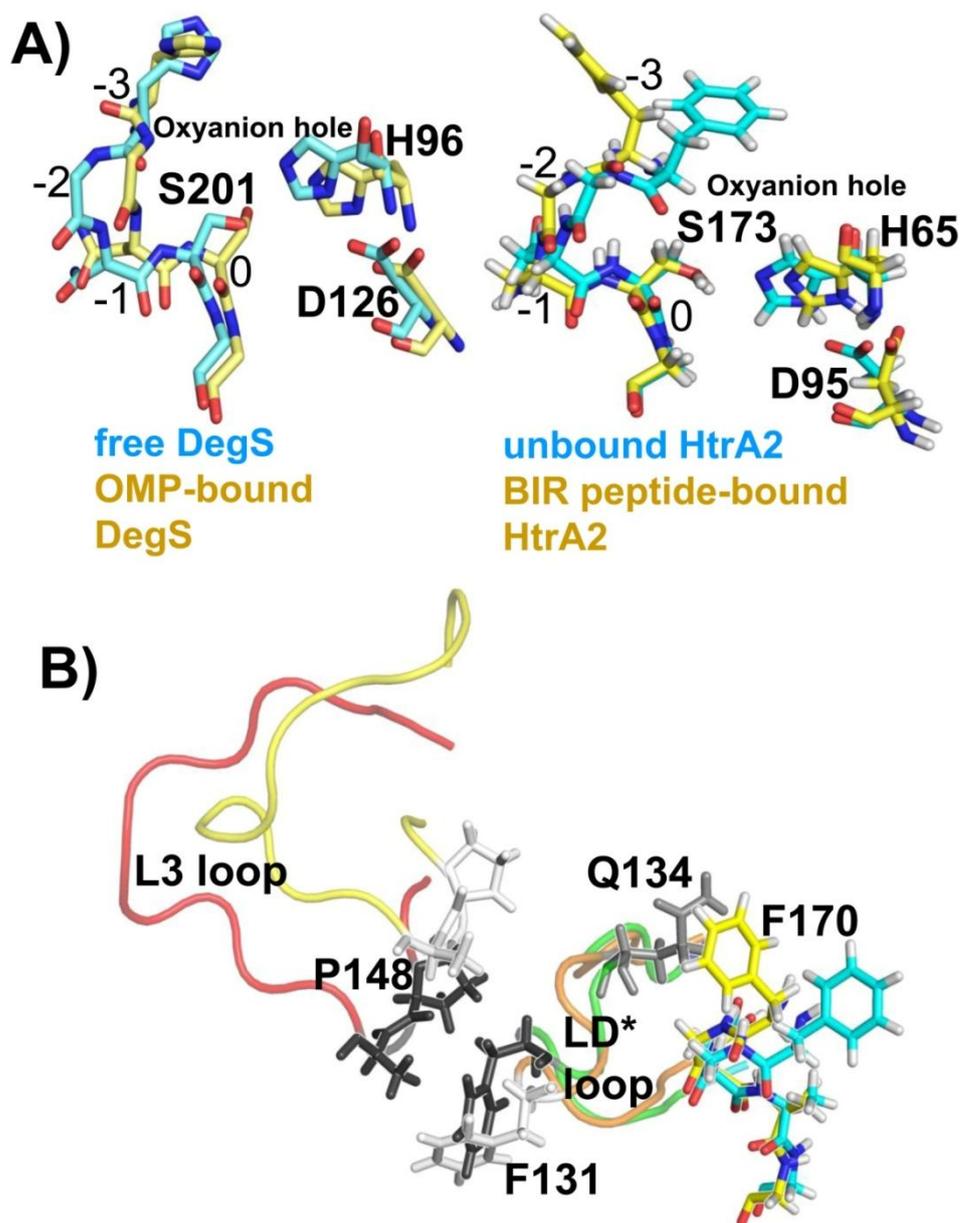


Figure 5.9 Monitoring conformational transition and rearrangements in BIR peptide-bound and unbound trimer HtrA2^{APDZ}. A) In left panel, comparison of oxyanion hole and active-site residues for inactive (OMP-free) and active (OMP-bound) forms of *E.coli* DegS shown as sticks. The residues in oxyanion are indicated as 0 for active-site Ser and -1, -2,-3 for preceding residues. In the inactive form, oxyanion hole is distorted especially in amide nitrogen atom of Gly at -2 position. Also, the active-site residues vary substantially. Similarly, in right panel, unbound HtrA2 shows distorted

oxyanion hole and malformed active-site whereas BIR peptide-bound complex shows re-orientation of Gly to form catalytically active HtrA2. B) Conformational changes between the bound and unbound structure. Upon peptide binding, L3 loop (red) of one subunit moves closer to LD* loop (orange) of adjacent unit. The side chain of residues (shown as black sticks) from L3 loop form non-bonded interactions with F131 of LD* loop leading to rearrangement of oxyanion hole due to its hydrophobic packing with LD loop. However, L3 (yellow) and LD* (green) are separated in unbound structure (* indicates residue from the adjacent subunit).

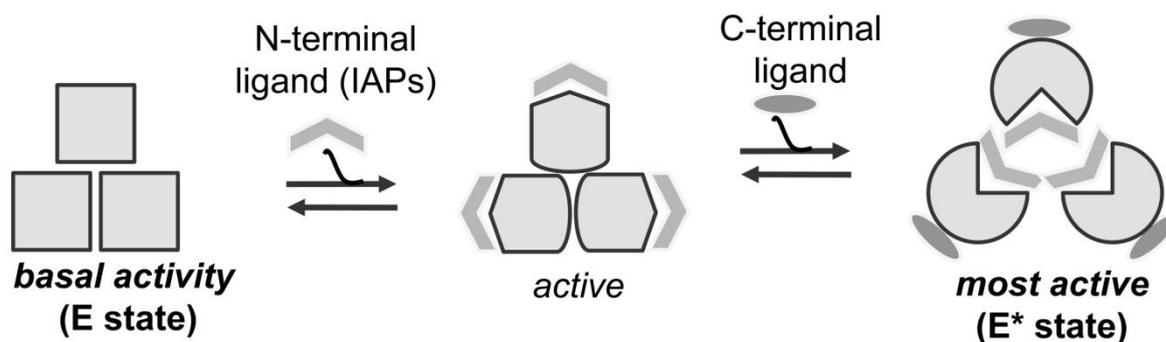


Figure 5.10 Proposed simplistic model for HtrA2 activation. The model assumes that HtrA2 exists in an equilibrium of inactive (E) and most active (E*) states. The complex allosteric propagation mediated upon binding of N- and C-terminal ligands synergistically transforms the protease into the most active conformer. N-terminal ligand (IAP) shifts the conformational equilibrium of HtrA2 to an active state that is further stabilized by reaction with C-terminal ligands/substrate.

5.4 Discussion

HtrA2 acts as a stress-induced serine protease in mammalian cells performing multiple functions [241]. Under physiological conditions, it participates in maintenance of mitochondrial homeostasis however it switches its role from a protector to a pro-apoptotic factor in response to stress-inducing agents. [4]. It induces apoptosis through classical pathways via caspase activation and also by a less understood caspase-independent mechanism [10]. This non-caspase dependent pathway of apoptosis largely depends on its proteolytic activity regulated by protein-protein interactions. To date, wide repertoire of

proteins binding to the C-terminal PDZ domain have been found to stimulate the protease activity thereby promoting cell death [242]. Interestingly, HtrA2 exhibits an additional level of functional modulation mediated through its N-terminus which is evident from IAP binding and its cleavage. This phenomenon emphasizes multiple modes of HtrA2 activation and regulation, the precise mechanism for which remains elusive so far. Furthermore, its association with severe pathophysiological conditions such as cancer and neurodegenerative disorders makes it a promising therapeutic target thereby emphasizing the need to delineate the basis and global mode of its activation.

The available crystal structure of inactive unbound HtrA2 [14] provides a broader picture of its structural organization and mechanism of activation. However, the importance of conformational flexibility required for protease function cannot be unambiguously resolved by the snapshots of selected conformational states provided by crystallography data. The regions with increased flexibility, such as loop segments appear to be preferred for allosteric modulation of serine proteases [243]. The trimeric HtrA2 comprises unique N-terminus and a complex network of regulatory loops (LA, LD, L1, and L2) which forms the 'activation domain'. The relative orientations of this activation domain along with flexible loop L3 and PDZ-protease interface linker might be critical in defining its functions. However, with the partially missing N-terminus and flexible loops, the solved unbound structure could not explain the dynamics that regulates HtrA2 activity as evident from the existing proposed model of activation [14]. Here, with an *in-silico*, enzymology and biochemical approach, we provide evidence of a novel mechanism that regulates HtrA2 activity through its N-terminus.

Enzymology studies with wild-type HtrA2 and its variants using substrate β -casein in presence of varying effectors such as full length XIAP, its minimal binding domains (BIR2 or BIR3) and IBM groove peptides showed significant increment in cleavage rate of the protease. Binding of XIAP proteins or peptides at N-terminus allosterically modulates the

protease and shifts the intrinsic equilibrium from tense to relaxed conformation thereby increasing the overall catalytic efficiency in a PDZ-independent yet synergistic activation mechanism. This N-terminal mediated regulation of HtrA2 seems to be novel among HtrA family of serine proteases. For example, *E coli* DegS, is proteolytically inactive due to inhibitory interaction of PDZ with the serine protease domain which keeps the protease in the basal state [235]. This inhibition is relieved either by addition of OMPs that bind to PDZ and rearrange the active-site or by deletion of PDZ domain to form an active enzyme suggesting that PDZ is the sole regulatory domain [229, 244]. However in HtrA2, we have established that the unique N-terminal region imposes an additional control in regulating its activity. The complexity in its behavior might be required for targeting apoptosis through multiple pathways as well as for myriad of other functions it might perform. Based on our observations we hypothesize that during an apoptotic stress, HtrA2 facilitates ‘caspase-dependent pathway’ by relieving the inhibition of IAPs on active caspases [12, 95]. This IAP-protease interaction then might trigger the non-classical ‘caspase-independent mechanism’ by up-regulating protease activity of HtrA2, which in conjunction with PDZ binding proteins promotes substrate cleavage and hence cell death.

In HtrA family of serine proteases dynamic inter-molecular processes play crucial role in activation and hence functions [226]. The remodeling is initiated in most, if not all HtrA proteins by sensor loop L3 and subsequently a cascade of conformational changes occur along $L3 \rightarrow LD^* \rightarrow L1^*/L2^*$ (the asterisk denotes contribution from a neighboring subunit) which enables the active-site to switch into a ‘proteolytically *ON*’ state. Although the general mechanism is conserved among the family; the activation signal detection might vary [245]. Conformational dynamic analyses revealed that, in unliganded HtrA2 specifically the regulatory loops L1, L3 and LD are disordered, rendering the catalytic site and oxyanion hole dysfunctional. A closer look at and around the active-site of bound HtrA2 shows stabilization

of oxyanion hole and proper positioning of the catalytic triad. These regulatory loops undergo disorder-to-order transition thus allowing formation of a well-defined activation pocket. The occurrence of such structural and conformational flexibility provides a molecular explanation for allosteric activation of HtrA2. From our studies we propose that HtrA2 exists in an equilibrium between an inactive (E) and most active conformation (E*). In E state, the protease exhibits poor enzymatic properties, however, a conformational switch triggered by N-terminal and C-terminal ligands shifts the intrinsic equilibrium towards E* state (**Fig. 5.10**). Thus, the equilibrium between the inactive and active forms of HtrA2 could be used to tune the enzyme in 'ON' or 'OFF' states upon binding of suitable activator or inhibitor molecules.

Since the time of its inception [234], the allosteric mechanism of enzyme regulation has intrigued biochemists and structural biologists, which has resulted in devising ways for manipulating this behavior with desired characteristics. An event occurring at any distal site that affects the catalytic efficiency of an enzyme via long-range communication provides an apt explanation of linkage and cooperativity [246]. Recent studies have emphasized the importance of allostery as an intrinsic property of all dynamic proteins that exist in multiple states at equilibrium and tend to show large conformational changes linked to ligand binding or substrate catalysis [247]. Due to inherent protein-cleaving ability of proteases and their vital biological roles in all living organisms, their enzymatic activity needs to be strictly regulated. HtrA family of proteases have developed an intrinsic allosteric mechanism for precise control of their enzyme activity and specificity which is mediated via regulatory PDZ domains. It can be presumed that the overall mechanism is conserved across the family from *E. coli* DegS, DegP, DegQ, *Mycobacterium* HtrA2, to *Arabidopsis thaliana* Deg1 [235, 245, 248-250]. Interestingly, in human HtrA2, evolution has possibly endowed it with an additional regulatory switch via N-terminal binding of substrates to impart rigid control on its

activation which might be crucial for different roles it plays under normal and diseased conditions [2, 74-76]. This distinctive mode of regulation offers new promising opportunities for specific control and functional modulation of wide biological processes associated with HtrA2. In the recent past, allosteric regulators have been demonstrated to have several advantages over their orthosteric counterparts with lower dose requirements and greater subsite specificity [251, 252]. It would therefore be possible to design and construct suitable peptidomimetics or small molecule analogs, especially BIR2 due to its higher affinity, which could bind the regulatory domains and stimulate protease activity thus promoting apoptosis. Recently, proteolytic activity of HtrA2 was shown to be critical in inducing apoptosis in prostate cancer cells mediated via integrin alpha [253]. Furthermore, its role in degradation of Wilms's tumor suppressor protein WT1 shows the possibility of targeting HtrA2 in therapy of Wilms tumor where WT1 is overexpressed [254, 255]. The E- E* equilibrium therefore proposes new ways to enhance activity of HtrA2 by stabilization of the E* form as shown recently for thrombin [256] and apoptotic procaspase-3 and -6 [98]. This advancement in our understanding toward HtrA2 mechanism of action will help targeting its allosteric site with tailored effectors hence representing a powerful approach for devising therapeutic strategies against varied diseases associated with it.

CHAPTER-6

Concluding remarks

Apoptosis induction in cancer cells provides one of the greatest possibilities to fight this life-threatening disease. Classically, external apoptotic stimuli or signals generated from within the cell activate different signal transduction pathways involving a family of cysteine proteases (caspases), which are the key players in apoptosis. However, the complexity of cancer biology draws interest in search of new molecules that can potentially modulate the cell death pathway. The findings that viral protein E2 induces adaptor independent caspase-8 activation and a serine protease HtrA2 mediates caspase-independent cell death open new avenues in apoptosis and cancer research. Here, with a goal to understand these alternate pathways of apoptosis in depth and to delineate their potential for therapeutic implications, we characterized the structural and functional properties of these proapoptotic proteins using inter-disciplinary approaches.

Part-I: Structural and functional characterization of human papillomavirus E2 protein

The interaction of E2 with procaspase-8 has been found to be independent of DISC formation and therefore represents a novel pathway of apoptosis induction. In this section, we provide fascinating insight into the molecular and structural basis of E2 – procaspase-8 and the classical FADD – procaspase-8 interactions, and their roles in caspase-8 activation and cell death. Understanding the underlying mechanism by which high risk HPV E2 proteins activate caspase-8, and delineating the driving force of the interaction might help develop novel E2 analogs with desired characteristics for therapeutic purposes.

Using *in silico*, site-directed mutagenesis, biochemical, biophysical and *ex vivo* studies, we have defined the surface of physical interaction for both these complexes. E2 – procaspase-8 interaction is mediated via a non-homotypic association between the helix $\alpha 2/\alpha 3$ of E2 TAD with $\alpha 2/\alpha 5$ helices of procaspase-8 DED-B. Sequence alignment of E2 proteins from different HPV types demonstrate that the key residues involved in this

interaction are highly conserved or show conservative substitution. The ability of E2 proteins to induce apoptosis is reported to be independent of its other virus associated functions and is majorly dependent on whether it is localized in the cytoplasm. In support of this hypothesis, a study showed that a cytoplasmic mutant of low risk HPV11 E2 induced apoptosis similar to high risk E2 proteins. We therefore speculate that E2 proteins from different viruses might bind DED-containing proteins via this conserved binding surface to induce apoptosis. For example, a recent study shows that apoptosis induced by HPV16 E2 involves direct interaction with cFLIP, a DED-containing protease-deficient caspase-8 homolog. Therefore, it would be interesting to test this hypothesis further to develop a model for E2-induced apoptosis in different HPV types thereby contributing to our understanding of HPV E2 and its role in cell death.

Caspase-8 activation is achieved through the oligomerization driven by its interaction with the adaptor molecule FADD. While E2 bound to caspase-8 DED-B domain, it did not form a complex with FADD, which contains a single DED. Co-expression of E2 and procaspase-8 result in colocalization of proteins in cytoplasmic punctate structures, which represent oligomerization of caspase-8, resulting into proximity induced auto-cleavage and release of active caspase-8. In order to gain further insight into the mechanism of caspase activation, it is important to establish a relation between the novel adaptor independent E2 induced apoptosis and the classical FADD-mediated pathway.

Although tremendous progress in understanding caspase-8 activation at the DISC has been made over last 10 to 15 years, several issues have remained open, especially at the molecular and structural levels of DED-containing proteins. Recently, a new paradigm in the DISC assembly added further to the complexity of the entire mechanism. Procaspase-8 is recruited not only through interaction with FADD but also by interacting with itself, resulting in formation of DED-chain which in turn facilitates its activation. We undertook the

challenge of dissecting the precise mode and the binding interface for DED-DED interaction which promote DED-chain assembly, and activation of caspase-8. An extensive mutagenesis screening and *in vitro* and *ex vivo* studies led to the revelation of hot-spots and structural modules essential for the intermolecular DED assembly. The homotypic DED-DED interaction involves a typical type-I (helix/helix) – (helix/helix) association between helices $\alpha 1/\alpha 4$ and $\alpha 2/\alpha 5$ for procaspase-8 tandem DED and $\alpha 1/\alpha 4$ and $\alpha 2/\alpha 3$ for the adaptor protein FADD DED. Further, we demonstrate that the signaling competency at DISC is dependent on FADD self-association and its interaction with procaspase-8, and both these events are not mutually exclusive.

Based on the experimental evidences, we propose a model for death-fold mediated procaspase-8 activation wherein, upon stimulation by appropriate death ligand or ectopic increase in local concentration, FADD forms the oligomeric signaling platform which initially recruits procaspase-8 via direct interaction with $\alpha 1/\alpha 4$ of DED-A, followed by sequential interaction mediated by helices $\alpha 2/\alpha 5$ of DED-B to form DED chain. This in turn facilitates the proximity induced auto-cleavage and release of active caspase-8 to promote cell death. As E2 promotes caspase-8 induced apoptosis via interaction at a site distinct to FADD – caspase-8 interaction, this suggests that E2 and FADD might concurrently interact and induce caspase-8 oligomerization, and hence activation. Further studies to establish the direct association between E2 and FADD-mediated signalling will provide important understanding to exploit the pseudo death-fold structure of E2 (helices $\alpha 2/\alpha 3$) for manipulating the caspase-8 cell death pathway in a new way.

Part-II: Structural and functional characterization of serine protease HtrA2

Human HtrA2 is a proapoptotic serine protease that is associated with several diseases such as cancer and neurodegenerative disorders and hence is an important therapeutic target. It

mediates apoptosis through classical caspase-dependent pathway primarily regulated by protein-protein interactions, and also by a less understood caspase-independent mechanism which depends on its proteolytic activity. The protease has a trimeric pyramidal architecture comprising a short N-terminal region, a serine protease domain and a C-terminal protein-protein interaction (PDZ) domain. The crystal structure of the unbound inactive form of the protease was reported in 2002 (Li et. al.) which proposed a model where the relative PDZ-protease movement was considered the primary reason for regulation of HtrA2 activity. However, the flexible PDZ-protease linker, N-terminal and several critical regulatory loops were missing from the structure thus limiting its ability to demonstrate dynamics of the mechanism of action. The structure-based activation model proclaims that stereo-chemical hindrance by PDZ restricts substrate accessibility in an otherwise proteolytically competent catalytic pocket thus hinting higher activity for Δ PDZ variant. However, recent reports from our laboratory demonstrated significant decrease in protease activity of this variant thus highlighting the importance of intricate co-ordination among different domains in defining HtrA2 activity and specificity. Interestingly, although C-terminal PDZ acts as a regulatory domain in all the members of HtrA family, uniqueness of human HtrA2 is manifested by its ability to bind subset of proteins such as IAPs through its N-terminus and subsequently cleave them. These observations accentuates existence of multiple or complex mechanism of HtrA2 activation involving PDZ and other regions of the protein which remains elusive so far.

Here, we demonstrate role of the short N-terminal region in modulation of HtrA2 protease activity in synergy with PDZ domain. The work showcases a combination of computational, biochemical, biophysical and functional enzymology approach to clearly show how precise coordination between ligand binding and flexible loop movements at a site distal from catalytic pocket regulates HtrA2 proteolytic activity. It is the first extensive

quantitative kinetics study carried out for HtrA2 with its N-terminal binding partner, X-linked IAP (XIAP) protein and its minimal binding peptides. Our findings distinctly demonstrate a novel N-terminal ligand mediated triggering of an allosteric switch essential for transforming HtrA2 to a proteolytically competent state in a PDZ-independent yet synergistic activation process. Dynamic analyses suggest that it occurs through a series of coordinated structural reorganizations at distal regulatory loops (L3, LD, L1) leading to population shift toward relaxed conformer. We believe that this precise synergistic coordination among different domains of HtrA2 might be physiologically relevant to bestow tighter control upon its activation, which perfectly commensurate with the variety of functions it performs in the cell. Furthermore, the peptide based activation analyses in our study highlight the possibility of designing suitable peptidomimetics or small molecule analogs of XIAP, so as to manipulate HtrA2 functions specifically for disease intervention.

CHAPTER-7

Future perspective

The current study has set up foundation for exploring several intriguing questions:

1. The residue level detailed information obtained from HPV18 E2 and procaspase-8 interaction analyses has shed new lights on role of E2 in apoptosis. The conservation of the interface residues across different HPV types suggests similar mechanism of apoptosis induction by E2 proteins. It will be therefore interesting to perform analogous studies with E2 protein from other viruses (including both low and high risk types) particularly targeting the key residues identified in the present study, and attempt to develop a model for E2-induced apoptosis.
2. Knowledge of the structure and dynamics of proteins and protein assemblies is critical both for understanding the molecular basis of physiological and patho-physiological processes and for guiding drug design. Structural analysis of a novel E2-procaspase-8 interaction in atomic detail therefore becomes imperative. Owing to the poor solution behavior of these proteins, a high resolution structure of the full length complex is however impractical. Strikingly, the information on minimal binding regions now could be utilized to determine the structure and dynamics of the E2-procaspase-8 complex. These atomic details could be utilized to design E2 analogs with enhanced proapoptotic properties that might have potential for disease intervention.
3. Evidence from the current study suggests that initiation of apoptosis by E2 and FADD-mediated procaspase-8 activation occur due to oligomerization of caspase-8 induced through distinct mechanisms. It would be thus interesting to determine whether E2 induced cell death can co-operate with FADD to enhance caspase activation. Furthermore, it would be important to address how the activation of procaspase-8 is regulated in the DED chain. Establishing a direct link between these pathways would be necessary to understand the effect of the protein-protein interactions on DISC

formation and its overall function with an aim at targeting the extrinsic cell death pathway from a different angle.

4. Multidomain proteins due to their structural complexity require different levels of regulatory mechanisms for executing cellular functions efficiently within a specified time period. Allosteric modulation is one such way which often helps a protein such as an enzyme to regulate functional behavior. Although we could provide the first evidence of the complex allosteric regulation of HtrA2, the mechanism of allosteric propagation and the underlying conformational plasticity needs to be delineated. Identification of the control switch regulating the protease activity will provide mechanistic details of its association in binding to myriads of cellular substrate and their cleavage. Furthermore, it will provide an idea of how HtrA2 might be activated *in vivo* in presence of different stimuli. These advances in our understanding of HtrA2 mechanism and function will help design allosteric modulators to manipulate its functions with desired characteristics.

References

1. Thierry, F. and C. Demeret, *Direct activation of caspase 8 by the proapoptotic E2 protein of HPV18 independent of adaptor proteins*. Cell Death Differ, 2008. **15**(9): p. 1356-63.
2. Suzuki, Y.T.-N., K., et al., *Mitochondrial protease Omi/HtrA2 enhances caspase activation through multiple pathways*. Cell Death Differ., 2004. **11**: p. 208-216.
3. Yang, Q., et al., *Omi/HtrA2 catalytic cleavage of inhibitor of apoptosis (IAP) irreversibly inactivates IAPs and facilitates caspase activity in apoptosis*. Genes Dev., 2003. **17**: p. 1487-1496.
4. Muller, M. and C. Demeret, *The HPV E2-Host Protein-Protein Interactions: A Complex Hijacking of the Cellular Network*. Open Virol J, 2012. **6**: p. 173-89.
5. Parish, J.L., et al., *E2 proteins from high- and low-risk human papillomavirus types differ in their ability to bind p53 and induce apoptotic cell death*. J Virol, 2006. **80**(9): p. 4580-90.
6. Desaintes, C., et al., *Expression of the papillomavirus E2 protein in HeLa cells leads to apoptosis*. EMBO. J., 1997. **16**: p. 504-514.
7. Blachon, S. and C. Demeret, *The regulatory E2 proteins of human genital papillomaviruses are pro-apoptotic*. Biochimie, 2003. **85**(8): p. 813-9.
8. Thierry, F. and C. Demeret, *Direct activation of caspase 8 by the proapoptotic E2 protein of HPV18 independent of adaptor proteins*. Cell Death Differ. , 2008. **15**: p. 1356-1363.
9. Singh, N., R.R. Kuppili, and K. Bose, *The structural basis of mode of activation and functional diversity: a case study with HtrA family of serine proteases*. Arch Biochem Biophys, 2011. **516**(2): p. 85-96.
10. Vande Walle, L., M. Lamkanfi, and P. Vandenabeele, *The mitochondrial serine protease HtrA2/Omi: an overview*. Cell Death Differ, 2008. **15**(3): p. 453-60.
11. Deveraux, Q.L. and J.C. Reed, *IAP family proteins--suppressors of apoptosis*. Genes Dev, 1999. **13**(3): p. 239-52.
12. Hegde, R., et al., *Identification of Omi/HtrA2 as a mitochondrial apoptotic serine protease that disrupts inhibitor of apoptosis protein-caspase interaction*. J Biol Chem, 2002. **277**: p. 432-438.
13. Trencia, A., et al., *Omi/HtrA2 promotes cell death by binding and degrading the anti-apoptotic protein ped/pea-15*. J Biol Chem., 2004. **279**: p. 46566-46572.
14. Li, W., et al., *Structural insights into the pro-apoptotic function of mitochondrial serine protease HtrA2/Omi*. Nat. Struct. Biol., 2002. **9**(6): p. 436-441.
15. Srinivasula, S.M., et al., *Inhibitor of apoptosis proteins are substrates for the mitochondrial serine protease Omi/HtrA2*. J Biol Chem, 2003. **278**(34): p. 31469-72.
16. Fuentes-Gonzalez, A.M., et al., *The modulation of apoptosis by oncogenic viruses*. Virol J, 2013. **10**: p. 182.
17. McLaughlin-Drubin, M.E. and K. Munger, *Viruses associated with human cancer*. Biochim Biophys Acta, 2008. **1782**(3): p. 127-50.
18. Molho-Pessach, V. and M. Lotem, *Viral carcinogenesis in skin cancer*. Curr Probl Dermatol, 2007. **35**: p. 39-51.
19. Munoz, N., et al., *Epidemiologic classification of human papillomavirus types associated with cervical cancer*. N Engl J Med, 2003. **348**(6): p. 518-27.
20. Burd, E., *Human Papillomavirus and Cervical Cancer*. Clin Microbiol Rev. , 2003. **16**: p. 1-17.

21. Bruni, L., et al., *Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings*. J Infect Dis, 2010. **202**(12): p. 1789-99.
22. Williams, V.M., et al., *HPV-DNA integration and carcinogenesis: putative roles for inflammation and oxidative stress*. Future Virol, 2011. **6**(1): p. 45-57.
23. Graham, S.V., *Human papillomavirus: gene expression, regulation and prospects for novel diagnostic methods and antiviral therapies*. Future Microbiol, 2010. **5**(10): p. 1493-506.
24. Sakakibara, N., R. Mitra, and A.A. McBride, *The papillomavirus E1 helicase activates a cellular DNA damage response in viral replication foci*. J Virol, 2011. **85**(17): p. 8981-95.
25. Venuti, A., et al., *Papillomavirus E5: the smallest oncoprotein with many functions*. Mol Cancer, 2011. **10**: p. 140.
26. Fehrman, F., D.J. Klumpp, and L.A. Laimins, *Human papillomavirus type 31 E5 protein supports cell cycle progression and activates late viral functions upon epithelial differentiation*. J Virol, 2003. **77**(5): p. 2819-31.
27. Howie, H.L., R.A. Katzenellenbogen, and D.A. Galloway, *Papillomavirus E6 proteins*. Virology, 2009. **384**(2): p. 324-34.
28. Wise-Draper, T.M. and S.I. Wells, *Papillomavirus E6 and E7 proteins and their cellular targets*. Front Biosci, 2008. **13**: p. 1003-17.
29. Cason, J., et al., *Expression of human papillomavirus type 16 (HPV-16) major (L1) and minor (L2) capsid proteins in insect cells as polyhistidine fusion proteins*. Biochem Soc Trans, 1994. **22**(3): p. 336S.
30. Longworth, M.S. and L.A. Laimins, *Pathogenesis of human papillomaviruses in differentiating epithelia*. Microbiol Mol Biol Rev, 2004. **68**(2): p. 362-72.
31. Moody, C.A. and L.A. Laimins, *Human papillomavirus oncoproteins: pathways to transformation*. Nat Rev Cancer, 2010. **10**(8): p. 550-60.
32. Bergvall, M., T. Melendy, and J. Archambault, *The E1 proteins*. Virology, 2013. **445**(1-2): p. 35-56.
33. Kasukawa, H., P.M. Howley, and J.D. Benson, *A fifteen-amino-acid peptide inhibits human papillomavirus E1-E2 interaction and human papillomavirus DNA replication in vitro*. J Virol, 1998. **72**(10): p. 8166-73.
34. White, P.W., A.M. Faucher, and N. Goudreau, *Small molecule inhibitors of the human papillomavirus E1-E2 interaction*. Curr Top Microbiol Immunol, 2011. **348**: p. 61-88.
35. Dowhanick, J.J., A.A. McBride, and P.M. Howley, *Suppression of cellular proliferation by the papillomavirus E2 protein*. J Virol, 1995. **69**(12): p. 7791-9.
36. Bontkes, H.J., et al., *Human papillomavirus type 16 E2-specific T-helper lymphocyte responses in patients with cervical intraepithelial neoplasia*. J Gen Virol, 1999. **80** (Pt 9): p. 2453-9.
37. Han, R., et al., *Protection of rabbits from viral challenge by gene gun-based intracutaneous vaccination with a combination of cottontail rabbit papillomavirus E1, E2, E6, and E7 genes*. J Virol, 1999. **73**(8): p. 7039-43.
38. Gu, W., et al., *Inhibition of cervical cancer cell growth in vitro and in vivo with lentiviral-vector delivered short hairpin RNA targeting human papillomavirus E6 and E7 oncogenes*. Cancer Gene Ther, 2006. **13**(11): p. 1023-32.
39. Kanodia, S., D.M. Da Silva, and W.M. Kast, *Recent advances in strategies for immunotherapy of human papillomavirus-induced lesions*. Int J Cancer, 2008. **122**(2): p. 247-59.

40. Yim, E.K. and J.S. Park, *The role of HPV E6 and E7 oncoproteins in HPV-associated cervical carcinogenesis*. *Cancer Res Treat*, 2005. **37**(6): p. 319-24.
41. Demeret, C., A. Garcia-Carranca, and F. Thierry, *Transcription-independent triggering of the extrinsic pathway of apoptosis by human papillomavirus 18 E2 protein*. *Oncogene*, 2003. **22**(2): p. 168-75.
42. Blachon, S., et al., *Nucleo-cytoplasmic shuttling of high risk human Papillomavirus E2 proteins induces apoptosis*. *J Biol Chem*, 2005. **280**(43): p. 36088-98.
43. Adams, A.K., T.M. Wise-Draper, and S.I. Wells, *Human papillomavirus induced transformation in cervical and head and neck cancers*. *Cancers (Basel)*, 2014. **6**(3): p. 1793-820.
44. Elmore, S., *Apoptosis: a review of programmed cell death*. *Toxicol Pathol*, 2007. **35**(4): p. 495-516.
45. Jeong, S.Y. and D.W. Seol, *The role of mitochondria in apoptosis*. *BMB Rep*, 2008. **41**(1): p. 11-22.
46. Fulda, S. and K.M. Debatin, *Extrinsic versus intrinsic apoptosis pathways in anticancer chemotherapy*. *Oncogene*, 2006. **25**(34): p. 4798-811.
47. Massimi, P., et al., *Interaction between the HPV-16 E2 transcriptional activator and p53*. *Oncogene*, 1999. **18**: p. 7748-7754.
48. Demeret, C., A. Garcia-Carranca, and F. Thierry, *Transcription-independent triggering of the extrinsic pathway of apoptosis by human papillomavirus 18 E2 protein*. *Oncogene*, 2003. **22**(2): p. 168-175.
49. Wang, W., et al., *The relationship between c-FLIP expression and human papillomavirus E2 gene disruption in cervical carcinogenesis*. *Gynecol Oncol.* , 2007. **105**: p. 571-577.
50. Park, H.H., et al., *The death domain superfamily in intracellular signaling of apoptosis and inflammation*. *Annu Rev Immunol*, 2007. **25**: p. 561-86.
51. Mace, P.D. and S.J. Riedl, *Molecular cell death platforms and assemblies*. *Curr Opin Cell Biol*, 2010. **22**(6): p. 828-36.
52. Scott, F.L., et al., *The Fas-FADD death domain complex structure unravels signalling by receptor clustering*. *Nature*, 2009. **457**(7232): p. 1019-22.
53. Berglund, H., et al., *The three-dimensional solution structure and dynamic properties of the human FADD death domain*. *J Mol Biol*, 2000. **302**(1): p. 171-88.
54. Sukits, S.F., et al., *Solution structure of the tumor necrosis factor receptor-1 death domain*. *J Mol Biol*, 2001. **310**(4): p. 895-906.
55. Liepinsh, E., et al., *NMR structure of the death domain of the p75 neurotrophin receptor*. *EMBO J*, 1997. **16**(16): p. 4999-5005.
56. Lasker, M.V., M.M. Gajjar, and S.K. Nair, *Cutting edge: molecular structure of the IL-1R-associated kinase-4 death domain and its implications for TLR signaling*. *J Immunol*, 2005. **175**(7): p. 4175-9.
57. Park, H.H. and H. Wu, *Crystal structure of RAIDD death domain implicates potential mechanism of PIDDosome assembly*. *J Mol Biol*, 2006. **357**(2): p. 358-64.
58. Lin, S.C., Y.C. Lo, and H. Wu, *Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling*. *Nature*, 2010. **465**(7300): p. 885-90.
59. Park, H.H., et al., *Death domain assembly mechanism revealed by crystal structure of the oligomeric PIDDosome core complex*. *Cell*, 2007. **128**(3): p. 533-46.

60. Martin, D.A., et al., *Defective CD95/APO-1/Fas signal complex formation in the human autoimmune lymphoproliferative syndrome, type Ia*. Proc Natl Acad Sci U S A, 1999. **96**(8): p. 4552-7.
61. Hill, J.M., et al., *Identification of an expanded binding surface on the FADD death domain responsible for interaction with CD95/Fas*. J Biol Chem, 2004. **279**(2): p. 1474-81.
62. Park, A. and V.R. Baichwal, *Systematic mutational analysis of the death domain of the tumor necrosis factor receptor 1-associated protein TRADD*. J Biol Chem, 1996. **271**(16): p. 9858-62.
63. Telliez, J.B., et al., *Mutational analysis and NMR studies of the death domain of the tumor necrosis factor receptor-1*. J Mol Biol, 2000. **300**(5): p. 1323-33.
64. Eberstadt, M., et al., *NMR structure and mutagenesis of the FADD (Mort1) death-effector domain*. Nature, 1998. **392**(6679): p. 941-5.
65. Hill, J.M., et al., *Recognition of ERK MAP kinase by PEA-15 reveals a common docking site within the death domain and death effector domain*. EMBO J, 2002. **21**(23): p. 6494-504.
66. Yang, J.K., et al., *Crystal structure of MC159 reveals molecular mechanism of DISC assembly and FLIP inhibition*. Mol Cell, 2005. **20**(6): p. 939-49.
67. Carrington, P.E., et al., *The structure of FADD and its mode of interaction with procaspase-8*. Mol Cell, 2006. **22**(5): p. 599-610.
68. Hill, J., et al., *Recognition of ERK MAP kinase by PEA-15 reveals a common docking site within the death domain and death effector domain*. EMBO J., 2002. **21**: p. 6494-6504.
69. Xiao, T., et al., *Three-dimensional structure of a complex between the death domains of Pelle and Tube*. Cell, 1999. **99**(5): p. 545-55.
70. Chou, J.J., et al., *Solution structure of the RAIDD CARD and model for CARD/CARD interaction in caspase-2 and caspase-9 recruitment*. Cell, 1998. **94**(2): p. 171-80.
71. Pallen, M.J. and B.W. Wren, *The HtrA family of serine proteases*. Mol Microbiol, 1997. **26**(2): p. 209-21.
72. Spiess, C., A. Beil, and M. Ehrmann, *A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein*. Cell, 1999. **97**(3): p. 339-47.
73. Krojer, T., et al., *Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine*. Nature, 2002. **416**(6879): p. 455-459.
74. Chien, J., et al., *HtrA serine proteases as potential therapeutic targets in cancer*. Curr Cancer Drug Targets, 2009. **9**: p. 451-468.
75. Polur, I., et al., *Role of HTRA1, a serine protease, in the progression of articular cartilage degeneration*. Histol Histopathol, 2010. **25**: p. 599-608.
76. Zurawa-Janicka, D., J. Skorko-Glonek, and B. Lipinska, *HtrA proteins as targets in therapy of cancer and other diseases*. Expert Opin Ther Targets, 2010. **14**: p. 665-679.
77. Lipinska, B., et al., *Identification, characterization, and mapping of the Escherichia coli htrA gene, whose product is essential for bacterial growth only at elevated temperatures*. J Bacteriol, 1989. **171**: p. 1574-1584.
78. Zumbunn, J. and B. Trueb, *Primary structure of a putative serine protease specific for IGF-binding proteins*. FEBS Lett, 1996. **398**(2-3): p. 187-92.
79. Faccio, L., et al., *Characterization of a Novel Human Serine Protease That Has Extensive Homology to Bacterial Heat Shock Endoprotease HtrA and Is Regulated by Kidney Ischemia*. J Biol Chem., 2000. **275**: p. 2581-2588.

80. Nie, G.Y., et al., *Identification and cloning of two isoforms of human high-temperature requirement factor A3 (HtrA3), characterization of its genomic structure and comparison of its tissue distribution with HtrA1 and HtrA2*. *Biochem J*, 2003. **371**(Pt 1): p. 39-48.
81. Ota, T., et al., *Complete sequencing and characterization of 21,243 full-length human cDNAs*. *Nat Genet*, 2004. **36**(1): p. 40-5.
82. Chien, J., et al., *HtrA serine proteases as potential therapeutic targets in cancer*. *Curr Cancer Drug Targets*, 2009. **9**(4): p. 451-68.
83. Murwantoko., et al., *Binding of proteins to the PDZ domain regulates proteolytic activity of HtrA1 serine protease*. *Biochem J*, 2004. **381**: p. 895-904.
84. Grau, S., et al., *Implications of the serine protease HtrA1 in amyloid precursor protein processing*. *Proc Natl Acad Sci USA*, 2005. **102**: p. 6021-6026.
85. Tsuchiya, A., et al., *Expression of mouse HtrA1 serine protease in normal bone and cartilage and its upregulation in joint cartilage damaged by experimental arthritis*. *Bone*, 2005. **37**: p. 323-336.
86. Milner, J., A. Patel, and A. Rowan, *Emerging roles of serine proteinases in tissue turnover in arthritis*. *Arthritis Rheum.*, 2008. **58**: p. 3644-3656.
87. Tocharus, J., et al., *Developmentally regulated expression of mouse HtrA3 and its role as an inhibitor of TGF-beta signaling*. *Dev Growth Differ*, 2004. **46**(3): p. 257-74.
88. Martins, L., et al., *The serine protease Omi/HtrA2 regulates apoptosis by binding XIAP through a reaper-like motif*. *J Biol Chem*, 2002. **277**: p. 439-444.
89. Verhagen, A.M. and D.L. Vaux, *Cell death regulation by the mammalian IAP antagonist Diablo/Smac*. *Apoptosis*, 2002. **7**(2): p. 163-6.
90. Kadomatsu, T., M. Mori, and K. Terada, *Mitochondrial import of Omi: the definitive role of the putative transmembrane region and multiple processing sites in the amino-terminal segment*. *Biochem Biophys Res Commun*, 2007. **361**(2): p. 516-21.
91. Jones, J., et al., *Loss of Omi mitochondrial protease activity causes the neuromuscular disorder of mnd2 mutant mice*. *Nature*, 2003. **425**: p. 721-725.
92. Krick, S., et al., *Mpv17l protects against mitochondrial oxidative stress and apoptosis by activation of Omi/HtrA2 protease*. *Proc Natl Acad Sci U S A*, 2008. **105**(37): p. 14106-11.
93. Moiso, N., et al., *Mitochondrial dysfunction triggered by loss of HtrA2 results in the activation of a brain-specific transcriptional stress response*. *Cell Death Differ*, 2009. **16**(3): p. 449-64.
94. Ding, X., et al., *Enhanced HtrA2/Omi expression in oxidative injury to retinal pigment epithelial cells and murine models of neurodegeneration*. *Invest Ophthalmol Vis Sci*, 2009. **50**(10): p. 4957-66.
95. Suzuki, Y., et al., *A serine protease, HtrA2, is released from the mitochondria and interacts with XIAP, inducing cell death*. *Mol. Cell.*, 2001. **8**(3): p. 613-621.
96. Eckelman, B.P., et al., *The mechanism of peptide-binding specificity of IAP BIR domains*. *Cell Death and Differentiation*, 2008. **15**(5): p. 920-928.
97. Wilkinson, J.C., et al., *Upstream regulatory role for XIAP in receptor-mediated apoptosis*. *Mol Cell Biol*, 2004. **24**(16): p. 7003-14.
98. Wolan, D.W., et al., *Small-molecule activators of a proenzyme*. *Science*, 2009. **326**(5954): p. 853-8.
99. Bhuiyan, M. and F. Kohji, *Mitochondrial Serine Protease HtrA2/Omi as a Potential Therapeutic Target*. *Curr Drug Targets*, 2009. **10**: p. 372-383.

100. Bhuiyan, M.S. and K. Fukunaga, *Activation of HtrA2, a mitochondrial serine protease mediates apoptosis: current knowledge on HtrA2 mediated myocardial ischemia/reperfusion injury*. Cardiovasc Ther, 2008. **26**(3): p. 224-32.
101. Krijnen, P.A., et al., *Apoptosis in myocardial ischaemia and infarction*. J Clin Pathol, 2002. **55**(11): p. 801-11.
102. Jin, S., et al., *CIAP1 and the serine protease HTRA2 are involved in a novel p53-dependent apoptosis pathway in mammals*. Genes Dev, 2003. **17**(3): p. 359-67.
103. Bhuiyan, M.S. and K. Fukunaga, *Inhibition of HtrA2/Omi ameliorates heart dysfunction following ischemia/reperfusion injury in rat heart in vivo*. Eur J Pharmacol, 2007. **557**(2-3): p. 168-77.
104. Su, D., et al., *UCF-101, a novel Omi/HtrA2 inhibitor, protects against cerebral ischemia/reperfusion injury in rats*. Anat Rec (Hoboken), 2009. **292**(6): p. 854-61.
105. Qiu, X.B. and A.L. Goldberg, *The membrane-associated inhibitor of apoptosis protein, BRUCE/Apollon, antagonizes both the precursor and mature forms of Smac and caspase-9*. J Biol Chem, 2005. **280**(1): p. 174-82.
106. Sekine, K., et al., *HtrA2 cleaves Apollon and induces cell death by IAP-binding motif in Apollon-deficient cells*. Biochem Biophys Res Commun, 2005. **330**(1): p. 279-85.
107. Kuninaka, S., et al., *Serine protease Omi/HtrA2 targets WARTS kinase to control cell proliferation*. Oncogene, 2007. **26**: p. 2395-2406.
108. Kuninaka, S., et al., *The tumor suppressor WARTS activates the Omi / HtrA2-dependent pathway of cell death*. Oncogene, 2005. **24**(34): p. 5287-98.
109. Hong, S.K., M.K. Cha, and I.H. Kim, *Specific protein interaction of human Pag with Omi/HtrA2 and the activation of the protease activity of Omi/HtrA2*. Free Radic Biol Med, 2006. **40**(2): p. 275-84.
110. Suzuki, Y., et al., *Mitochondrial protease Omi/HtrA2 enhances caspase activation through multiple pathways*. Cell Death Differ, 2004. **11**(2): p. 208-16.
111. Marabese, M., et al., *HtrA2 enhances the apoptotic functions of p73 on bax*. Cell Death Differ, 2008. **15**(5): p. 849-58.
112. Lee, S.H., et al., *Immunohistochemical analysis of Omi/HtrA2 expression in stomach cancer*. APMIS, 2003. **111**(5): p. 586-90.
113. Narkiewicz, J., et al., *Changes in mRNA and protein levels of human HtrA1, HtrA2 and HtrA3 in ovarian cancer*. Clin Biochem, 2008. **41**(7-8): p. 561-9.
114. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications*. Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10869-74.
115. Sorlie, T., et al., *Repeated observation of breast tumor subtypes in independent gene expression data sets*. Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8418-23.
116. Hu, X.Y., et al., *Immunohistochemical analysis of Omi/HtrA2 expression in prostate cancer and benign prostatic hyperplasia*. APMIS, 2006. **114**(12): p. 893-8.
117. Lee, H.J. and J.J. Zheng, *PDZ domains and their binding partners: structure, specificity, and modification*. Cell Commun Signal, 2010. **8**: p. 8.
118. Bejugam, P.R., et al., *Allosteric regulation of serine protease HtrA2 through novel non-canonical substrate binding pocket*. PLoS One, 2013. **8**(2): p. e55416.
119. Chaganti, L.K., R.R. Kuppili, and K. Bose, *Intricate structural coordination and domain plasticity regulate activity of serine protease HtrA2*. FASEB J, 2013. **27**(8): p. 3054-66.

120. Gupta, S., *The C-terminal Tail of Presenilin Regulates Omi/HtrA2 Protease Activity*. Journal of Biological Chemistry, 2004. **279**(44): p. 45844-45854.
121. Martins, L., et al., *Binding specificity and regulation of the serine protease and PDZ domains of HtrA2/Omi*. J Biol Chem., 2003. **278**: p. 49417-49427.
122. Verhagen, A.M., et al., *HtrA2 promotes cell death through its serine protease activity and its ability to antagonize inhibitor of apoptosis proteins*. J Biol Chem, 2002. **277**(1): p. 445-54.
123. Inoue, H., H. Nojima, and H. Okayama, *High efficiency transformation of Escherichia coli with plasmids*. Gene, 1990. **96**(1): p. 23-8.
124. Jordan, M., A. Schallhorn, and F.M. Wurm, *Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation*. Nucleic Acids Res, 1996. **24**(4): p. 596-601.
125. Walters, J., S.L. Milam, and A.C. Clark, *Practical approaches to protein folding and assembly: spectroscopic strategies in thermodynamics and kinetics*. Methods Enzymol, 2009. **455**: p. 1-39.
126. Scheffner, M., et al., *The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53*. Cell, 1993. **75**(3): p. 495-505.
127. Filippova, M., L. Parkhurst, and P.J. Duerksen-Hughes, *The human papillomavirus 16 E6 protein binds to Fas-associated death domain and protects cells from Fas-triggered apoptosis*. J Biol Chem, 2004. **279**(24): p. 25729-44.
128. Desaintes, C., et al., *Expression of the papillomavirus E2 protein in HeLa cells leads to apoptosis*. EMBO J, 1997. **16**(3): p. 504-14.
129. McMurray, H.R., et al., *Biology of human papillomaviruses*. Int J Exp Pathol, 2001. **82**(1): p. 15-33.
130. Wang, W., et al., *Triggering of death receptor apoptotic signaling by human papillomavirus 16 E2 protein in cervical cancer cell lines is mediated by interaction with c-FLIP*. Apoptosis, 2011. **16**(1): p. 55-66.
131. Blachon, S., et al., *Nucleo-cytoplasmic shuttling of high-risk HPV E2 proteins induces apoptosis*. J. Biol. Chem., 2005. **280**: p. 36088-36098.
132. Sessler, T., et al., *Structural determinants of DISC function: new insights into death receptor-mediated apoptosis signalling*. Pharmacol Ther, 2013. **140**(2): p. 186-99.
133. Shi, Y., *Caspase activation: revisiting the induced proximity model*. Cell, 2004. **117**(7): p. 855-8.
134. Slee, E.A., C. Adrain, and S.J. Martin, *Executioner caspase-3, -6, and -7 perform distinct, non-redundant roles during the demolition phase of apoptosis*. J Biol Chem, 2001. **276**(10): p. 7320-6.
135. Chang, D.W., et al., *Oligomerization is a general mechanism for the activation of apoptosis initiator and inflammatory procaspases*. J Biol Chem, 2003. **278**(19): p. 16466-9.
136. Medema, J.P., et al., *FLICE is activated by association with the CD95 death-inducing signaling complex (DISC)*. EMBO J, 1997. **16**(10): p. 2794-804.
137. Kumar, S. and P.A. Colussi, *Prodomains--adaptors--oligomerization: the pursuit of caspase activation in apoptosis*. Trends Biochem Sci, 1999. **24**(1): p. 1-4.
138. Gervais, F.G., et al., *Recruitment and activation of caspase-8 by the Huntingtin-interacting protein Hip-1 and a novel partner Hipp1*. Nat Cell Biol, 2002. **4**(2): p. 95-105.
139. Webb, B. and A. Sali, *Comparative Protein Structure Modeling Using MODELLER*. Curr Protoc Bioinformatics, 2014. **47**: p. 5 6 1-5 6 32.

140. Laskowski, R.A., D.S. Moss, and J.M. Thornton, *Main-chain bond lengths and bond angles in protein structures*. J Mol Biol, 1993. **231**(4): p. 1049-67.
141. Harris, S.F. and M.R. Botchan, *Crystal structure of the human papillomavirus type 18 E2 activation domain*. Science, 1999. **284**(5420): p. 1673-7.
142. Antson, A.A., et al., *Structure of the intact transactivation domain of the human papillomavirus E2 protein*. Nature, 2000. **403**(6771): p. 805-9.
143. de Vries, S.J., M. van Dijk, and A.M. Bonvin, *The HADDOCK web server for data-driven biomolecular docking*. Nat Protoc, 2010. **5**(5): p. 883-97.
144. Vriend, G., *WHAT IF: a molecular modeling and drug design program*. J Mol Graph, 1990. **8**(1): p. 52-6, 29.
145. Laskowski, R.A., V.V. Chistyakov, and J.M. Thornton, *PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids*. Nucleic Acids Res, 2005. **33**(Database issue): p. D266-8.
146. de Beer, T.A., et al., *PDBsum additions*. Nucleic Acids Res, 2014. **42**(Database issue): p. D292-6.
147. Tina, K.G., R. Bhadra, and N. Srinivasan, *PIC: Protein Interactions Calculator*. Nucleic Acids Res, 2007. **35**(Web Server issue): p. W473-6.
148. Reichmann, D., et al., *The molecular architecture of protein-protein binding sites*. Curr Opin Struct Biol, 2007. **17**(1): p. 67-76.
149. Bogan, A.A. and K.S. Thorn, *Anatomy of hot spots in protein interfaces*. J Mol Biol, 1998. **280**(1): p. 1-9.
150. Clackson, T. and J.A. Wells, *A hot spot of binding energy in a hormone-receptor interface*. Science, 1995. **267**(5196): p. 383-6.
151. Webster, K., et al., *The human papillomavirus (HPV) 16 E2 protein induces apoptosis in the absence of other HPV proteins and via a p53-dependent pathway*. J Biol Chem, 2000. **275**(1): p. 87-94.
152. Cavallo, L., J. Kleinjung, and F. Fraternali, *POPS: A fast algorithm for solvent accessible surface areas at atomic and residue level*. Nucleic Acids Res, 2003. **31**(13): p. 3364-6.
153. Robert, X. and P. Gouet, *Deciphering key features in protein structures with the new ENDscript server*. Nucleic Acids Res, 2014. **42**(Web Server issue): p. W320-4.
154. Siegel, R.M., et al., *Death-effector filaments: novel cytoplasmic structures that recruit caspases and trigger apoptosis*. J Cell Biol, 1998. **141**(5): p. 1243-53.
155. Dickens, L.S., et al., *A death effector domain chain DISC model reveals a crucial role for caspase-8 chain assembly in mediating apoptotic cell death*. Mol Cell, 2012. **47**(2): p. 291-305.
156. Yang, X., H.Y. Chang, and D. Baltimore, *Autoproteolytic activation of pro-caspases by oligomerization*. Mol Cell, 1998. **1**(2): p. 319-25.
157. Muzio, M., et al., *An induced proximity model for caspase-8 activation*. J Biol Chem, 1998. **273**(5): p. 2926-30.
158. Schleich, K., et al., *Stoichiometry of the CD95 death-inducing signaling complex: experimental and modeling evidence for a death effector domain chain model*. Mol Cell, 2012. **47**(2): p. 306-19.
159. Majkut, J., et al., *Differential affinity of FLIP and procaspase 8 for FADD's DED binding surfaces regulates DISC assembly*. Nat Commun, 2014. **5**: p. 3350.
160. Kersse, K., et al., *The death-fold superfamily of homotypic interaction motifs*. Trends Biochem Sci, 2011. **36**(10): p. 541-52.

161. Wang, L., et al., *The Fas-FADD death domain complex structure reveals the basis of DISC assembly and disease mutations*. Nat Struct Mol Biol, 2010. **17**(11): p. 1324-9.
162. Marikar, F.M., et al., *Expression of recombinant human FADD, preparation of its polyclonal antiserum and the application in immunoassays*. Cell Mol Immunol, 2008. **5**(6): p. 471-4.
163. Comeau, S.R., et al., *ClusPro: a fully automated algorithm for protein-protein docking*. Nucleic Acids Res, 2004. **32**(Web Server issue): p. W96-9.
164. Comeau, S.R., et al., *ClusPro: an automated docking and discrimination method for the prediction of protein complexes*. Bioinformatics, 2004. **20**(1): p. 45-50.
165. Whitmore, L. and B.A. Wallace, *DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data*. Nucleic Acids Res, 2004. **32**(Web Server issue): p. W668-73.
166. Li, F.Y., et al., *Crystal structure of a viral FLIP: insights into FLIP-mediated inhibition of death receptor signaling*. J Biol Chem, 2006. **281**(5): p. 2960-8.
167. Yang, J., et al., *Crystal structure of MC159 reveals molecular mechanism of DISC assembly and FLIP inhibition*. Mol Cell., 2005. **20**: p. 939-949.
168. Ferrao, R. and H. Wu, *Helical assembly in the death domain (DD) superfamily*. Curr Opin Struct Biol, 2012. **22**(2): p. 241-7.
169. MacFarlane, M., et al., *Active caspases and cleaved cytokeratins are sequestered into cytoplasmic inclusions in TRAIL-induced apoptosis*. J Cell Biol, 2000. **148**(6): p. 1239-54.
170. Yuan, R.T., et al., *Caspase-8 isoform 6 promotes death effector filament formation independent of microtubules*. Apoptosis, 2012. **17**(3): p. 229-35.
171. Barnhart, B.C., et al., *The death effector domain protein family*. Oncogene, 2003. **22**(53): p. 8634-44.
172. Kortemme, T., D.E. Kim, and D. Baker, *Computational alanine scanning of protein-protein interfaces*. Sci STKE, 2004. **2004**(219): p. pl2.
173. Kortemme, T. and D. Baker, *A simple physical model for binding energy hot spots in protein-protein complexes*. Proc Natl Acad Sci U S A, 2002. **99**(22): p. 14116-21.
174. Sandu, C., et al., *FADD self-association is required for stable interaction with an activated death receptor*. Cell Death Differ, 2006. **13**(12): p. 2052-61.
175. Morrisett, J.D., et al., *Interaction of an apolipoprotein (apoLP-alanine) with phosphatidylcholine*. Biochemistry, 1973. **12**(7): p. 1290-9.
176. Tourneur, L. and G. Chiochia, *FADD: a regulator of life and death*. Trends Immunol, 2010. **31**(7): p. 260-9.
177. Barbero, S., et al., *Identification of a critical tyrosine residue in caspase 8 that promotes cell migration*. J Biol Chem, 2008. **283**(19): p. 13031-4.
178. Garvey, T.L., et al., *Binding of FADD and caspase-8 to molluscum contagiosum virus MC159 v-FLIP is not sufficient for its antiapoptotic function*. J Virol, 2002. **76**(2): p. 697-706.
179. Neumann, L., et al., *Dynamics within the CD95 death-inducing signaling complex decide life and death of cells*. Mol Syst Biol, 2010. **6**: p. 352.
180. Lavrik, I.N., *Systems biology of death receptor networks: live and let die*. Cell Death Dis, 2014. **5**: p. e1259.
181. Muppidi, J.R., et al., *Homotypic FADD interactions through a conserved RXDLL motif are required for death receptor-induced apoptosis*. Cell Death Differ, 2006. **13**(10): p. 1641-50.

182. Zhou, P., et al., *Solution structure of Apaf-1 CARD and its interaction with caspase-9 CARD: a structural basis for specific adaptor/caspase interaction*. Proc Natl Acad Sci U S A, 1999. **96**(20): p. 11265-70.
183. Yuan, S., et al., *Structure of an apoptosome-procaspase-9 CARD complex*. Structure, 2010. **18**(5): p. 571-83.
184. Narayanan, K.B., T.H. Jang, and H.H. Park, *Self-oligomerization of ASC PYD domain prevents the assembly of inflammasome in vitro*. Appl Biochem Biotechnol, 2014. **172**(8): p. 3902-12.
185. Lu, A., et al., *Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes*. Cell, 2014. **156**(6): p. 1193-206.
186. Bae, J.Y. and H.H. Park, *Crystal structure of NALP3 protein pyrin domain (PYD) and its implications in inflammasome assembly*. J Biol Chem, 2011. **286**(45): p. 39528-36.
187. Yan, Q., et al., *Structural insight for the roles of fas death domain binding to FADD and oligomerization degree of the Fas-FADD complex in the death-inducing signaling complex formation: a computational study*. Proteins, 2013. **81**(3): p. 377-85.
188. Sun, H., et al., *A heterotrimeric death domain complex in Toll signaling*. Proc Natl Acad Sci U S A, 2002. **99**(20): p. 12871-6.
189. Qu, Q., et al., *Structural characterization of the self-association of the death domain of p75(NTR.)*. PLoS One, 2013. **8**(3): p. e57839.
190. Schiffmann, D.A., et al., *Formation and biochemical characterization of tube/pelle death domain complexes: critical regulators of postreceptor signaling by the Drosophila toll receptor*. Biochemistry, 1999. **38**(36): p. 11722-33.
191. Reichmann, D., et al., *The modular architecture of protein-protein binding interfaces*. Proc Natl Acad Sci U S A, 2005. **102**(1): p. 57-62.
192. Wu, H., *Higher-order assemblies in a new paradigm of signal transduction*. Cell, 2013. **153**(2): p. 287-92.
193. Acehan, D., et al., *Three-dimensional structure of the apoptosome: implications for assembly, procaspase-9 binding, and activation*. Mol Cell, 2002. **9**(2): p. 423-32.
194. Shiozaki, E.N., J. Chai, and Y. Shi, *Oligomerization and activation of caspase-9, induced by Apaf-1 CARD*. Proc Natl Acad Sci U S A, 2002. **99**(7): p. 4197-202.
195. Guicciardi, M.E. and G.J. Gores, *Life and death by death receptors*. FASEB J, 2009. **23**(6): p. 1625-37.
196. Lavrik, I.N. and P.H. Krammer, *Life and Death Decisions in the CD95 System: Main Pro-and Anti-Apoptotic Modulators*. Acta Naturae, 2009. **1**(1): p. 80-3.
197. Tsukumo, S.I. and S. Yonehara, *Requirement of cooperative functions of two repeated death effector domains in caspase-8 and in MC159 for induction and inhibition of apoptosis, respectively*. Genes Cells, 1999. **4**(9): p. 541-9.
198. Yao, Z., et al., *Death effector domain DEDa, a self-cleaved product of caspase-8/Mch5, translocates to the nucleus by binding to ERK1/2 and upregulates procaspase-8 expression via a p53-dependent mechanism*. EMBO J, 2007. **26**(4): p. 1068-80.
199. Soung, Y.H., et al., *CASPASE-8 gene is inactivated by somatic mutations in gastric carcinomas*. Cancer Res, 2005. **65**(3): p. 815-21.
200. Soung, Y.H., et al., *Caspase-8 gene is frequently inactivated by the frameshift somatic mutation 1225_1226delTG in hepatocellular carcinomas*. Oncogene, 2005. **24**(1): p. 141-7.

201. Kim, H.S., et al., *Inactivating mutations of caspase-8 gene in colorectal carcinomas*. *Gastroenterology*, 2003. **125**(3): p. 708-15.
202. Ando, M., et al., *Cancer-associated missense mutations of caspase-8 activate nuclear factor-kappaB signaling*. *Cancer Sci*, 2013. **104**(8): p. 1002-8.
203. Muller, F., T. Giroglou, and M. Sapp, *Characterization of the DNA-binding activity of the E1 and E2 proteins and the E1/E2 complex of human papillomavirus type 33*. *J Gen Virol*, 1997. **78 (Pt 4)**: p. 911-5.
204. Helfer, C.M., R. Wang, and J. You, *Analysis of the papillomavirus E2 and bromodomain protein Brd4 interaction using bimolecular fluorescence complementation*. *PLoS One*, 2013. **8**(10): p. e77994.
205. Abbate, E.A., C. Voitenleitner, and M.R. Botchan, *Structure of the papillomavirus DNA-tethering complex E2:Brd4 and a peptide that ablates HPV chromosomal association*. *Mol Cell*, 2006. **24**(6): p. 877-89.
206. Abbate, E.A., J.M. Berger, and M.R. Botchan, *The X-ray structure of the papillomavirus helicase in complex with its molecular matchmaker E2*. *Genes Dev*, 2004. **18**(16): p. 1981-96.
207. Wang, Y., et al., *Crystal structure of the E2 transactivation domain of human papillomavirus type 11 bound to a protein interaction inhibitor*. *J Biol Chem*, 2004. **279**(8): p. 6976-85.
208. Sanders, C.M., et al., *Transcription activator structure reveals redox control of a replication initiation reaction*. *Nucleic Acids Res*, 2007. **35**(10): p. 3504-15.
209. Hooley, E., et al., *The recognition of local DNA conformation by the human papillomavirus type 6 E2 protein*. *Nucleic Acids Res*, 2006. **34**(14): p. 3897-908.
210. Rozenberg, H., et al., *Structural code for DNA recognition revealed in crystal structures of papillomavirus E2-DNA targets*. *Proc Natl Acad Sci U S A*, 1998. **95**(26): p. 15194-9.
211. Hernandez-Ramon, E.E., et al., *Dimerization of the human papillomavirus type 16 E2 N terminus results in DNA looping within the upstream regulatory region*. *J Virol*, 2008. **82**(10): p. 4853-61.
212. Hegde, R., *The papillomavirus E2 proteins: structure, function, and biology*. *Annu. Rev. Biophys. Biomol. Struct.*, 2002. **31**: p. 343-360.
213. Harris, S.F. and M.R. Botchan, *Crystal Structure of the Human Papillomavirus Type 18 E2 Activation Domain*. *Science*, 1999. **284**(5420): p. 1673-1677.
214. Gagnon, D., et al., *Genetic analysis of the E2 transactivation domain dimerization interface from bovine papillomavirus type 1*. *Virology*, 2013. **439**(2): p. 132-9.
215. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures*. *J. Appl. Cryst.*, 1993. **26**: p. 283-291.
216. Wiederstein, M. and M.J. Sippl, *ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins*. *Nucleic Acids Res*, 2007. **35**(Web Server issue): p. W407-10.
217. Eisenberg, D., R. Luthy, and J.U. Bowie, *VERIFY3D: assessment of protein models with three-dimensional profiles*. *Methods Enzymol*, 1997. **277**: p. 396-404.
218. Colovos, C. and T.O. Yeates, *Verification of protein structures: patterns of nonbonded atomic interactions*. *Protein Sci*, 1993. **2**(9): p. 1511-9.
219. Hernandez-Ramon, E., et al., *Dimerization of the human papillomavirus type 16 E2 N terminus results in DNA looping within the upstream regulatory region*. *J Virol*, 2008. **82**: p. 4853-4861.

220. Riley, P.W., et al., *Dimer dissociation and unfolding mechanism of coagulation factor XI apple 4 domain: spectroscopic and mutational analysis*. J Mol Biol, 2007. **367**(2): p. 558-73.
221. Antson, A., et al., *Structure of the intact transactivation domain of the human papillomavirus E2 protein*. Nature, 2000. **403**: p. 805-809.
222. Gohlke, H. and D.A. Case, *Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf*. J Comput Chem, 2004. **25**(2): p. 238-50.
223. Basdevant, N., H. Weinstein, and M. Ceruso, *Thermodynamic basis for promiscuity and selectivity in protein-protein interactions: PDZ domains, a case study*. J Am Chem Soc, 2006. **128**(39): p. 12766-77.
224. Grunberg, R., M. Nilges, and J. Leckner, *Flexibility and conformational entropy in protein-protein binding*. Structure, 2006. **14**(4): p. 683-93.
225. Beard, H., et al., *Applying physics-based scoring to calculate free energies of binding for single amino acid mutations in protein-protein complexes*. PLoS One, 2013. **8**(12): p. e82849.
226. Clausen, T., et al., *HTRA proteases: regulated proteolysis in protein quality control*. Nat Rev Mol Cell Biol, 2011. **12**: p. 152-162.
227. Cilenti, L., et al., *Regulation of HAX-1 anti-apoptotic protein by Omi/HtrA2 protease during cell death*. J Biol Chem., 2004. **279**: p. 50295-50301.
228. Nallamsetty, S., et al., *Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro*. Protein Expr Purif., 2004. **38**(1): p. 108-115.
229. Sohn, J. and R. Sauer, *OMP peptides modulate the activity of DegS protease by differential binding to active and inactive conformations*. Mol Cell, 2009. **33**: p. 64-74.
230. Himoe, A., et al., *Investigations of the chymotrypsin-catalyzed hydrolysis of specific substrates. IV. Pre-steady state kinetic approaches to the investigation of the catalytic hydrolysis of esters*. J Biol Chem, 1969. **244**(13): p. 3483-93.
231. Bowers, K., et al. *Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters*. in *Proceedings of the ACM/IEEE Conference on Supercomputing (SC06)*. 2006. Florida.
232. Savopoulos, J.W., et al., *Expression, purification, and functional analysis of the human serine protease HtrA2*. Protein Expr Purif, 2000. **19**(2): p. 227-34.
233. Mazat, J. and J. Patte, *Lysine-sensitive aspartokinase of Escherichia coli K12. Synergy and autosynergy in an allosteric V system*. Biochemistry, 1976. **15**: p. 4053-4058.
234. Monod, J., J. Wyman, and J.P. Changeux, *On the Nature of Allosteric Transitions: A Plausible Model*. J Mol Biol, 1965. **12**: p. 88-118.
235. Sohn, J., R. Grant, and R. Sauer, *Allosteric activation of DegS, a stress sensor PDZ protease*. Cell, 2007. **131**: p. 572-583.
236. Feixas, F., et al., *Exploring the role of receptor flexibility in structure-based drug discovery*. Biophys Chem, 2014. **186**: p. 31-45.
237. Wist, A.D., et al., *Structure-activity based study of the Smac-binding pocket within the BIR3 domain of XIAP*. Bioorg Med Chem, 2007. **15**(8): p. 2935-43.
238. Sun, H., et al., *Design of small-molecule peptidic and nonpeptidic Smac mimetics*. Acc Chem Res, 2008. **41**(10): p. 1264-77.
239. Hedstrom, L., *Serine protease mechanism and specificity*. Chem Rev, 2002. **102**(12): p. 4501-24.

240. Hunkapiller, M.W., M.D. Forgac, and J.H. Richards, *Mechanism of action of serine proteases: tetrahedral intermediate and concerted proton transfer*. Biochemistry, 1976. **15**(25): p. 5581-8.
241. Gray, C., et al., *Characterization of human HtrA2, a novel serine protease involved in the mammalian cellular stress response*. Eur. J. Biochem., 2000. **267**(18): p. 5699-5710.
242. Skorko-Glonek, J., et al., *HtrA protease family as therapeutic targets*. Curr Pharm Des, 2013. **19**(6): p. 977-1009.
243. Huber, R. and W. Bode, *Structural basis of the activation and action of trypsin*. Accounts of Chemical Research, 1978. **11**(3): p. 114-122.
244. Sohn, J., R. Grant, and S. RT., *Allostery is an intrinsic property of the protease domain of DegS: implications for enzyme function and evolution*. J Biol Chem., 2010. **285**: p. 34039-34047.
245. Hansen, G. and R. Hilgenfeld, *Architecture and regulation of HtrA-family proteins involved in protein quality control and stress response*. Cell Mol Life Sci, 2013. **70**(5): p. 761-75.
246. Gandhi, P.S., et al., *Structural identification of the pathway of long-range communication in an allosteric enzyme*. Proc Natl Acad Sci U S A, 2008. **105**(6): p. 1832-7.
247. Gunasekaran, K., B. Ma, and R. Nussinov, *Is allostery an intrinsic property of all dynamic proteins?* Proteins, 2004. **57**(3): p. 433-43.
248. Krojer, T., et al., *Interplay of PDZ and protease domain of DegP ensures efficient elimination of misfolded proteins*. Proc Natl Acad Sci USA, 2008. **105**: p. 7702-7707.
249. Mohamedmohaideen, N., et al., *Structure and function of the virulence-associated high-temperature requirement A of Mycobacterium tuberculosis*. Biochemistry, 2008. **47**: p. 6092-6102.
250. Kley, J., et al., *Structural adaptation of the plant protease Deg1 to repair photosystem II during light exposure*. Nat Struct Mol Biol, 2011. **18**: p. 728-731.
251. Urwyler, S., *Allosteric modulation of family C G-protein-coupled receptors: from molecular insights to therapeutic perspectives*. Pharmacol Rev, 2011. **63**(1): p. 59-126.
252. May, L., et al., *Allosteric modulation of G protein-coupled receptors*. Annu Rev Pharmacol Toxicol. , 2007. **47**: p. 1-51.
253. Zhu, Z.H., et al., *Integrin alpha 7 interacts with high temperature requirement A2 (HtrA2) to induce prostate cancer cell death*. Am J Pathol, 2010. **177**(3): p. 1176-86.
254. Ariyaratana, S. and D.M. Loeb, *The role of the Wilms tumour gene (WT1) in normal and malignant haematopoiesis*. Expert Rev Mol Med, 2007. **9**(14): p. 1-17.
255. Hartkamp, J. and S.G. Roberts, *HtrA2, taming the oncogenic activities of WT1*. Cell Cycle, 2010. **9**(13): p. 2508-14.
256. Pozzi, N., et al., *Rigidification of the autolysis loop enhances Na(+) binding to thrombin*. Biophys Chem, 2011. **159**(1): p. 6-13.

Publications



Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Equilibrium dissociation and unfolding of human papillomavirus E2 transactivation domain

Nitu Singh, Shruthi Kanthaje¹, Kakoli Bose*

Integrated Biophysics and Structural Biology Lab, ACTREC, Tata Memorial Centre, Navi Mumbai 410210, India

ARTICLE INFO

Article history:

Received 23 April 2015

Accepted 14 May 2015

Available online 16 June 2015

Keywords:

HPV

E2

Dimer

Equilibrium

Thermodynamic

Unfolding

ABSTRACT

Papillomavirus E2 protein that performs essential functions such as viral oncogene expression and replication represents specific target for therapeutic intervention. DNA-binding activity is associated with its C-terminal DNA-binding domain (DBD), while the N-terminal transactivation domain (TAD) is responsible for replication and transactivation functions. Although both demonstrate large dependence on dimerization for mediating their functions, K_D for N-terminal dimerization is significantly high suggesting more dynamic role of this domain. However, unlike DBD, very little information is available on TAD dimerization, its folding and stability. Therefore, with an aim at delineating the regulatory switch of its dimerization, we have characterized high-risk HPV18 E2 TAD. Our studies demonstrate that E2 TAD is a weak but thermodynamically stable dimer ($K_D \sim 1.8 \mu\text{M}$, $\Delta G^{\text{H}_2\text{O}} = 18.8 \text{ kcal mol}^{-1}$) with $\alpha 2$ – $\alpha 3$ helices forming the interface. It follows a three-state folding pathway, in which unfolding involves dissociation of a dimeric intermediate. Interestingly, 90% of the conformational free energy is associated with dimer dissociation (16.9 of $18.8 \text{ kcal mol}^{-1}$) suggesting dimerization significantly contributes to its overall thermodynamic stability. These revelations might be important toward designing inhibitors for targeting dimerization or folding intermediates and hence multiple functions that E2 performs.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Human papillomaviruses (HPVs) are double-stranded DNA viruses that have been strongly implicated in development of lesions ranging from skin or genital warts to cancer [1]. Based upon oncogenicity, they are divided into two subclasses, high-risk (HR) types including HPV16 or 18 that are associated with 70% cases of cervical cancer, and benign wart-causing low risk (LR) types such as HPV6 and 11. Tremendous divergence in pathogenesis of HPVs has evolved due to complex system of regulation mediated by protein–DNA and protein–protein interactions between viral and host factors [2]. Particularly, the early protein E2 acts as a master regulator of the viral life-cycle. While considered to be the principal transcription factor, the E2 protein is equally important for viral DNA replication as well as mitotic partitioning of the viral genome [3]. In addition, several groups have demonstrated myriads of other functions such as NF κ B activation [4], induction of apoptosis [5] or regulation of host cell cycle [6]. The correlation of E2 with

mechanisms linked with cellular transformation hypothesizes its direct involvement in early steps of carcinogenesis, and thus exemplifies it as a potent target for therapeutic intervention.

E2 proteins characterized from numerous human and animal serotypes have three distinct modules; an amino-terminal transactivation domain (TAD) followed by a flexible, proline-rich hinge region and a carboxyl-terminal DNA-binding domain (DBD). TAD is a protein–protein interaction domain that binds to the viral helicase E1 [7] and to various host proteins such as topoisomerase-I [2], Brd4 [8] and procaspase-8 [5]. Crystal structures of TAD reveal that it comprises two sub-domains, N1 and N2, packed antiparallel to one another and separated by two consecutive single helical turns [7,9,10]. Although the main dimerization interface is located within DBD, which strongly dimerizes (nanomolar K_D) and tightly binds to viral DNA sites [11], available structures suggest that TAD also homodimerizes, however, with a higher K_D [9,12]. Dimerization of TAD has been proposed to serve as a molecular switch between early gene expression and viral genome replication during HPV infection [9,13]. Studies have also shown that induction of apoptosis by HPV16 E2 requires formation of functional TAD dimers, thereby further emphasizing the requisite for dimerization [14]. For HPV16 E2, the dimerization of TAD involves residues from helices $\alpha 2/\alpha 3$ of

* Corresponding author.

E-mail address: kbose@actrec.gov.in (K. Bose).¹ Present address: PGIMER, Chandigarh 160012, India.

N1 and 142–144 from N2 domains [9]. Nonetheless, with respect to HR-HPV18 E2, the possibility of TAD dimerization is still obscure. The available crystal structure has 65 N-terminal residues (two-thirds of N1 domain) missing and therefore lacks key amino acids reported for dimerization [15]. Furthermore, despite numerous reports on the biophysical properties of DBD, till date, no information is available on TAD folding and stability. A precise understanding of folding as well as determination of the overall conformational stability will help delineate the regulatory switch for E2 TAD dimerization which is crucial to understand its biological functions.

Here, using molecular modeling, mutagenesis and biophysicochemical approaches, we distinctly show that HR-HPV18 E2 TAD dimerizes at micromolar concentration. The dimerization is mediated by R41 and W42 as opposed to a single key residue R37 reported for HPV16 E2 [13]. In addition, we provide the first evidence on unfolding and conformational stability of E2 TAD.

2. Materials and methods

2.1. Molecular modeling, docking and simulation

Using Modeller-9v11 [16], the missing region of HPV18 E2 TAD (PDB: 1QQH) was homology modeled using HPV16 E2 TAD (PDB: 1DTP) as a template [9]. To obtain TAD dimer, the modeled monomeric HPV18 E2 TAD was superimposed on dimeric HPV16 E2 TAD, having C2 biological symmetry, using PyMol (Schrödinger, LLC). To verify correct assembly of dimeric interface, molecular docking of the monomer was performed with ClusPro-2.0 [17]. The final structure was energy minimized with PRIME Minimization module (Schrödinger, LLC) and further subjected to molecular dynamic simulation for 20 ns as described previously [18]. The simulated models were verified and validated using PROCHECK [19], ProSA-web [20], and ERRAT algorithms [21].

2.2. Sub-cloning, protein expression and purification

PEGFPc1-HPV18 E2 vector was a kind gift from Dr. F. Thierry (Institute of Medical Biology, Singapore). E2 TAD (residues 1–201) was sub-cloned in pACYC-vector (Clontech), and a C-terminal his₆-tag was introduced to facilitate protein purification. Mutants were generated using site-directed mutagenesis kit (Stratagene) and verified by DNA sequencing. *Escherichia coli* BL21(DE3) cells transformed with TAD constructs were grown at 37 °C till the O.D₆₀₀ reached 0.5, further induced with 2 µg/ml of arabinose and cultured at 18 °C for 20 h. The Ni²⁺-affinity purification was performed in buffer-A (10 mM Na₂HPO₄/NaH₂PO₄, pH 7.5, 100 mM NaCl, 20 mM β-mercaptoethanol) as described previously [22]. The proteins were further purified and characterized by size exclusion chromatography using Superdex200-10/30 column (GE Healthcare). Protein identity was confirmed with Western blotting using anti-E2 antibody (sc-26939).

2.3. Glutaraldehyde cross-linking

20 µg of purified protein was treated with freshly prepared glutaraldehyde solution (final concentration 0.1%) for 2 min at 37 °C. The reaction was terminated by adding 5 µl of 2 M Tris–HCl, pH 8.0. Cross-linked product was mixed with Laemmli sample buffer and analyzed on 12% SDS-PAGE.

2.4. Dynamic light scattering (DLS)

1 mg/ml protein was loaded into a 45 µl quartz cuvette. Measurements were performed at 25 °C and at least 10 scans each of

12 s duration were collected (DynaPro NanoStar, Wyatt Technology). Histogram analyses were carried out using DYNAMICS v.6.0 software.

2.5. Fluorescence and circular dichroism (CD) spectroscopy

Protein sample preparations for equilibrium unfolding studies were done as described previously [23]. Fluorescence measurements were recorded on Fluorolog-3 spectrofluorometer (Horiba Scientific) with 280 and 295 nm excitation followed by emission between 310 and 400 nm. Far UV CD spectra were acquired using JASCO J-815 spectropolarimeter between 250 and 200 nm, at 20 nm/min scan speed.

2.6. Analysis of the equilibrium unfolding profile

E2 TAD wild-type and single mutants followed a three-state folding pathway with the formation of a dimeric intermediate, $N_2 \rightleftharpoons I_2 \rightleftharpoons 2U$. The data was normalized and fitted to the equations described earlier [23], for obtaining the thermodynamic parameters using Igor Pro 6.03A (WaveMetrics). The data for E2 TAD double mutant was fitted to a two-state monomeric model, $N \rightleftharpoons U$ [23].

3. Results

3.1. In silico analysis of E2 TAD dimerization

To assess homodimerization of HPV18 E2 TAD, a structural model was generated and its overall quality factor was assessed by ERRAT, which was found to be 82% implying good accuracy of the predicted structure. The stereochemical analysis revealed 100% of residues were in the most favorable and additional allowed regions of the Ramachandran plot (Fig. S1A). ProSA-web evaluation provided a compatible Z score value of –7.9 (Fig. S1B), which is well within the range of native conformations of crystal structures. The overall residue energies of the 3D model were largely found to be negative (Fig. S1C). These observations indicated that the model quality was reliable and could be considered for further analysis.

Superposition of HPV16 and 18 E2 TAD dimers showed an RMSD of less than 0.8 Å, suggesting similar conformation for the complexes (Fig. 1A). HPV18 E2 TAD dimer had 1806 Å² of buried surface area, comparable to the 2026 Å² reported for HPV16 E2 [9]. At the dimer interface, each subunit primarily contributed residues R41, W42, F48, I77, M81, Q84, Q88 from helices α2 and α3 (Table S1). Notably, residues R41, W42 (helix α2) and Q84, Q88 (helix α3) seem to play central roles in stabilizing the dimer interface (Fig. 1B). Dimerization studies on HPV16 E2 TAD demonstrated that substitution of R37 (equivalent to R41 of HPV18) abolished dimer formation [2]. The residue R41 is preserved while W42 is conserved in terms of its hydrophobicity across different papillomavirus types (Fig. S2). Thus, based on these observations, R41 and W42 of HPV18 E2 TAD were targeted for mutagenesis and further characterization.

3.2. In vitro characterization of wild-type and mutant E2 TAD

To determine whether HPV18 E2 TAD forms dimer in solution, we tested the wild-type and α2 helix mutants (R41A, W42A, and R41A/W42A) for their ability to homodimerize *in vitro*. As the first step, the purity of the proteins was checked on SDS-PAGE and the identity was confirmed with Western blotting using anti-E2 antibody (Fig. S3). The gel filtration profile for E2 TAD variants is presented in Fig. 1C. It shows that the wild-type and mutant proteins exist predominantly as monomers, however, except for the double mutant, a peak corresponding to the dimer is also observed (at a

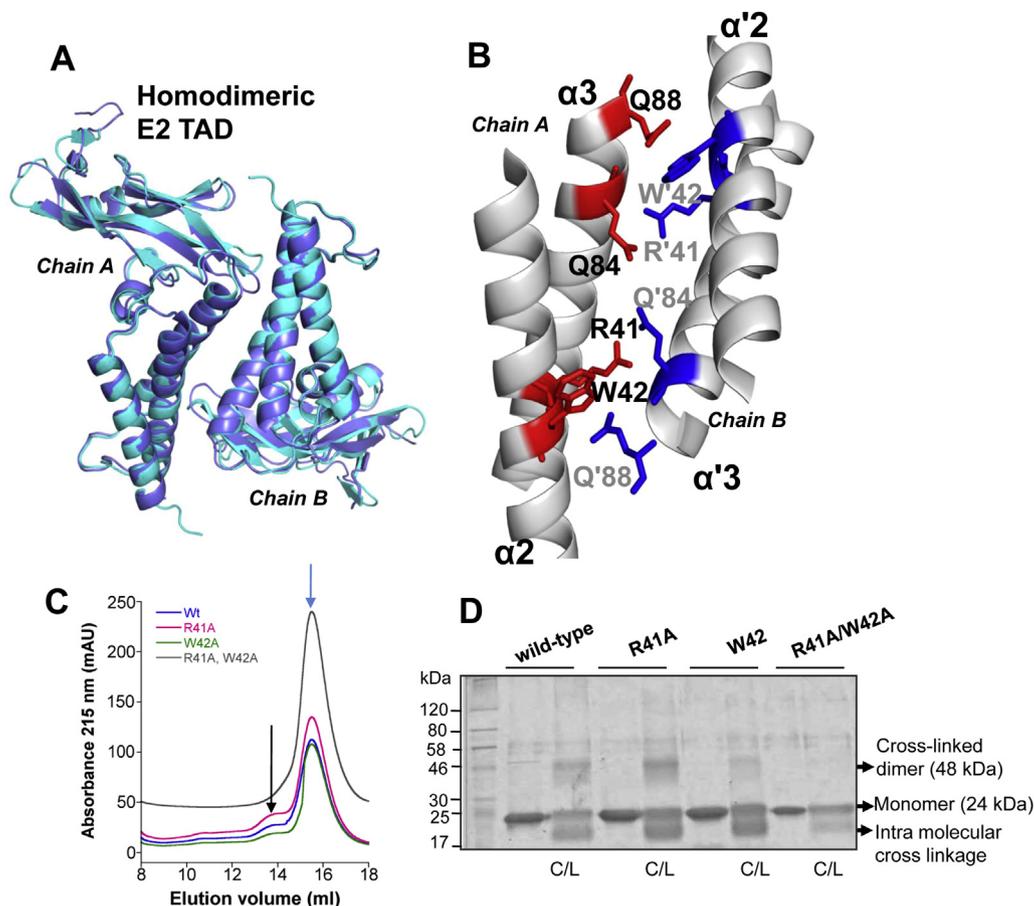


Fig. 1. *In silico* and *in vitro* characterization of HPV18 E2 TAD. A) Superposition of HPV18 and 16 E2 TAD shown in cyan and blue respectively. B) A close-up view of the interface with selected side chains from chain A and B in red and blue respectively. C) Size exclusion chromatography for E2 TAD variants at protein concentrations $<10 \mu\text{M}$. A major peak (blue arrow) corresponding to the calculated molecular weight of a monomer (24.5 kDa) and a minor peak (black arrow), except for the double mutant, related to the dimeric species (49 kDa) are observed. D) Glutaraldehyde cross-linking reaction product resolved on SDS-PAGE. A band corresponding to size lower than 24 kDa in the cross-linked lane is due to intra-molecular protein cross-linking resulting in increased mobility. C/L represents cross-linked.

final eluant concentration of $\sim 1 \mu\text{M}$) suggesting TAD forms a weak dimer. To confirm the presence of the dimeric species, $20 \mu\text{g}$ of wild-type and mutant proteins were cross-linked with glutaraldehyde and run on SDS-PAGE. The wild-type and single mutants show mixture of monomeric and dimeric species, however, the double mutant was found to exist only in the monomeric form (Fig. 1D). These results demonstrate that at low protein concentrations, TAD wild-type and single mutants are present in monomer–dimer equilibrium.

To substantiate these findings, we determined the heterogeneity of the proteins at higher concentrations (1 mg/ml) using DLS. We observed that the average hydrodynamic radii of TAD wild-type and single mutants were larger compared to the double mutant (Fig. S4). Based upon the radii, the calculated molecular weight of the proteins is given in Table 1. For wild-type and single mutants, the calculated molecular weight closely resembled the theoretical molecular weight of a dimeric E2 TAD, whereas the double mutant showed a homogenous monomeric population.

3.3. Urea-dependent unfolding of E2 TAD

Equilibrium unfolding studies provide an excellent tool to delineate protein stability and determine whether dimerization is a folding event [24]. We therefore studied unfolding of E2 TAD as a function of urea concentration using different spectroscopic probes. In case of wild-type TAD, the data show a first step of unfolding

which occurs between 0 and $\sim 2.5 \text{ M}$ urea, followed by a second transition between ~ 3 and 8 M urea. All the three spectroscopic signals showed similar trends in the unfolding (Fig. 2A). To examine reversibility, the protein unfolded in buffer containing 8 M urea was diluted to final urea concentrations shown in Fig. 2A. The data suggest that the folding transitions are completely reversible. Similar experiments were performed with TAD single and double mutant proteins. For both R41A and W42A single mutants, a clear biphasic denaturation curve with reversible folding was observed (Fig. 2B and C). Interestingly, in case of TAD double mutant (R41A/W42A), the data showed only a single reversible transition. A pre-transition occurs between 0 and $\sim 1 \text{ M}$ urea which is followed by a cooperative decrease in signal with a post transition baseline between ~ 5 and 8 M urea (Fig. 2D).

Overall, these results demonstrate that E2 TAD wild-type and single mutants, follow a three-state pathway in which a folding intermediate is in equilibrium with the native and unfolded states. However, strikingly, the double mutant (R41A/W42A) seems to undergo a two-state folding process.

3.4. Protein concentration dependence of unfolding

The mechanism by which oligomeric proteins unfold can be delineated by monitoring the profile of each equilibrium denaturation and its protein concentration dependence. Since E2 TAD homodimerizes, we determined the concentration dependence of

Table 1
Dynamic light scattering analysis of E2 TAD variants.

E2 TAD	Hydrodynamic radius ^a	% Polydispersity ^b	Calculated mean residual molecular weight (kDa)	Theoretical molecular weight (expsy protparam) (kDa)	
				Monomer	Dimer
Wild-type	2.98	34.9	55.4	24.2	48.4
R41A	2.93	31.3	52.7	24.1	48.2
W42A	2.90	18.4	51.1	24.1	48.2
R41A, W42A	2.42	4.5	26.2	24.0	48.0

^a Mean hydrodynamic radius derived from the measured translational diffusion coefficient using the Stokes–Einstein equation.

^b Polydispersity divided by the hydrodynamic radius.

the unfolding process. As shown in Fig. 3A–C, there is a shift in the midpoint of the second transition with the increase in protein concentration. However, the first transition is concentration independent and the $U_{1/2}$ value (concentration of urea at which 50% of the protein is unfolded) remains constant. This pattern is consistent with a scheme where a native dimer (N_2) partially unfolds to a dimeric intermediate (I_2), which then unfolds to denatured monomer ($2U$), $N_2 \rightleftharpoons I_2 \rightleftharpoons 2U$. Thus, fitting of the data to a three-state folding model with a dimeric intermediate was carried out

to obtain the conformational free energies (ΔG^{H_2O}) and the cooperativity indices, or m -values (Table 2). The data show that the initial unfolding step ($N_2 \rightleftharpoons I_2$) has a $\Delta G_1^{H_2O}$ of <3 kcal mol⁻¹. Thus, the majority of E2 TAD stability resides in its dimerization ($I_2 \rightleftharpoons 2U$), which requires an additional ~ 17 kcal mol⁻¹ to unfold.

By using ΔG^{H_2O} and m -values, we calculated equilibrium distribution of the species for all the protein concentrations. As shown in Fig. 4A, between 0 and 3.5 M urea, there is decrease in the population of native dimer with concomitant increase in the population of the dimeric intermediate. The mid-point of the transition is 1.8 M urea that is consistent with the experimental results shown in Fig. 3A. The relative concentration of the dimeric species is dependent on the protein concentration and reaches a maximum at ~ 3.1 M urea. The second transition, I_2 to U , has midpoint values of 4.3, 4.7 and 4.9 M for the concentration of the proteins indicated in the figure legend. By observing this second-order transition, both as a function of denaturant and protein concentration, we calculated an apparent K_D (K_D^{app}) for the dissociation using equation, $\Delta G_2^{H_2O} = -RT \ln K_D^{app}$. The calculated K_D^{app} at 4.5 M urea was found to be ~ 1.8 μ M which is in close agreement with the dissociation constant K_D , $8.1 \pm 4.2 \times 10^{-6}$ M reported for HPV16 E2 TAD dimer [9]. Similar pattern for distribution of three species (native, intermediate, and unfolded) is observed for dimeric R41A and W42A single mutants as shown in Fig. 4B and C.

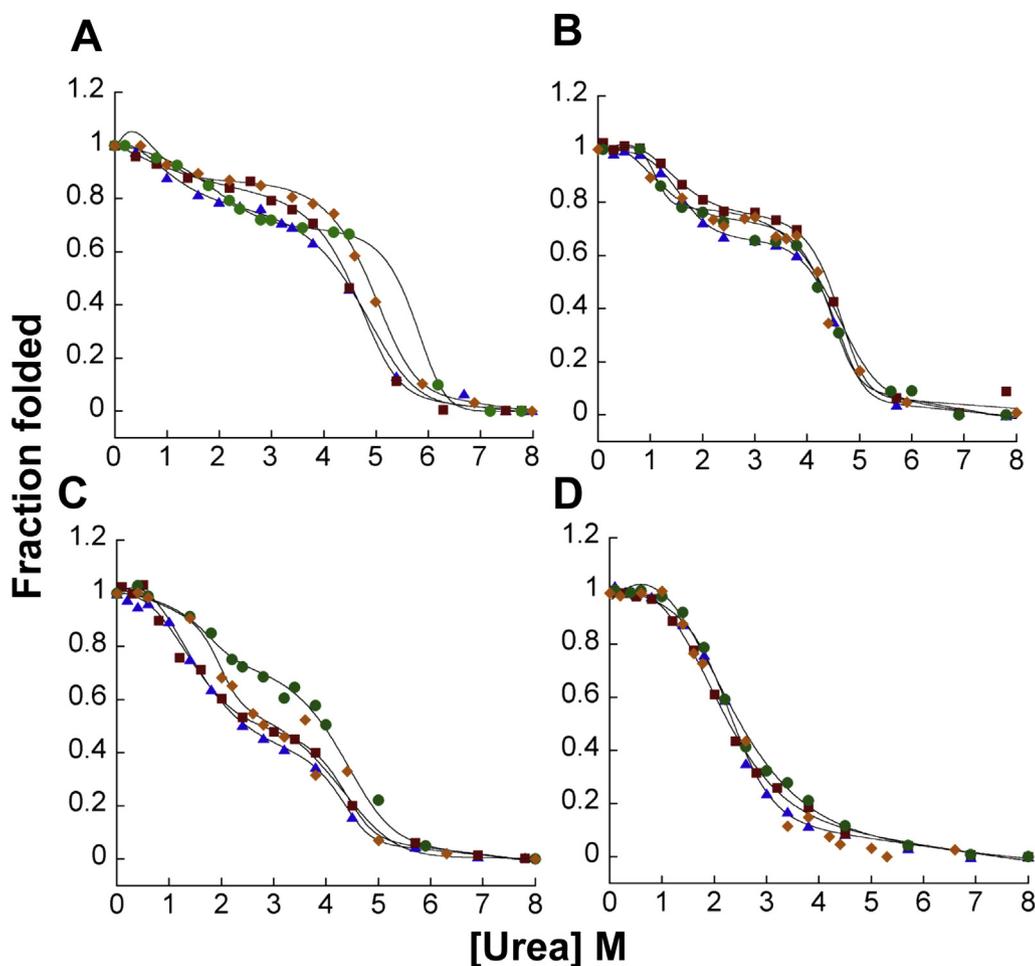


Fig. 2. Equilibrium unfolding of E2 TAD variants. Unfolding measured by fluorescence emission with excitation at 280 (▲), or 295 nm (◆), CD at 228 nm (●) and refolding measured by fluorescence emission with excitation at 280 nm (■). For fluorescence and CD measurements, the protein concentrations were 2.5 μ M and 5 μ M respectively. For clarity, error bars and refolding data are not shown for all the three data sets. The solid lines represent fit using a three-state (A–C) or two-state folding model (D).

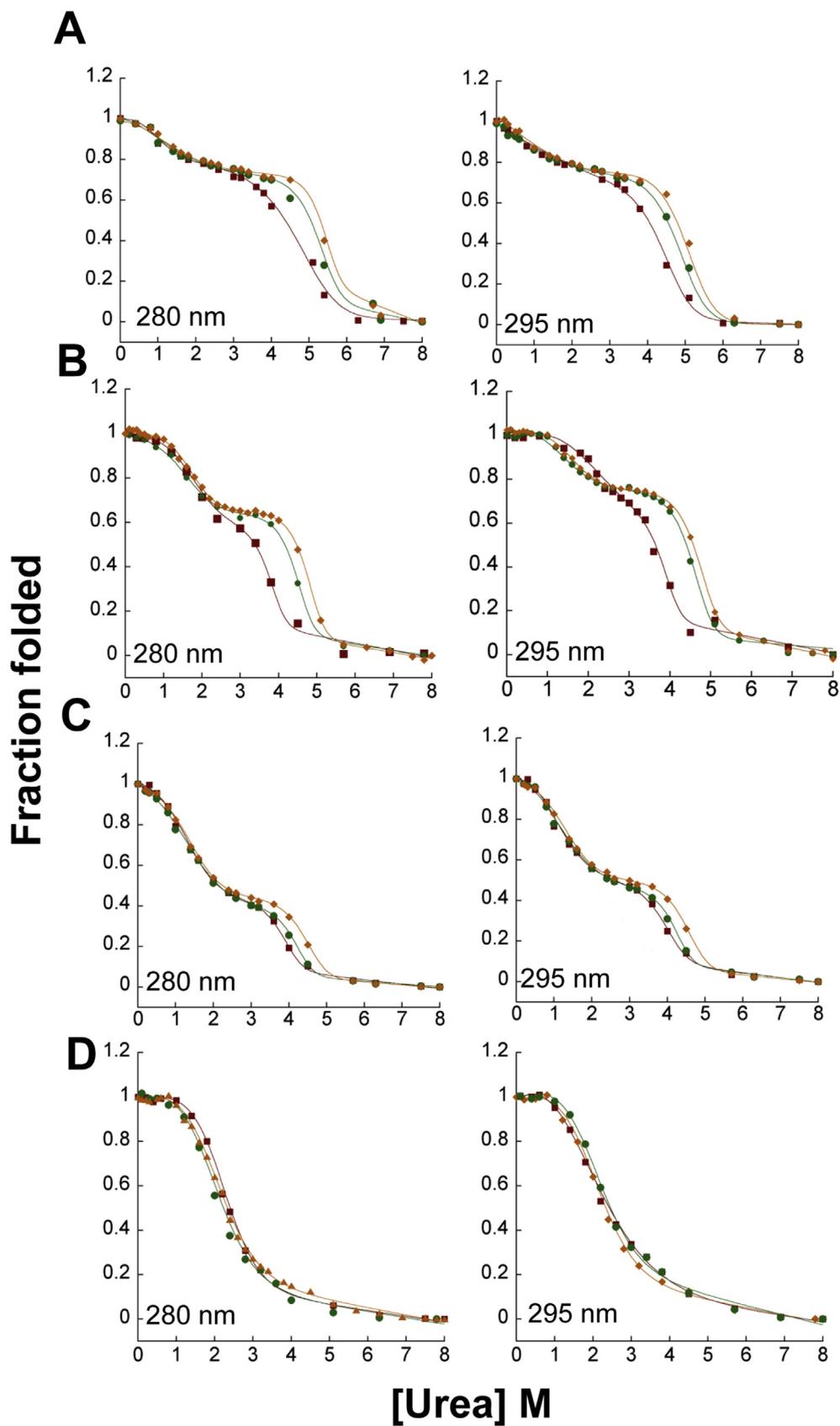


Fig. 3. Protein concentration dependence of unfolding. Equilibrium unfolding monitored by fluorescence emission with excitation at 280 and 295 nm for TAD wild-type (A), R41A (B), W42A (C), and R41A/W42A (D) proteins. The protein concentrations are 1.5, 2.5, and 3.5 μM , each represented by symbols \blacksquare , \bullet , \blacklozenge respectively. The solid lines represent global fits of the data using Igor Pro. For clarity, error bars are not shown for all the data sets.

Table 2
Thermodynamic parameters for E2 TAD wild-type and single mutants.

E2 TAD	$m_1 (N_2 \rightleftharpoons I_2)$ (kcal mol ⁻¹ M ⁻¹)	$m_2 (I_2 \rightleftharpoons 2U)$ (kcal mol ⁻¹ M ⁻¹)	$\Delta G_1^{\text{H}_2\text{O}} (N_2 \rightleftharpoons I_2)$ (kcal mol ⁻¹)	$\Delta G_2^{\text{H}_2\text{O}} (I_2 \rightleftharpoons 2U)$ (kcal mol ⁻¹)
WT	0.98 ± 0.1	1.97 ± 0.06	1.98 ± 0.1	16.9 ± 0.1
R41A	1.81 ± 0.1	2.39 ± 0.1	2.62 ± 0.6	18.2 ± 0.7
W42A	1.32 ± 0.1	2.18 ± 0.2	2.11 ± 0.2	17.3 ± 0.9

Each entry corresponds to the mean ± S.E. (standard error) of three independent experiments at three different concentrations of the protein: 1.5, 2.5, and 3.5 μM.

We also performed equilibrium denaturation experiments with several protein concentrations of R41A/W42A double mutant. Interestingly, a single sharp transition with no protein concentration dependence is observed between native and denatured states demonstrating that the protein exists as a monomer (Fig. 3D). The total free energy was therefore calculated by fitting the data to a two-state equilibrium model for monomeric proteins. The $U_{1/2}$ and m -value of the transition are 2.5 ± 0.1 M urea and 1.4 ± 0.08 kcal mol⁻¹ M⁻¹, respectively. The $\Delta G^{\text{H}_2\text{O}} (N \rightleftharpoons U)$ of this monomeric variant is thus barely 3.3 ± 0.3 kcal mol⁻¹ ($\Delta G^{\text{H}_2\text{O}} = m [U_{1/2}]$). This stability is similar to the initial unfolding of the dimeric E2 TAD (Table 2). The fraction of species at different urea concentrations is shown in Fig. 4D. Overall, these results suggest that the stability of the dimer contributes significantly to the conformational free energy of E2 TAD protein.

4. Discussion

The results presented here highlight four key properties of HPV18 E2 TAD. Firstly, it is a weak dimer ($K_D \sim 1.8$ μM) that readily isomerizes with a lesser cooperativity compared to the second

transition, to a partially unfolded dimeric intermediate. Second, presence of a dimeric intermediate suggests dimerization is a folding event. Third, dimerization contributes significantly towards stability of the protein (>89%). Lastly, monomeric TAD variant is substantially destabilized compared to the dimeric protein and unfolds via a two-state pathway without any significantly populated intermediate at the given concentrations.

The biphasic nature of HPV18 E2 TAD denaturation curve and protein concentration-dependent second transition is consistent with the existence of a dimeric intermediate. The overall free energy change involved in the transition from native dimer to the intermediate corresponds to only ~10% of the total stabilization energy (Table 2). However, the second part of the transition is characterized by 8–9 times larger values of $\Delta G_2^{\text{H}_2\text{O}}$. Moreover, based on the m -values ($m_1 \ll m_2$), the first transition ($N_2 \rightleftharpoons I_2$) shows lack of co-operative unfolding highlighting that the structure is particularly flexible and can easily adopt partially unfolded conformations. Several regions of E2 TAD modular structure, in particular, unstructured key joints and exposed loops at the surface might provide such flexibility [9,15,25]. Preferential interactions of the denaturant with these external regions can be the principal reason

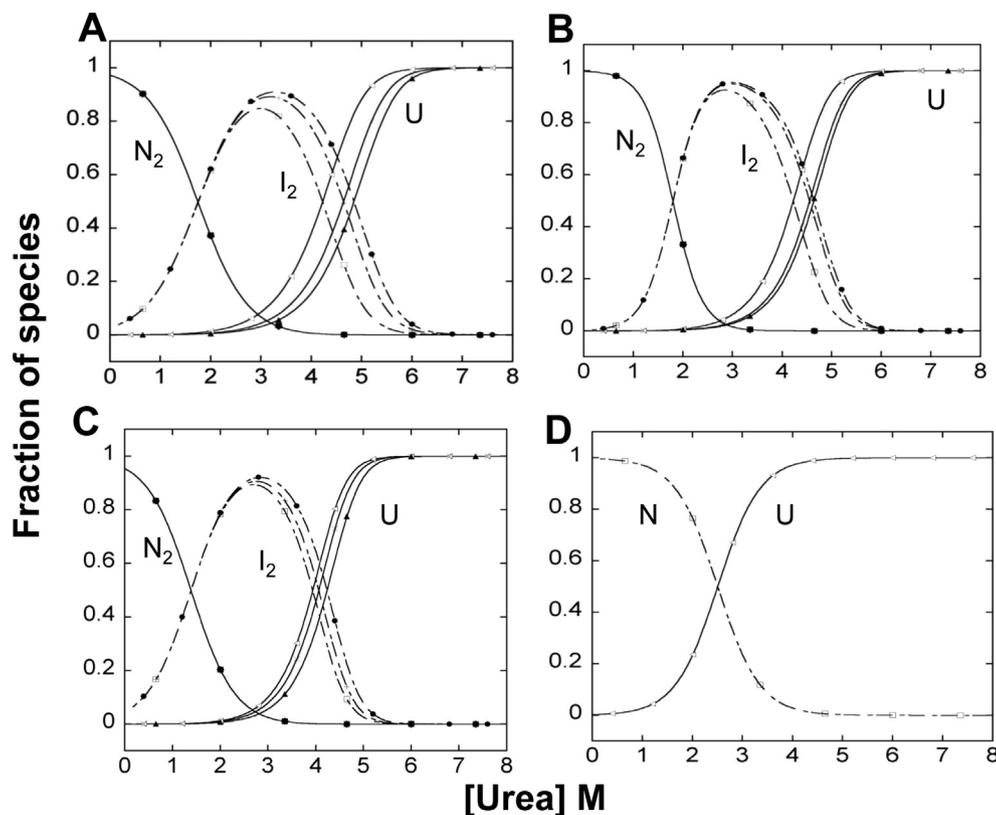


Fig. 4. Fraction of species in the unfolding process. The fraction of native, intermediate and unfolded protein calculated as a function of urea concentration for wild-type (A), R41A (B), W42A (C), and R41A/W42A (D). The protein concentrations are 1.5, 2.5 and 3.5 μM, and N_2 and N refers to the native protein, I_2 is the dimeric intermediate and U refers to the unfolded species.

for local unfolding of the E2 TAD structure far from the dimeric interface.

Denaturant m -values can be related to the global changes in accessible surface area (Δ ASA) of the protein exposed to solvent during unfolding [26]. Scholtz and co-workers [27] have described empirical relationships that correlate Δ ASA with experimental m -values, given as, $y = 368 + 0.11(\Delta$ ASA). Using this equation and the m -values for E2 TAD wild-type (Table 2), the calculated Δ ASAs for the first and second transitions are 5500 Å² and 14,500 Å², respectively. This is in agreement with our finding that subunit dissociation and conversion of the intermediate to unfolded E2 TAD monomers occur in second transition, where maximum surface exposure would be expected.

Our data show that HPV18 E2 TAD is reasonably stable with a total free energy change of 18.8 kcal mol⁻¹ that is comparable with other dimeric proteins of similar size [28]. Studies on the stability of dimeric and monomeric proteins highlight that the former are, on an average, 4–15 kcal mol⁻¹ more stable than the later [24]. Increased stability in case of dimers seems to occur due to burial of additional surfaces upon subunit association [29]. Consistent to this, the monomeric E2 TAD shows a dramatic decrease in the conformational stability. These results demonstrate that the interfacial residues are critically involved in structural stability. Furthermore, both the residues R41 and W42 seem to be crucial for homodimerization, unlike HPV16 E2, where R37A mutation alone resulted in dimer destabilization. Based upon HPV16 E2 TAD dimer structure (PDB ID: 1D7O) and our model for HPV18 E2 TAD, we understand that the possible difference could be due to the type of interactions at the interface (Tables S1 and S2). In HPV16 E2, R37 makes extensive hydrogen bonding and ionic interactions with E80 to stabilize the symmetrical dimer interface. However, in HPV18 E2, E80 is replaced with Q84 which although makes hydrogen contact with R41, does not participate in forming a stabilizing salt-bridge. The additional hydrogen bonds between side chains of W42 and Q88 might be required to provide stability to the dimer interface. Thus, although E2 TAD shares homologous native conformation, the residues and interactions at dimer interface might not be conserved. In support of this observation, recently, studies on BPV1 E2 TAD demonstrated that the dimerization is mediated through redox interactions involving different residues [12,30]. Therefore, while the native structures of E2 are highly conserved, the folding pathways and the intermediates may differ.

In HPV E2 proteins, the extended (~80 amino acids) linker between DBD and TAD provides flexibility to both the domains so as to enable them to function independently [31]. Based on our results, we hypothesize that at lower concentrations, the monomeric TAD species will be populated, which would dimerize only beyond a threshold micromolar concentration. Although E2 is present in a dimeric form even at lower concentrations (mediated by DBD), the more dynamic N-terminal TAD might exist in a monomer–dimer equilibrium. Post viral replication, when the concentration of this protein considerably increases within the host cell, TAD dimerizes through its critical residues which might be a prerequisite for transactivating E6 and E7 oncogenes, and hence viral propagation [32]. Therefore, identifying this regulatory switch might be important towards devising specific targets for regulating TAD dimerization and pathogenesis.

Acknowledgment

We thank Mr. Ali Hassan for help in computational modeling, Dr. G. Krishnamoorthy (TIFR, India) for providing the DLS facility and ACTREC for funding the project.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.bbrc.2015.05.057>.

Transparency document

Transparency document related to this article can be found online at <http://dx.doi.org/10.1016/j.bbrc.2015.05.057>.

References

- [1] S.V. Graham, Human papillomavirus: gene expression, regulation and prospects for novel diagnostic methods and antiviral therapies, *Future Microbiol.* 5 (2010) 1493–1506.
- [2] M. Muller, C. Demeret, The HPV E2-host protein-protein interactions: a complex hijacking of the cellular network, *Open Virol. J.* 6 (2012) 173–189.
- [3] A.A. McBride, The papillomavirus E2 proteins, *Virology* 445 (2013) 57–79.
- [4] M. Boulabiar, S. Boubaker, M. Favre, C. Demeret, Keratinocyte sensitization to tumour necrosis factor-induced nuclear factor kappa B activation by the E2 regulatory protein of human papillomaviruses, *J. Gen. Virol.* 92 (2011) 2422–2427.
- [5] S. Blachon, C. Demeret, The regulatory E2 proteins of human genital papillomaviruses are pro-apoptotic, *Biochimie* 85 (2003) 813–819.
- [6] S. Bellanger, S. Blachon, F. Mechali, C. Bonne-Andrea, F. Thierry, High-risk but not low-risk HPV E2 proteins bind to the APC activators Cdh1 and Cdc20 and cause genomic instability, *Cell Cycle* 4 (2005) 1608–1615.
- [7] E.A. Abbate, J.M. Berger, M.R. Botchan, The X-ray structure of the papillomavirus helicase in complex with its molecular matchmaker E2, *Genes Dev.* 18 (2004) 1981–1996.
- [8] C.M. Helfer, R. Wang, J. You, Analysis of the papillomavirus E2 and bromodomain protein Brd4 interaction using bimolecular fluorescence complementation, *PLoS One* 8 (2013) e77994.
- [9] A.A. Antson, J.E. Burns, O.V. Moroz, D.J. Scott, C.M. Sanders, I.B. Bronstein, G.G. Dodson, K.S. Wilson, N.J. Maitland, Structure of the intact transactivation domain of the human papillomavirus E2 protein, *Nature* 403 (2000) 805–809.
- [10] E.A. Abbate, C. Voitenleitner, M.R. Botchan, Structure of the papillomavirus DNA-tethering complex E2:Brd4 and a peptide that ablates HPV chromosomal association, *Mol. Cell* 24 (2006) 877–889.
- [11] H. Rozenberg, D. Rabinovich, F. Frolow, R.S. Hegde, Z. Shakked, Structural code for DNA recognition revealed in crystal structures of papillomavirus E2-DNA targets, *Proc. Natl. Acad. Sci. U. S. A.* 95 (1998) 15194–15199.
- [12] D. Gagnon, H. Senechal, C.M. D'Abramo, J. Alvarez, A.A. McBride, J. Archambault, Genetic analysis of the E2 transactivation domain dimerization interface from bovine papillomavirus type 1, *Virology* 439 (2013) 132–139.
- [13] E.E. Hernandez-Ramon, J.E. Burns, W. Zhang, H.F. Walker, S. Allen, A.A. Antson, N.J. Maitland, Dimerization of the human papillomavirus type 16 E2 N terminus results in DNA looping within the upstream regulatory region, *J. Virol.* 82 (2008) 4853–4861.
- [14] K. Webster, J. Parish, M. Pandya, P.L. Stern, A.R. Clarke, K. Gaston, The human papillomavirus (HPV) 16 E2 protein induces apoptosis in the absence of other HPV proteins and via a p53-dependent pathway, *J. Biol. Chem.* 275 (2000) 87–94.
- [15] S.F. Harris, M.R. Botchan, Crystal structure of the human papillomavirus type 18 E2 activation domain, *Science* 284 (1999) 1673–1677.
- [16] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234 (1993) 779–815.
- [17] S.R. Comeau, D.W. Gatchell, S. Vajda, C.J. Camacho, ClusPro: a fully automated algorithm for protein-protein docking, *Nucleic Acids Res.* 32 (2004) W96–W99.
- [18] N. Singh, A. D'Souza, A. Cholleti, G.M. Sastry, K. Bose, Dual regulatory switch confers tighter control on HtrA2 proteolytic activity, *FEBS J.* 281 (2014) 2456–2470.
- [19] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Cryst.* 26 (1993) 283–291.
- [20] M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, *Nucleic Acids Res.* 35 (2007) W407–W410.
- [21] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, *Protein Sci.* 2 (1993) 1511–1519.
- [22] J.E. Burns, O.V. Moroz, A.A. Antson, C.M. Sanders, K.S. Wilson, N.J. Maitland, Expression, crystallization and preliminary X-ray analysis of the E2 transactivation domain from papillomavirus type 16, *Acta Crystallogr. D. Biol. Crystallogr.* 54 (1998) 1471–1474.
- [23] J. Walters, S.L. Milam, A.C. Clark, Practical approaches to protein folding and assembly: spectroscopic strategies in thermodynamics and kinetics, *Methods Enzymol.* 455 (2009) 1–39.
- [24] K.E. Neet, D.E. Timm, Conformational stability of dimeric proteins: quantitative studies by equilibrium denaturation, *Protein Sci.* 3 (1994) 2167–2174.

- [25] Y. Wang, R. Coulombe, D. Cameron, L. Thauvette, M. Massariol, L. Amon, D. Fink, S. Titolo, E. Welchner, C. Yoakim, J. Archambault, P. White, Crystal structure of the E2 transactivation domain of human papillomavirus type 11 bound to a protein interaction inhibitor, *J. Biol. Chem.* 279 (2004) 6976–6985.
- [26] D.O. Alonso, K.A. Dill, Solvent denaturation and stabilization of globular proteins, *Biochemistry* 30 (1991) 5974–5985.
- [27] J.K. Myers, C.N. Pace, J.M. Scholtz, Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding, *Protein Sci.* 4 (1995) 2138–2148.
- [28] J.A. Rumpfheldt, C. Galvagnion, K.A. Vassall, E.M. Meiering, Conformational stability and folding mechanisms of dimeric proteins, *Prog. Biophys. Mol. Biol.* 98 (2008) 61–84.
- [29] S. Jones, J.M. Thornton, Protein-protein interactions: a review of protein dimer structures, *Prog. Biophys. Mol. Biol.* 63 (1995) 31–65.
- [30] C.M. Sanders, D. Sizov, P.R. Seavers, M. Ortiz-Lombardia, A.A. Antson, Transcription activator structure reveals redox control of a replication initiation reaction, *Nucleic Acids Res.* 35 (2007) 3504–3515.
- [31] S. Bellanger, C.L. Tan, Y.Z. Xue, S. Teissier, F. Thierry, Tumor suppressor or oncogene? A critical role of the human papillomavirus (HPV) E2 protein in cervical cancer progression, *Am. J. Cancer Res.* 1 (2011) 373–389.
- [32] J. Doorbar, Molecular biology of human papillomavirus infection and cervical cancer, *Clin. Sci. (Lond)* 110 (2006) 525–541.

Dual regulatory switch confers tighter control on HtrA2 proteolytic activity

Nitu Singh¹, Areetha D'Souza^{1,*}, Anuradha Cholleti², G. Madhavi Sastry² and Kakoli Bose¹

¹ Advanced Centre for Treatment, Research and Education in Cancer (ACTREC), Tata Memorial Centre, Kharghar, Navi Mumbai, India

² Schrödinger, Sanali Infopark, Banjara Hills, Hyderabad, India

Keywords

allostery; IAP; N-terminus; PDZ; serine protease

Correspondence

K. Bose, Integrated Biophysics and Structural Biology Laboratory, ACTREC, Tata Memorial Centre, Sector 22, Kharghar, Navi Mumbai 410210, India
Fax: +91 22 2740 5085
Tel: +91 22 2740 5109
E-mail: kbose@actrec.gov.in

*Present address

School of Biological Sciences, Nanyang Technological University, Singapore

(Received 22 February 2014, revised 25 March 2014, accepted 28 March 2014)

doi:10.1111/febs.12799

High-temperature requirement protease A2 (HtrA2), a multitasking serine protease that is involved in critical biological functions and pathogenicity, such as apoptosis and cancer, is a potent therapeutic target. It is established that the C-terminal post-synaptic density protein, *Drosophila* disc large tumor suppressor, zonula occludens-1 protein (PDZ) domain of HtrA2 plays pivotal role in allosteric modulation, substrate binding and activation, as commonly reported in other members of this family. Interestingly, HtrA2 exhibits an additional level of functional modulation through its unique N-terminus, as is evident from 'inhibitor of apoptosis proteins' binding and cleavage. This phenomenon emphasizes multiple activation mechanisms, which so far remain elusive. Using conformational dynamics, binding kinetics and enzymology studies, we addressed this complex behavior with respect to defining its global mode of regulation and activity. Our findings distinctly demonstrate a novel N-terminal ligand-mediated triggering of an allosteric switch essential for transforming HtrA2 to a proteolytically competent state in a PDZ-independent yet synergistic activation process. Dynamic analyses suggested that it occurs through a series of coordinated structural reorganizations at distal regulatory loops (L3, LD, L1), leading to a population shift towards the relaxed conformer. This precise synergistic coordination among different domains might be physiologically relevant to enable tighter control upon HtrA2 activation for fostering its diverse cellular functions. Understanding this complex rheostatic dual switch mechanism offers an opportunity for targeting various disease conditions with tailored site-specific effector molecules.

Structured digital abstract

- [HtrA2 cleaves beta casein](#) by [enzymatic study](#) (1, 2, 3)
- [HtrA2 binds to XIAP-BIR2](#) by [surface plasmon resonance](#) ([View interaction](#))
- [HtrA2 binds to XIAP-BIR2 pep](#) by [surface plasmon resonance](#) ([View interaction](#))
- [HtrA2 binds to XIAP-BIR3](#) by [surface plasmon resonance](#) ([View interaction](#))
- [HtrA2 binds to XIAP-BIR3 pep](#) by [surface plasmon resonance](#) ([View interaction](#))
- [XIAP binds to HtrA2](#) by [pull down](#) ([View interaction](#))
- [XIAP-BIR2 binds to HtrA2](#) by [pull down](#) ([View interaction](#)) ([View interaction](#))
- [XIAP-BIR3 binds to HtrA2](#) by [pull down](#) ([View interaction](#)) ([View interaction](#)) ([View interaction](#))

Abbreviations

BIR, baculovirus IAP repeat; GST, glutathione S-transferase; HtrA2, high-temperature requirement protease A2; IAP, inhibitor of apoptosis proteins; IBM, IAP-binding motif; ITC, isothermal titration calorimetry; MBP, maltose-binding protein; MD, molecular dynamics; OMP, outer membrane porins; PDZ, post-synaptic density protein, *Drosophila* disc large tumor suppressor, zonula occludens-1 protein; rmsf, root mean square fluctuation; SPR, surface plasmon resonance; XIAP, X-linked inhibitor of apoptosis protein.

Introduction

High-temperature requirement protease A (HtrA) comprises a family of multidomain serine proteases well-known for structural complexities and involvement in numerous cellular processes such as protein quality control, the unfolded protein response, cell growth and apoptosis. This multitasking ability associates them with a number of pathophysiological conditions, such as cancer, arthritis and neurodegenerative disorders [1–4]. Despite differences in functions, HtrAs share a common architecture, comprising a serine protease domain and one or more C-terminal post-synaptic density protein, *Drosophila* disc large tumor suppressor, zonula occludens-1 protein (PDZ) domains arranged in a pyramidal oligomeric assembly ranging from trimer to 24-mer structure [5]. Although there is an overall structural conservation, significant diversity at the N-terminus, such as in the signal sequences, construes their cellular localization and functionality. These variations, along with intricate conformational dynamics, lead to subtle differences and hence provide a basis for their specificity and distinct behavior [6].

Human HtrA2, one of the key members of this family, has evolved into a mitochondrial proapoptotic molecule with the ability to induce apoptosis via multiple pathways [7]. It is synthesized as a 458 amino acid precursor protein, which includes a short N-terminal region, a serine protease and a C-terminal PDZ domain. HtrA2 undergoes maturation upon removal of first 133 residues comprising a transmembrane anchor and a mitochondrial localization signal. This cleavage exposes an inhibitor of apoptosis protein (IAP) recognizing tetrapeptide motif (AVPS) unique to this member of the HtrA family. Several members of human IAPs, including X-linked inhibitor of apoptosis protein (XIAP), cellular IAP1 and cellular IAP2, are potent endogenous inhibitors of caspase-3, -7 and -9 [8]. During apoptosis, mature HtrA2 enters the cytoplasm where it interacts with baculovirus IAP repeat (BIR) domains of IAPs such as XIAP and relieves their inhibition on active caspases [9,10] thereby promoting apoptosis [11]. However, its caspase-independent proapoptotic property is manifested by an ability to proteolytically cleave important cellular proteins, such as cytoskeleton-associated actin, α -tubulin, β -tubulin, vimentin, proteins related to translational machinery [7] and a few antiapoptotic proteins [12,13]. This unique property has made it an interesting target, especially where caspase activation is prevented either as a result of an inhibition of caspase activity or genetic inactivation of a key caspase.

The available crystal structure of inactive unbound HtrA2 [14] provides an overview of its global structural organization and mode of activation. The mature trimeric protease has a pyramidal architecture where the core serine protease domains are gated by C-terminal PDZ domains at the base and short N-terminal regions on the top, thus creating an extremely buried active-site milieu. The structure-based model suggests that PDZ is stereochemically limiting the accessibility of substrate in an otherwise proteolytically competent active-site pocket, thus hinting at a higher activity for the Δ PDZ variant [14]. By contrast, a recent study demonstrated a significant decrease in the protease activity of a PDZ lacking variant [15]. This observation highlights the importance of PDZ-mediated allosteric behavior, interdomain cross-talk and conformational plasticity as prerequisites for substrate cleavage [16]. Although PDZ acts as a regulatory domain in all of the members of HtrA family, the uniqueness of HtrA2 is manifested by its ability to bind a subset of proteins, such as IAPs, through its N-terminus and subsequently cleave them [17]. This phenomenon unequivocally accentuates the existence of multiple mechanisms of HtrA2 activation involving PDZ and other regions of the protein.

In the present study, which aimed to delineate the structural correlates of HtrA2 activation, as well as to develop a universal model for its mechanism of action, we performed comprehensive binding and enzyme kinetic analyses of the HtrA2–XIAP interaction using full-length protein, its minimal binding region and the corresponding peptides. Our studies clearly demonstrate the role of XIAP as an allosteric activator of HtrA2, thus establishing the importance of the N-terminus in the regulation of protease activity. The protease–ligand interaction is proposed to drive a series of subtle conformational alterations and coordinated structural reorganizations at the distal regulatory L3, LD and L1 loops to form a catalytically competent active-site and oxyanion hole, thus enhancing its activity by several fold. These observations suggest how structural plasticity, as mediated by its substrate binding process, complex trimeric architecture and inter-subunit networking, drive HtrA2 function. This ‘control switch’, unique to HtrA2, might be important for the variety of physiological roles that it plays in the cell. Thus, learning to specifically regulate these switches would help target distinct pathophysiological conditions without affecting normal HtrA2 functions.

Results

XIAP activates HtrA2 protease activity

Apart from deactivating IAPs [9,18], the complex organization of HtrA2 has enabled it to promote apoptosis via a caspase-independent pathway that is mediated through its serine protease activity [3,7]. To assess any role of IAPs on HtrA2 proteolytic activity, we determined its initial velocity (v_0) as a function of increasing XIAP concentration, using fixed amount of enzyme and substrate β -casein, a well established model substrate of HtrA proteases [19,20]. It was observed that, in the absence of XIAP, HtrA2 displayed a maximum velocity (V_{\max}) of $0.8 \times 10^{-9} \text{ M}\cdot\text{s}^{-1}$ (Table 1), which increased in the presence of XIAP, and the half maximal activation constant (K_{act}) was determined to be $10 \pm 0.7 \text{ nM}$ (Fig. S1A). The full-length XIAP is a homodimer consisting of three tandem BIR domains, followed by single ubiquitin associated and RING domains [21]. The BIR2 and/or BIR3 domains have been reported to be essential for the HtrA2–XIAP interaction [9]. To further explore XIAP-mediated HtrA2 activation, we tested protease activity in presence of BIR2 or BIR3, as well as peptides corresponding to the IAP-binding motif (IBM) groove of these two minimal binding domains. We observed that, in

the presence of BIR2, the cleavage rate was similar to full-length XIAP; however, the former achieves so at the expense of four-fold greater protein concentration ($58 \pm 10 \text{ nM}$) (Fig. S1B). In the case of BIR3, the rate was slightly less compared to the full length and also exhibited a high K_{act} concentration ($200 \pm 25 \text{ nM}$) (Fig. S1C). These results suggest that the BIR2 domain might be the key player in XIAP-mediated HtrA2 activation. However, the increase in the half maximal activation constant of BIR domains compared to the full length suggests that homodimeric XIAP with tandem BIRs may be important with respect to enhancing its avidity towards HtrA2 through a two-site interaction driven mechanism.

To quantify the kinetics of the complex formation between HtrA2 and BIR ligands in real time, a surface plasmon resonance (SPR) technique was used. It was observed that BIR proteins and peptides readily bound to immobilized HtrA2 (Fig. 1). Kinetic analysis of the binding data for BIR2 and BIR3 demonstrated apparent equilibrium dissociation constants (K_D) of $54 \pm 8 \text{ nM}$ and $254 \pm 18 \text{ nM}$, respectively (Table 2). A decrease in K_D for BIR2 was mainly the result of a higher association combined with moderately lower dissociation rates compared to BIR3. We correlated the binding properties for BIR proteins with those of the IBM groove peptides (Fig. 1C,D), which showed

Table 1. Steady-state enzymatic parameters for HtrA2 and its variants. The initial rates for substrate cleavage in the presence or absence of XIAP proteins or peptides were measured and fitted to the Hill form of the Michaelis–Menten equation to determine the steady-state kinetic parameters (Figs 2 and S2). Values are the mean \pm SEM and are generated from data points obtained from at least three independent experiments. NA, not applicable.

Enzyme	Activator (protein/peptide)	$K_{0.5}$ (μM)	Hill constant	Maximum velocity V_{\max} ($\text{M}\cdot\text{s}^{-1}$) $\times 10^{-9}$	Catalytic efficiency $k_{\text{cat}}/K_{0.5}$ ($\text{M}^{-1}\cdot\text{s}^{-1}$)
Wild-type	None	2.3 ± 0.1	2.8 ± 0.2	0.84 ± 0.1	1.6×10^3
Wild-type	BIR3	2.3 ± 0.2	1.8 ± 0.1	1.4 ± 0.1	3.1×10^3
Wild-type	BIR2	2.3 ± 0.2	1.6 ± 0.1	1.8 ± 0.2	3.8×10^3
Wild-type	XIAP FL	2.7 ± 0.1	1.5 ± 0.1	2.3 ± 0.1	4.7×10^3
G230A	None	3.2 ± 0.2	NA	0.17 ± 0.06	1.6×10^2
G230A	BIR3	3.5 ± 0.2	NA	0.28 ± 0.01	2.6×10^2
G230A	BIR2	3.2 ± 0.1	NA	0.38 ± 0.09	4.4×10^2
G230A	XIAP FL	3.0 ± 0.1	NA	0.51 ± 0.13	5.1×10^2
HtrA2 ^{APDZ}	None	1.4 ± 0.04	1.7 ± 0.1	0.39 ± 0.12	2.9×10^2
HtrA2 ^{APDZ}	BIR3	1.7 ± 0.1	1.5 ± 0.04	0.82 ± 0.16	5.1×10^2
HtrA2 ^{APDZ}	BIR2	1.7 ± 0.1	1.4 ± 0.1	0.91 ± 0.11	6.2×10^2
HtrA2 ^{APDZ}	XIAP FL	1.8 ± 0.1	1.5 ± 0.1	1.08 ± 0.12	8.0×10^2
Peptide-mediated activation of HtrA2					
Wild-type	BIR3 _{pep}	2.5 ± 0.2	2.2 ± 0.2	1.3 ± 0.1	2.4×10^3
Wild-type	BIR2 _{pep}	2.8 ± 0.3	1.9 ± 0.1	1.9 ± 0.1	3.3×10^3
Wild-type	PDZ _{opt}	2.6 ± 0.3	2.5 ± 0.2	1.2 ± 0.1	2.2×10^3
Wild-type	PDZ _{opt} + XIAP	2.9 ± 0.3	1.9 ± 0.2	2.2 ± 0.2	4.0×10^3
HtrA2 ^{APDZ}	BIR3 _{pep}	1.8 ± 0.1	1.5 ± 0.1	0.75 ± 0.06	5.4×10^2
HtrA2 ^{APDZ}	BIR2 _{pep}	1.3 ± 0.2	1.2 ± 0.1	0.76 ± 0.15	6.1×10^2

Fig. 1. SPR binding analyses of HtrA2 with BIR domain proteins and IBM groove peptides. HtrA2 was immobilized on a CM5 sensor chip as described in the Materials and methods. The indicated concentrations of BIR proteins and peptides (analyte) were injected over immobilized HtrA2 at a flow rate of $20 \mu\text{L}\cdot\text{min}^{-1}$ at 25°C . The sensorgram shown is for the binding kinetics of BIR2 (A), BIR3 (B), BIR2_{pep} (C) and BIR3_{pep} (D). The specific binding response was obtained by subtracting the background signal over a reference surface without immobilized protein. Data were evaluated using BIAEVALUATION software and the constants were calculated by fitting the data using a 1 : 1 Langmuir binding model.

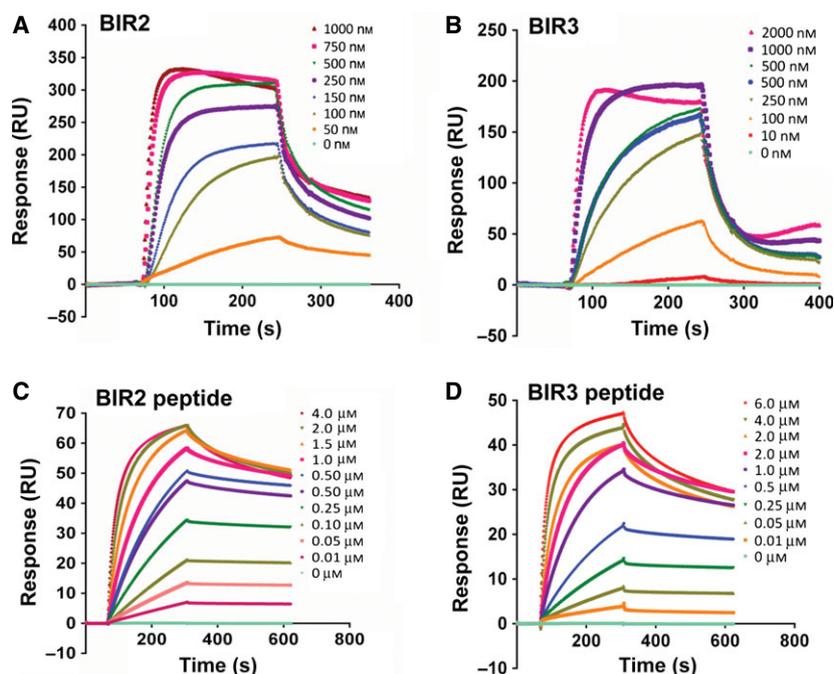


Table 2. SPR analysis of the interaction between HtrA2 and XIAP–BIR proteins and peptides. The kinetics of the interaction between HtrA2 and XIAP–BIR proteins or peptides was evaluated using a Biacore 3000 instrument. The ligand HtrA2 was immobilized on a surface that was challenged with different concentrations of BIR proteins/peptides and the rate constants were determined by applying a 1 : 1 Langmuir model to the data. The equilibrium dissociation constant (K_D) was calculated from the ratio of dissociation and association rate constants (k_d/k_a).

Ligand	Analyte	Association rate constant k_a ($1/\text{M}^{-1}\cdot\text{s}^{-1}$)	Dissociation rate constant k_d ($1/\text{s}^{-1}$)	Equilibrium dissociation constant K_D (M)
HtrA2	BIR2	1.6×10^5	8.5×10^{-3}	5.5×10^{-8}
HtrA2	BIR3	5.7×10^4	1.5×10^{-2}	2.5×10^{-7}
HtrA2	BIR2 _{pep}	1.3×10^4	2.0×10^{-4}	1.5×10^{-8}
HtrA2	BIR3 _{pep}	7.9×10^3	1.5×10^{-3}	2.0×10^{-7}

comparable apparent K_D values (Table 2). Thus, HtrA2 interacts with both the BIR2 and BIR3 domains of XIAP but shows a higher affinity for the former. In addition, these results establish that the IBM groove is critical and sufficient for mediating the XIAP–HtrA2 interaction.

Interaction with XIAP allosterically modulates HtrA2 activity

To determine the steady-state kinetic parameters, we measured the initial rates of substrate cleavage for different β -casein concentrations with or without saturating XIAP proteins or peptides. It was observed that HtrA2 cleavage of β -casein follows a sigmoidal curve with a Hill constant of 2.8 ± 0.2 (Fig. 2A). The substrate β -casein has a putative PDZ-binding consensus

(GPFPIIV) and has specific cleavage sites for serine protease, thus acting as a C-terminal self-activating substrate for HtrA2 [15]. Therefore, the positive cooperativity upon substrate cleavage suggests that binding of one substrate (β -casein) molecule to one subunit of HtrA2 trimer favors the consequent binding of others. Interestingly, upon pre-incubation of HtrA2 with XIAP, the catalytic efficiency ($k_{\text{cat}}/K_{0.5}$) increased by approximately three-fold and the Hill constant was reduced to 1.5 ± 0.1 , thus indicating greater stabilization of the relaxed (active) state of protease upon IAP binding [22]. However, the substrate concentration at half maximal velocity ($K_{0.5}$) remained unchanged, even in presence of effector (XIAP), implying the existence of a similar binding pocket. Comparable results were obtained with a saturating concentration of BIR domain proteins and peptides, as shown in Table 1.

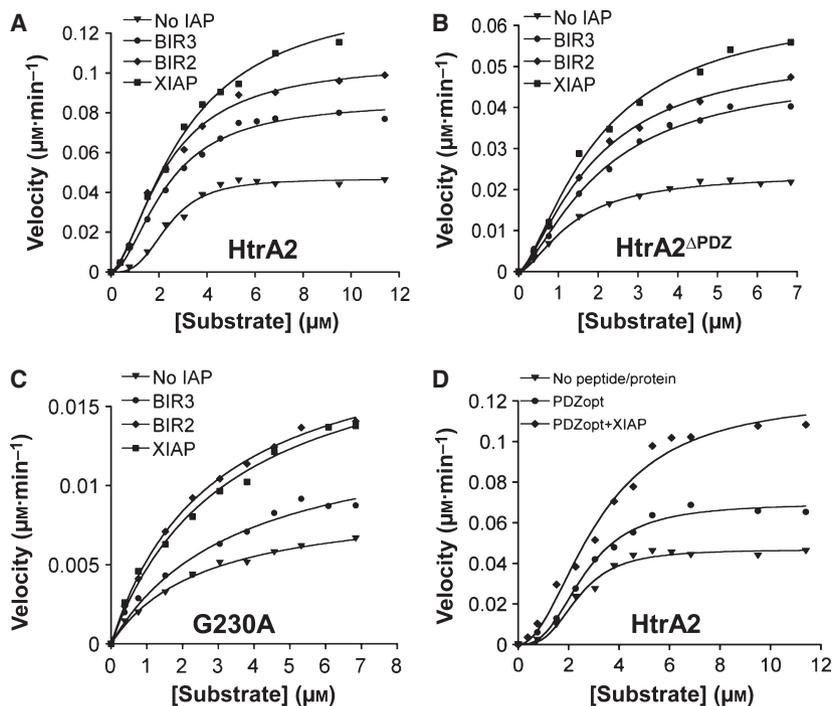


Fig. 2. Activity of wild-type HtrA2, its variant and mutant in the presence or absence of activator. The steady-state kinetics of β -casein cleavage by wild-type HtrA2 (200 nM) (A), PDZ lacking variant HtrA2 $^{\Delta$ PDZ (1000 nM) (B) and PDZ-peptide groove mutant G230A (200 nM) (C) in the presence of a saturating concentration of XIAP (20 nM), BIR2 (100 nM) and BIR3 (1000 nM) proteins. (D) Plot representing the activity of wild-type HtrA2 in the presence of either 5 μ M PDZ interacting peptide (PDZ_{opt}) and a combination of PDZ_{opt} and full-length XIAP (20 nM). The solid lines are the nonlinear least squares fit of the data to the Hill form of the Michaelis–Menten equation: $\text{velocity} = V_{\text{max}}/[1 + (K_m/[\text{substrate}])^n]$. For clarity, error bars are not shown for all of the data sets.

Here, in the presence of XIAP proteins or peptides, there is an increase in V_{max} without any significant alteration in K_m of HtrA2, which indicates that it likely follows the ‘V system’ of allosteric modulation [23]. In this model, the substrate has similar affinity for both relaxed (active) and tensed (inactive) states of enzyme; hence, there is no influence of the effector on substrate binding (and vice versa). It depends greatly on whether the effector has maximum affinity for the active state (acting as an activator) or for the inactive state (here, an inhibitor) [24]. This differential affinity of the effector towards the enzyme, along with relaxed state stabilization, shifts the equilibrium towards the active state [22,25]. Thus, XIAP possibly acts as an activator to allosterically modulate HtrA2 to form a catalytically active enzyme, which then cleaves β -casein.

PDZ domain is dispensable for XIAP-mediated HtrA2 activation

To identify whether the mechanism of HtrA2 activation via N- and C-termini effector binding is additive, we carried out an enzyme cleavage assay in the presence of optimal HtrA2 PDZ domain binding peptide (NH₂-GQYYFV-COOH), termed PDZ_{opt} [26]. It was observed that pre-incubation of HtrA2 with PDZ_{opt} peptide increased the cleavage rate of substrate (Fig. 2D). In addition, the presence of a saturating

concentration of XIAP exhibited an approximately two-fold further increase in turnover number and catalytic efficiency (Table 1), demonstrating synergy between XIAP and PDZ_{opt} induced activation.

To understand whether N-terminal XIAP-mediated activation requires any direct involvement of the C-terminal PDZ domain, we first determined the enzymatic parameters for a PDZ lacking variant (HtrA2 $^{\Delta$ PDZ) (Fig. 2B). HtrA2 $^{\Delta$ PDZ cleaved β -casein with steady-state kinetic parameters of $K_{0.5} = 1.4 \pm 0.04 \mu\text{M}$, $V_{\text{max}} = 0.4 \times 10^{-9} \text{ M}\cdot\text{s}^{-1}$ and Hill constant = 1.7 ± 0.1 (Table 1). The $k_{\text{cat}}/K_{0.5}$ of this variant was 5.5-fold less than intact trimeric HtrA2 and the apparent K_m and co-operativity were also reduced. Because β -casein is a C-terminal binding substrate, deletion of PDZ domain possibly affects the initial substrate binding process, allosteric modulation and cleavage, thereby leading to a decrease in co-operativity and catalytic efficiency. However, interestingly, the addition of XIAP (proteins or peptides) to HtrA2 $^{\Delta$ PDZ enhanced the β -casein cleavage rate to a fold increase similar to full-length HtrA2 (approximately 2.8-fold) with a concomitant decrease in cooperativity (Figs 2B and S2). This suggests that the N-terminal ligand is capable of independently inducing conformational changes in the protease domain to form an active conformer.

To further clarify the role of XIAP in HtrA2 activation, we tested the activity of a PDZ-binding groove mutant. PDZ domains have a canonical binding site

comprising a highly conserved ‘G- Φ -G- Φ motif’, where Φ denotes hydrophobic residues. The first Gly residue is highly variable among the PDZ domains, whereas the second and fourth residues are hydrophobic (Val, Ile, Leu or Phe) [27]. This recognition sequence is represented by YIGV in HtrA2. We generated a YIGV groove mutant (YIAV), denoted here as G230A, to prevent substrate interaction with the PDZ domain, which was validated using isothermal titration calorimetry (ITC) using PDZ_{opt} peptide. It was found that, although wild-type HtrA2 interacts with the peptide, the mutant fails to do so, thus confirming the importance of YIGV in the initial substrate binding (Fig. 3). We then determined the enzymatic properties of this mutant with or without a saturating concentration of IAPs. As anticipated, the steady-state kinetic parameters showed a significant decrease in the catalytic efficiency of G230A compared to wild-type (Table 1). Surprisingly, it did not exhibit any cooperativity and the plot was found to be hyperbolic (Fig. 2C). This Michaelis–Menten behavior of the enzyme can be attributed to the lack of substrate-induced stabilization of the active-site, as well as to the steric block imparted by the PDZ domain to the specificity pocket as a result of a nonfunctional

‘YIGV’ binding groove. This could have led to loss in cooperativity and decreased activity. However, the increase in V_{\max} by 2.4-fold upon the addition of IAPs re-iterates the role of the N-terminus with respect to the regulation of HtrA2 activity in a PDZ independent manner (Table 1).

Disruption of the XIAP–HtrA2 interaction prevents protease activation

HtrA2 interacts with XIAP via an N-terminal reaper-like motif (AVPS) similar to Smac/DIABLO [11]. To identify the role of the XIAP–HtrA2 interaction in the activation of the protease, we created an AVPS mutant (Δ , V, P3A, S; termed A1 Δ P3A). As expected, this variant did not interact with XIAP, as observed by a glutathione *S*-transferase (GST) pull-down assay (Fig. 4A). The proteolytic activity of A1 Δ P3A was found to be comparable to wild-type and remained unchanged even in the presence of XIAP proteins (Fig. 4D). This result confirms the requirement of the XIAP–protease interaction for N-terminus-mediated allosteric activation of the protease. Similarly, we targeted the IBM groove of XIAP that is responsible for its interaction with HtrA2. Since two IBM groove residue pairs (Glu, His and Gln, Trp) confer selectivity to BIR2 and BIR3, respectively, for interaction with caspase-3 and -7 [28], we generated mutants of XIAP–BIR2 (E219R, H223V) and XIAP–BIR3 (Q319R, W323V), aiming to explore the importance of these residues in their interaction with HtrA2. Pull-down studies with the IBM groove mutants of BIR domains showed a complete abrogation of the interaction, whereas it was partially retained in the case of full-length XIAP with a single IBM groove mutation (Fig. 4C). Moreover, these mutants failed to enhance the proteolytic activity, probably as a result of their inability to interact with HtrA2 (Fig. 4D). Altogether, these studies strongly demonstrate the role that the IBM groove plays in mediating the HtrA2 activation.

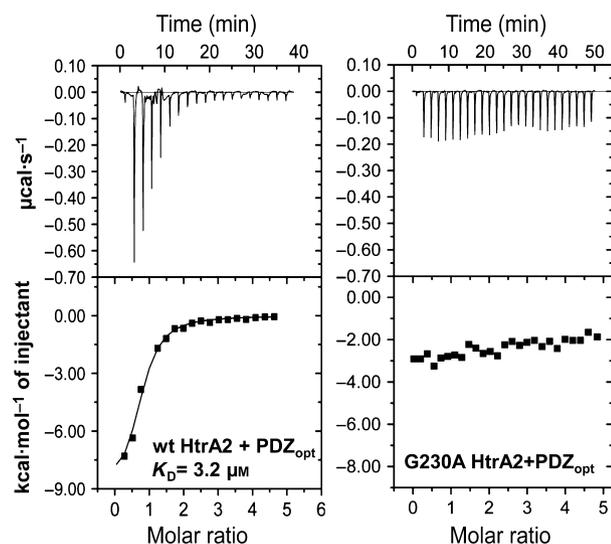


Fig. 3. Interaction analyses of PDZ_{opt} with HtrA2 wild-type and G230A mutant measured using ITC. Binding thermograms for PDZ_{opt} peptide (H₂N-GQYYFV-COOH) (0.5 mM) with HtrA2 wild-type (0.03 mM) (left) and PDZ-peptide groove mutant, G230A (0.033 mM) (right). Data were recorded at 298 K. In each titration, the upper panel shows the raw data, whereas the lower panel corresponds to corrected integrated binding isotherm together with the fitted binding curve. The dissociation constant was calculated to be 3.2 μ M for wild-type, whereas no significant heat change was observed for the PDZ-peptide groove mutant.

Conformational transitions and flexibility in regulatory loops govern HtrA2 activity

With the aim of understanding the conformational changes that mediate allosteric activation, we adopted an *in silico* approach. Molecular dynamic (MD) simulation is a powerful tool for predicting the conformational plasticity and intricate molecular motions of large enzymes and their complexes [29]. To gain insight into the structural rearrangements that occur in a nanosecond time scale upon activator binding, MD analyses of BIR3 peptide-bound HtrA2 complex and

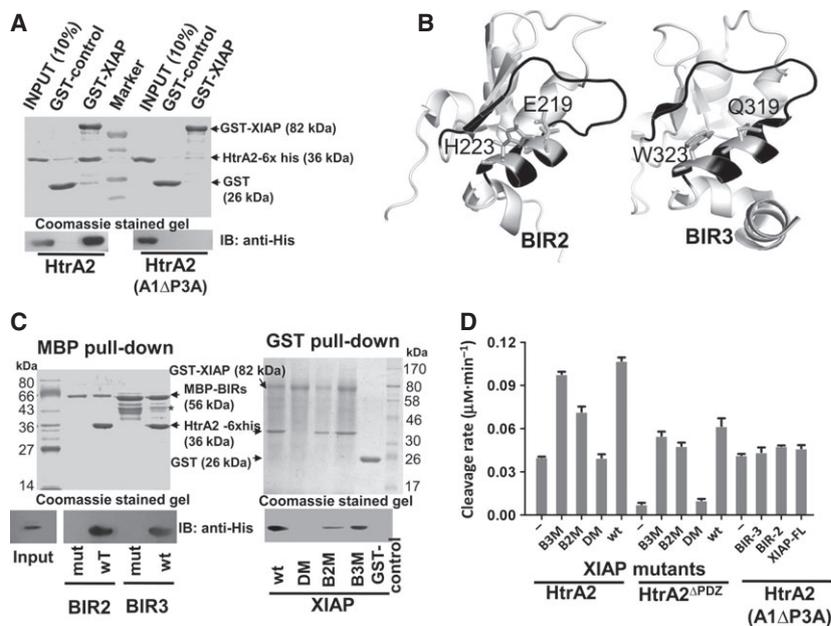


Fig. 4. Mutational analysis of the XIAP–HtrA2 interaction. (A) GST pull-down assay of wild-type HtrA2 and AVPS mutant (A1ΔP3A) with XIAP. The A1ΔP3A mutant (prey) was defective in binding to GST-XIAP (bait) compared to wild-type. GST control did not show any binding to HtrA2. (B) Structure of BIR2 (Protein Data Bank code: [1c9q](#)) compared to BIR3 (Protein Data Bank code: [1f9x](#)), highlighting the IBM groove in black. Residues for IBM groove mutations E219, H223 of BIR2 and Q319, W323 of BIR3 are shown as sticks. (C) Pull-down assay of IBM groove mutants using full-length XIAP, BIR2 and BIR3 domains. Left: MBP pull-down was carried out with MBP-BIR2 and -BIR3 wild-type and IBM groove mutants BIR2 (E219A, H223A) and BIR3 (Q319A, W323A). A representative coomassie stained SDS/PAGE and corresponding immunoblotting with anti-His for detection of His-tagged HtrA2 shows that HtrA2 failed to associate with mutants (*MBP contamination). Right: GST pull-down was performed with XIAP and its mutants. XIAP with both BIR domains mutated (DM) showed a complete loss of interaction, whereas the single domain mutants XIAP–BIR2 (E219A, H223A) (B2M) and XIAP–BIR3 (Q319A, W323A) (B3M) retained binding with HtrA2. (D) Basal cleavage rate of β -casein by HtrA2 and HtrA2^{APDZ} in the presence of XIAP and its mutants. The DM failed to increase the enzymatic activity of HtrA2, whereas B3M and B2M enhanced the activity similar to wild-type BIR2 and BIR3, respectively. The A1ΔP3A mutant incapable of binding to XIAP did not show any increment in proteolytic activity in the presence of IAP.

unbound form were performed. Because the BIR3–Smac mimetic complex has already been reported in the literature [30,31], demonstrating the interaction of a similar molecule (Smac) with the IBM groove, we carried out our MD simulation studies with BIR3 peptide. A trimeric HtrA2^{APDZ} (residues 1–211) was generated to perform docking and MD simulation, resulting in the BIR3 peptide-bound complex (BIR3_{pep}–HtrA2^{APDZ}). Simultaneously, peptide-unbound HtrA2^{APDZ} was subjected to MD simulation for comparison with the bound complex. We observed a significant difference in the conformation of regulatory loops of the protease upon IAP binding. A root mean square fluctuation (rmsf) plot of the trajectories for BIR3_{pep}-bound complex showed higher relative fluctuations in loop LA (residues 37–41), LD (residues 126–140) and L3 (residues 142–162) (Figs 5A and S3). Domain wise rmsd plots demonstrated major changes in sensor loop L3 (Fig. 5B,C). However, loop L2 (190–196), which harbors the specificity pocket,

remained unchanged in the bound or unliganded states, suggesting the presence of a well-formed substrate binding pocket (Fig. 5A). This observation corroborates very well with the substrate binding affinity (apparent K_m), which was not appreciably different in the presence or absence of the activator. The overall structure for peptide-bound and unbound forms is very similar, with a mean rmsd of 1.21 Å. The difference could mainly be attributed to structural deviation near the activation domain loops, as well as some variations at the N-terminus (residues 1–9) (Fig. 6A,B). Further characterization of the structural changes was performed to understand the mechanistic details of activation.

In serine proteases, the formation of catalytic triad in an arrangement sufficiently close for electron transfer from Asp to Ser through His is a prerequisite for an active enzyme [32]. For mature HtrA2, the catalytic triad that includes His65, Asp95 and Ser173 is malformed in the available unbound crystal structure [14].

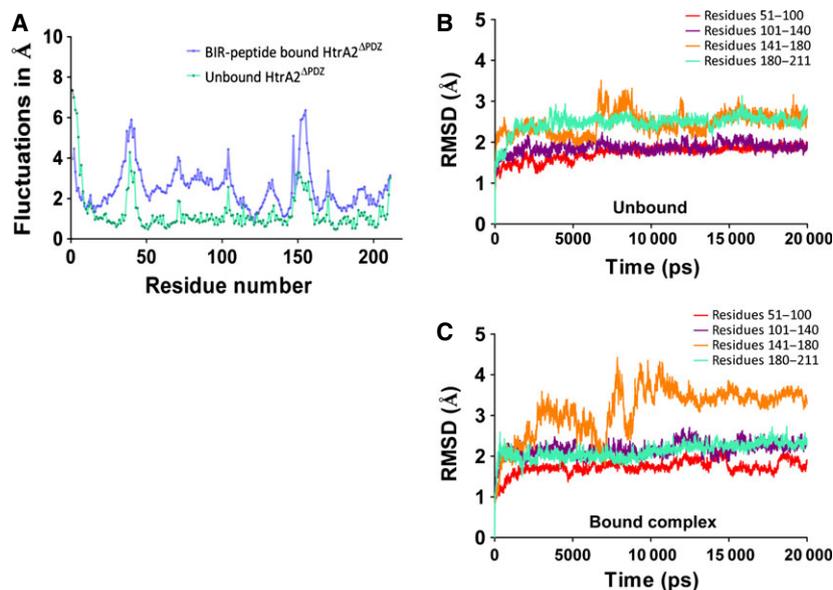


Fig. 5. Graphical representation of movements in the peptide bound and unbound form of trimeric HtrA2^{APDZ}. (A) rmsf plot of 20-ns MD simulation trajectory for BIR3 peptide-bound and unbound HtrA2^{APDZ}. (B) Domain wise rmsd plots for peptide unbound and (C) bound HtrA2^{APDZ}. Residues 140–180, comprising loops LD, L3 and L1, display significant deviations upon peptide binding.

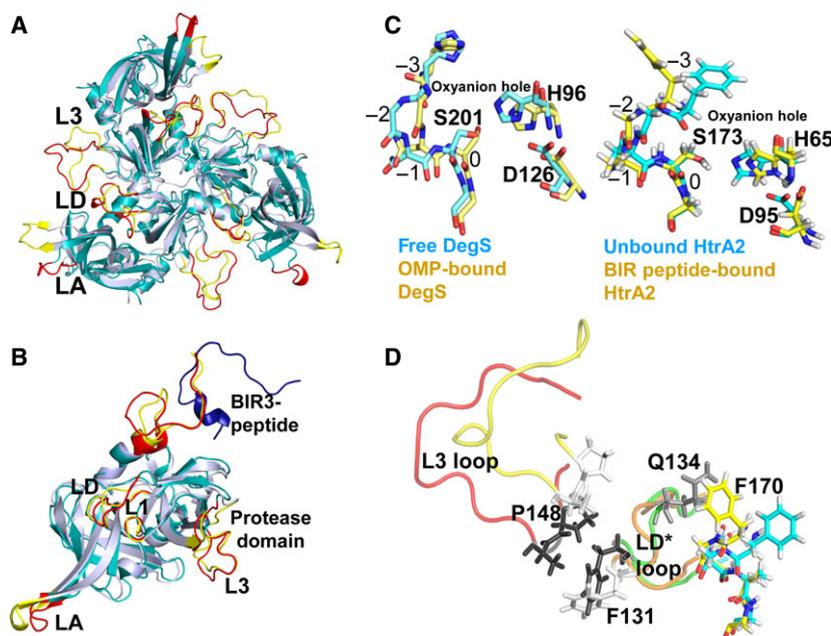


Fig. 6. Monitoring conformational transition and rearrangements in BIR peptide-bound and unbound trimer HtrA2^{APDZ}. (A) Structural alignment of BIR peptide bound-HtrA2^{APDZ} (cyan) and unbound (light blue) in cartoon representation. Movement in loops LA, LD and L3 is highlighted in red (peptide bound) and in yellow for the unbound structure. (B) Structural overlay highlighting a single subunit of BIR3 peptide (dark blue) complex (cyan) and unbound (light blue) with the re-oriented regions shown in red (bound) and yellow (unbound). (C) Left: comparison of oxyanion hole and active-site residues for inactive (OMP-free) and active (OMP-bound) forms of *E. coli* DegS shown as sticks. The residues in oxyanion are indicated as 0 for active-site Ser and –1, –2 and –3 for the preceding residues. In the inactive form, the oxyanion hole is distorted, especially in the amide nitrogen atom of Gly at the –2 position. Also, the active-site residues vary substantially. Similarly, on the right, unbound HtrA2 shows a distorted oxyanion hole and a malformed active-site, whereas the BIR peptide-bound complex shows re-orientation of Gly to form catalytically active HtrA2. (D) Conformational changes between the bound and unbound structure. Upon peptide binding, the L3 loop (red) of one subunit moves closer to the LD* loop (orange) of the adjacent unit. The side chain of residues (shown as black sticks) from the L3 loop form nonbonded interactions with F131 of the LD* loop, leading to rearrangement of oxyanion hole as a result of its hydrophobic packing with the LD loop. However, L3 (yellow) and LD* (green) are separated in the unbound structure (*residue from the adjacent subunit).

Structural comparison of the unbound protease with the BIR3_{pep}-bound complex shows relative movements in the active-site triad residues (Fig. 6C). It has been well established that the distance between the N^ε atom of His and the O^γ atom of Ser in active proteases is < 3.5 Å, whereas, in the inactive form, it is > 3.5 Å [5]. The atomic distances between the nitrogen (ε) atom of His65 and the oxygen (γ) of Ser173 decreases from 5.1 Å to 3.2 Å in the bound form and that between the nitrogen (δ) of His65 and the oxygen (δ) of Asp95 for the peptide-bound complex is 0.3 Å shorter compared to the unbound structure (Table S1).

Apart from the catalytic triad, another key factor for proteolysis by serine proteases is stabilization of the negative charge on the carbonyl oxygen of the tetrahedral alkoxide intermediate (oxyanion hole) [33], which resides in loop L1 (169–173) of HtrA2. Structural alignment of the oxyanion hole residues for HtrA2 (BIR3_{pep}-bound and unbound forms) in comparison with the free and outer membrane porins (OMP)-bound *Escherichia coli* DegS (closest structural ortholog of HtrA2) is shown in Fig. 6C. In the BIR-bound form, the nitrogen atom of Gly171 (−2 position) acts as a hydrogen donor to form proper oxyanion hole, whereas, in the unbound form, it is oriented in the opposite direction, creating a malformed structure. Based upon structural overlay, the re-orientation of loop L1 appears to occur as a result of its increased proximity towards another mechanistically important loop LD via a hydrophobic interaction. It thereby shifts the phenyl ring of F170 to accommodate the LD loop and form a competent oxyanion hole, as shown in Fig. 6D. Hence, the orientation of loops that accommodate the oxyanion hole and the catalytic triad residues might render the protease active or inactive in the presence or absence of ligand, respectively.

In the HtrA group of serine proteases, intricate intermolecular interaction among different subunits of an oligomer leads to a proper active-site formation [6]. From the structural alignment of BIR3_{pep}-HtrA2^{ΔPDZ} with the unbound form, we observed significant deviations in loop L3 and LD upon peptide binding. This movement in particular appears to allow loop L3 of one subunit to move towards the LD of a neighboring unit. Specifically, in the bound complex, P148 and A149 of the L3 loop form hydrogen bond and Van der Waal contacts with P130 and F131 of the LD loop from an adjacent subunit, which otherwise is separated in the unbound structure (Fig. 6D). Thus, all of these loop movements might coordinate to bring about the observed concerted conformational rearrangements in

(and around) the active-site, leading to protease activation upon IAP binding.

Discussion

HtrA2 has been described as a stress-induced serine protease in mammalian cells [34] with multiple functions. Under physiological conditions, it participates in the maintenance of mitochondrial homeostasis, where its protective function transforms into proapoptotic behavior in response to stress-inducing agents. It mediates apoptosis through classical pathways via caspase activation and also by a less understood caspase-independent mechanism [7]. This nonclassical pathway of apoptosis depends on the proteolytic activity of HtrA2, which is primarily regulated by protein–protein interactions. To date, a wide repertoire of proteins binding to the C-terminal PDZ domain has been found to stimulate the protease activity, thereby promoting cell death [35]. Interestingly, in addition to the regulatory role played by the PDZ domain, HtrA2 exhibits an additional level of functional modulation mediated through its N-terminus, which is evident from its IAP binding and cleavage. This phenomenon emphasizes multiple modes of HtrA2 activation and regulation. Furthermore, its association with severe pathophysiological conditions, such as cancer and neurodegenerative disorders, makes it a promising therapeutic target for which delineation of the basis and global mode of its activation is required.

The high-resolution crystal structure of unbound HtrA2 [14] provides a broader picture of its structural organization and the mechanism of activation. However, the importance of conformational flexibility that is required for serine protease function cannot be unambiguously resolved by the snapshots of selected conformational states provided by crystallography data. The regions with increased flexibility, such as loop segments, appear to be preferred for allosteric modulation of the serine protease domain [36]. The trimeric HtrA2 has a unique N-terminus and a complex network of regulatory loops (LA, LD, L1 and L2), which forms the ‘activation domain’. The relative orientations of this activation domain along with flexible loop L3 and the PDZ–protease interface linker might be critical for defining its functions. However, the solved unbound structure with the partially missing flexible regions could not explain the dynamics that regulate HtrA2 activity, as is evident from the existing proposed model of activation [14]. In the present study, we have identified a novel mechanism that regulates HtrA2 activity through its N-terminal region.

Enzyme kinetic studies with wild-type HtrA2 and its variants suggest that binding of XIAP proteins or peptides at the N-terminus allosterically modulates the protease and shifts the intrinsic equilibrium from a tense to relaxed conformation, thereby increasing the overall catalytic efficiency in a PDZ-independent yet synergistic activation mechanism. This N-terminal-mediated regulation of HtrA2 appears to be novel among the HtrA family of serine proteases. For example, *E. coli* DegS is proteolytically inactive as a result of inhibitory interaction of PDZ with the serine protease domain, which keeps the protease in its basal state [25]. This inhibition is relieved either by the addition of OMPs that bind to PDZ and rearrange the active-site or by deletion of the PDZ domain to form an active enzyme, suggesting that PDZ might act as the sole regulatory domain [22,37]. However, in HtrA2, we have established that the unique N-terminal region imposes an additional control in regulating its activity. The complexity in its behavior might be required for targeting apoptosis through multiple pathways, as well as for a myriad of other functions that it might perform. Based on our observations, we hypothesize that, during apoptotic stress, HtrA2 facilitates the ‘caspase-dependent pathway’ by relieving the inhibition of IAPs on active caspases [9,10]. This IAP–protease interaction might then trigger the nonclassical ‘caspase-independent mechanism’ by up-regulating the protease activity of HtrA2, which, in conjunction with PDZ binding proteins, promotes substrate cleavage and hence cell death.

In the HtrA family of serine proteases, dynamic intermolecular processes play a crucial role in activation and hence functions [6]. The remodeling is initiated in most, if not all, HtrA proteins by sensor loop L3 and, subsequently, a cascade of conformational changes occur along $L3 \rightarrow LD^* \rightarrow L1^*/L2^*$ (the asterisk denotes the contribution from a neighboring subunit), which enables the active-site to switch into a ‘proteolytically ON’ state. Although the general mechanism is conserved among the family, the activation signal

detection might vary [38]. Conformational dynamic analyses revealed that, in unliganded HtrA2, specifically the regulatory loops L1, L3 and LD are disordered, rendering the catalytic site and oxyanion hole dysfunctional. A closer look at (and around) the active-site of bound HtrA2 shows stabilization of the oxyanion hole and proper positioning of the catalytic triad. These regulatory loops undergo disorder-to-order transition, thus allowing the formation of a well-defined activation pocket. The occurrence of such structural and conformational flexibility provides a molecular explanation for the allosteric activation of HtrA2. From our studies, we propose that HtrA2 exists in an equilibrium between an inactive (E) and most active conformation (E*). In the E state, the protease exhibits poor enzymatic properties; however, a conformational switch triggered by N-terminal and C-terminal ligands shifts the intrinsic equilibrium towards the E* state (Fig. 7). Thus, the equilibrium between the inactive and active forms of HtrA2 could be used to tune the enzyme in the ‘ON’ or ‘OFF’ states upon binding of suitable activator or inhibitor molecules.

Subsequent to its inception [24], the allosteric mechanism of enzyme regulation has intrigued biochemists and structural biologists, which has resulted in the derivation of ways to manipulate this behavior with desired characteristics. An event occurring at any distal site that affects the catalytic efficiency of an enzyme via long-range communication provides an appropriate explanation of linkage and cooperativity [39]. Recent studies have emphasized the importance of allostery as an intrinsic property of all of the dynamic proteins that exist in multiple states at equilibrium and tend to show large conformational changes linked to ligand binding or substrate catalysis [40]. As a result of the inherent protein-cleaving ability of proteases and their vital biological roles in all living organisms, their enzymatic activity needs to be strictly regulated. The HtrA family of proteases has developed an intrinsic allosteric mechanism for the precise control of their enzyme activity

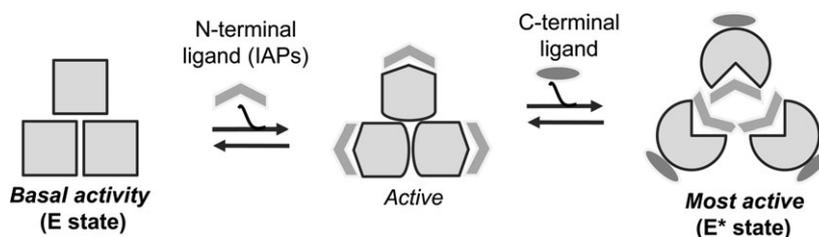


Fig. 7. Proposed simplistic model for HtrA2 activation. The model assumes that HtrA2 exists in an equilibrium of inactive (E) and most active (E*) states. The complex allosteric propagation mediated upon binding of the N- and C-terminal ligands synergistically transforms the protease into the most active conformer. N-terminal ligand (IAP) shifts the conformational equilibrium of HtrA2 to an active state that is further stabilized by reaction with C-terminal ligands/substrate.

and specificity, which is mediated via regulatory PDZ domains. It can be presumed that the overall mechanism is conserved across the family from *E. coli* DegS, DegP, DegQ and *Mycobacterium* HtrA2 to *Arabidopsis thaliana* Deg1 [25,38,41–43]. Interestingly, in human HtrA2, evolution has possibly endowed it with an additional regulatory switch via the N-terminal binding of substrates to impart rigid control on its activation, which might be crucial for the different roles that it plays under normal and diseased conditions [1–4]. This distinctive mode of regulation offers promising opportunities for the specific control and functional modulation of a wide range of biological processes associated with HtrA2. In the recent past, allosteric regulators were shown to have several advantages over their orthosteric counterparts, with lower dose requirements and greater subsite specificity [44,45]. It would therefore be possible to design and construct suitable peptidomimetics or small molecule analogs that could bind the regulatory domains and stimulate protease activity, thus promoting apoptosis. Recently, the proteolytic activity of HtrA2 was shown to be critical for inducing apoptosis in prostate cancer cells mediated via integrin α [46]. Furthermore, its role in the degradation of Wilm's tumor suppressor protein, WT1, shows the possibility of targeting HtrA2 in the therapy of Wilm's tumor, where WT1 is overexpressed [47,48]. The E–E* equilibrium therefore proposes new ways for enhancing the activity of HtrA2 by stabilization of the E* form, as shown recently for thrombin [49] and apoptotic procaspase-3 and -6 [50]. This advancement in our understanding of the mechanism of action of HtrA2 will help in the targeting of its allosteric site with tailored effectors, hence representing a powerful approach for devising therapeutic strategies against the varied diseases associated with it.

In summary, the results of the present study provide the molecular basis of a novel N-terminal-mediated regulation of HtrA2 activity. Ligand binding at the N-terminal IAP-binding motif of HtrA2 leads to the activation of an allosteric signal through a series of subtle conformational rearrangements essential for the formation of the most active state of the protease. The present study showcases a combination of computational, biochemical, biophysical and functional enzymology approaches aiming to clearly demonstrate how precise coordination between ligand binding and flexible loop movements at a site distal from catalytic pocket regulates HtrA2 proteolytic activity. Moreover, the peptide-based activation process highlights the possibility of designing suitable peptidomimetics or small molecule analogs for manipulating HtrA2 functions, specifically for disease intervention.

Materials and methods

Peptides

All of the peptides were commercially synthesized (USV Ltd, Mumbai, India) with > 95% purity. The primary sequences of the peptides were: HtrA2-PDZ_{opt} (NH₂-GQYYFV-COOH) [26], BIR2_{pep} residues (NH₂-²⁰⁴GGKLKNWPCDRAWSEHRRHF²²⁴-COOH), BIR3_{pep} residues (NH₂-³⁰⁵GGLTDWKPSDPWEQHAKWY³²⁴-COOH) and control nonspecific peptide (NH₂-KNNPNNAHQN-COOH).

Recombinant protein production and purification

Mature (Δ 133) HtrA2 full length with C-terminal his₆-tag in pET-20b (Addgene, Cambridge, MA, USA) and PDZ lacking variant HtrA2 ^{Δ PDZ} were expressed and purified as described previously [15]. The canonical PDZ-peptide groove mutant (G230A) was generated using site-directed mutagenesis (Stratagene, Austin, TX, USA). Full-length XIAP clone in pGEX-4T with GST-tag was obtained from Addgene. XIAP–BIR2, BIR3 domains and mutants were subcloned into modified pMALc5-TEV vector (New England Biolabs, Ipswich, MA, USA) that provides an N-terminal maltose-binding protein (MBP) tag. The BIR domain and IBM groove mutant tested were XIAP–BIR2: 124–240, E219R/H223V; XIAP–BIR3: 241–356, Q319R/W323V. The IBM groove mutations were also introduced in GST–XIAP full length as XIAP–E219R/H223V (BIR2 domain mutant), XIAP–Q319R/W323V (BIR3 domain mutant) and XIAP–E219R/H223V, Q319R/W323V (BIR2 and BIR3 domain mutant). Recombinant proteins were expressed in *E. coli* strain Rosetta (DE3) (Novagen, Billerica, MA, USA). Cells were grown at 37 °C until D_{600} of 0.6 was reached and then induced with 0.4 mM isopropyl thio- β -D-galactoside. Cells were further cultured at 18 °C for 20 h post induction. Proteins with the GST-tag were purified by affinity chromatography using GST-sepharose resin (Novagen), whereas MBP-tag proteins were purified using amylose resin (New England Biolabs) in buffer: 20 mM Na₂HPO₄/NaH₂PO₄ (pH 7.8) containing 100 mM NaCl, 25 mM β -mercaptoethanol, 50 μ M Zn(Ac)₂. The MBP-tag was cleaved using TEV protease [51] and was further subjected to size exclusion chromatography. All of the fractions with > 95% purity as estimated by SDS/PAGE were stored in aliquots at –80 °C.

Enzyme assay

The protease activity of wild-type and its variants was determined using substrate fluorescein isothiocyanate labelled β -casein (Sigma Chemicals, St Louis, MO, USA) as described previously [15]. The fluorescent substrate cleavage was measured by incubating respective concentration of enzymes in the presence or absence of XIAP (proteins or peptides) with

increasing concentrations (0–25 μM) of β -casein at 37 °C in cleavage buffer (20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$, pH 7.8, containing 100 mM NaCl, 0.1 mM dithiothreitol). Fluorescein isothiocyanate fluorescence was monitored in a multi-well plate reader (Berthold Technologies, Oak Ridge, TN, USA) using an excitation wavelength of 485 nm and emission at 545 nm. Reaction rates (v_0) were calculated using linear regression analysis. XIAP and BIR domain protein activation curves were fitted to the equation: velocity = basal + $\max/[1 + [K_{\text{act}}/(\text{peptide})]^n]$, where K_{act} is the half maximal activation constant and n is the Hill constant [22]. The steady-state kinetic parameters were calculated from the reaction rates by fitting the data to the Hill form of the Michaelis–Menten equation, velocity = $V_{\text{max}}/[1 + (K_{0.5}/[\text{substrate}])^n]$, where V_{max} is the maximum velocity and $K_{0.5}$ is the substrate concentration at half maximal velocity using KALEIDAGRAPH (Synergy Software, Reading, PA, USA). All of the experiments were carried out independently in triplicate and the mean \pm SEM values are reported in the plots.

SPR analysis

Real-time kinetic analyses of HtrA2 and XIAP–BIR interaction were performed using SPR (Biacore 3000; GE Healthcare, Uppsala, Sweden). Carboxymethylated sensor chips (type CM5) were first activated with a 1 : 1 mixture of 0.2 M *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide and 0.05 M *N*-hydroxysuccinimide in water. The ligand, HtrA2 (50 $\mu\text{g}\cdot\text{mL}^{-1}$ in 10 mM sodium acetate, pH 4.5), was immobilized on the activated CM5 sensor chip by amine coupling (Biacore) and the unreacted sites were blocked with 1.0 M ethanolamine–HCl (pH 8.5). Control flow cells were activated and blocked in the absence of HtrA2 protein. The analytes, BIR proteins (BIR2 and BIR3) and corresponding to IBM groove peptides of these two domains were then injected as a function of increasing concentration in running buffer (10 mM Hepes pH 7.5, 200 mM NaCl) under continuous flow of 20 $\mu\text{L}\cdot\text{min}^{-1}$ at 25 °C. The chip was regenerated after each cycle by washing it with 2 M NaCl followed by running buffer. The sensorgrams for specific interactions were obtained by subtracting the reference unimmobilized flow cell response from that of the sample and were evaluated using the BIA-EVALUATION software package, version 4.1 (GE Healthcare). Data were analyzed by fitting to a 1 : 1 Langmuir binding model of both the association and dissociation phases. The apparent equilibrium dissociation constants (K_D) were determined from the ratio of the dissociation and association rate constants (k_d/k_a).

ITC

ITC measurements were performed at 25 °C using a MicroCal ITC200 (GE Healthcare). All of the solutions were

degassed before titration. The calorimetry cell contained 200 μL of 33 μM wild-type or G230A mutant HtrA2 in 20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ buffer (pH 7.8) containing 100 mM NaCl. The titration was initiated with one 0.4- μL injection followed by 20 injections of 1.5 μL each of PDZ_{opt} peptide reconstituted in the same buffer. To correct for the heat of dilution, peptide injection into a cell containing sample buffer was carried out under identical conditions and subtracted from the raw data prior to analysis. The data were analyzed using MicroCal ORIGIN software (GE Healthcare) with the integrated heat peaks fitted to a one site-binding model.

Pull-down experiment

For the binding assay, MBP-tag BIR domain and GST-tag XIAP wild-type and mutant proteins were incubated on amylose and GST-sepharose resins respectively for 1 h at 4 °C. Beads were washed three times with binding buffer (20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$, pH 7.8 containing 200 mM NaCl, 1 mM dithiothreitol, 0.1% TritonX-100). The washed beads were further incubated with equivalent amount of wild-type HtrA2 for 3 h at 4 °C. The unbound HtrA2 was removed by washing the beads three times with the binding buffer. The bound proteins were separated on 12% SDS/PAGE and were probed with anti-His antibody (Abcam, Cambridge, MA, USA) against HtrA2.

Protein modeling and docking

The crystal structure of inactive unbound HtrA2 (Protein Data Bank code: [1lcy](#)) was obtained from the Protein Data Bank. The protein has missing N-terminal residues (AVPSP), a flexible loop (¹⁴⁹ARDLGLPQT¹⁵⁷) and a linker region (²¹¹RGEKKNSSSGISGSQ²²⁵). The loop and linker region were modeled and refined using PRIME, version 3.0 [20] (Schrödinger, LLC, New York, NY, USA). Missing N-terminal residues (AVPSP) were modeled using the ‘Cross-Link Proteins’ module of BIOLUMINATE, version 1.2 (Schrödinger, LLC). The ‘Crosslink Proteins’ module allows cross-linking of prepositioned proteins by connecting chain termini with peptide linkers. The HtrA2 trimer model without PDZ domain (HtrA2 ^{Δ PDZ}) residues 1–211 was generated using the PRIME, version 3.0 ‘Homo-multimer program’, which allows the generation of a homomultimer from a monomeric protein. Docking of HtrA2 trimer with the BIR3 IBM groove peptide (GGLTDWKPSQDPWEQ-HAKWY) was performed using the ‘Protein–Protein Docking’ program [52] of BIOLUMINATE, version 1.2 (Schrödinger, LLC). Protein–protein docking calculations were performed by specifying 70 000 ligand rotations to probe, which involves sampling every 5° in the space of Euler angles. Protein–protein docking of the HtrA2 trimer with BIR3 produced N-terminal docked poses. Three poses showed

that the peptide docked at N terminal tetrapeptide AVPS region and the docked poses have no steric clashes. BIR3 peptide residues W323, W310, A321 and L307 showed interactions with each chain in the trimer.

MD simulation and analysis

MD simulation was performed and analyzed using DESMOND, version 3.4 (Schrödinger, LLC) as described previously [15,53]. Briefly, protein structures were solvated with a simple point charge orthorhombic water box with a 10 Å buffer space. The system was neutralized by replacing water molecules with sodium and chloride counter ions. Nose–Hoover thermostats and the Martina–Tobias–Klein method were used to maintain a constant simulation temperature and to control pressure, respectively. The entire system was equilibrated using the default protocol provided in DESMOND. Two rounds of steepest descent minimization were performed, followed by series of four MD simulations. The default equilibration was further followed by 5000 ps of NPT (i.e. constant number of particles, pressure and temperature) simulation to equilibrate the system. A total of 20 ns of NPT production simulation was then run and coordinates were saved every 2 ps. Simulation data were analyzed with the help of the analytical tools in the DESMOND package. For MD simulation quality analysis, the potential energy of the protein and the total energy of entire system was calculated. The lowest potential energy conformations were used for structural analysis of peptide bound and unbound structures. The trajectories of peptide bound complex and unbound forms were then compared based on their overall calculated domain wise rmsd and rmsf values. These are graphically represented using GRAPHPAD PRISM, version 5.0 (GraphPad Software, San Diego, CA, USA).

Acknowledgements

We thank Shruthi Kanthaje for help with protein purification, as well as members of our laboratory for helpful discussions. We are grateful to Dr Smita Mahale, NIRRH, for providing the SPR facility and Dr Pradip Chaudhari, ACTREC, for his critical feedback on the manuscript. This work is supported by the Science and Engineering Research Board (SERB), Department of Science and Technology (DST), Govt of India (grant number SERB/F/108/2012-13).

Author contributions

KB and NS conceived and designed the experiments. NS, AD and AC performed the experiments. NS, AD, AC, MS and KB analyzed the data. NS and KB wrote the paper. All of the authors read, revised and approved the manuscript submitted for publication.

References

- Chien J, Campioni M, Shridhar V & Baldi A (2009) HtrA serine proteases as potential therapeutic targets in cancer. *Curr Cancer Drug Targets* **9**, 451–468.
- Polur I, Lee P, Servais J, Xu L & Li Y (2010) Role of HTRA1, a serine protease, in the progression of articular cartilage degeneration. *Histol Histopathol* **25**, 599–608.
- Suzuki Y, Takahashi-Niki K, Akagi T, Hashikawa T & Takahashi R (2004) Mitochondrial protease Omi/HtrA2 enhances caspase activation through multiple pathways. *Cell Death Differ* **11**, 208–216.
- Zurawa-Janicka D, Skorko-Glonek J & Lipinska B (2010) HtrA proteins as targets in therapy of cancer and other diseases. *Expert Opin Ther Targets* **14**, 665–679.
- Singh N, Kuppili RR & Bose K (2011) The structural basis of mode of activation and functional diversity: a case study with HtrA family of serine proteases. *Arch Biochem Biophys* **516**, 85–96.
- Clausen T, Kaiser M, Huber R & Ehrmann M (2011) HTRA proteases: regulated proteolysis in protein quality control. *Nat Rev Mol Cell Biol* **12**, 152–162.
- Vande Walle L, Lamkanfi M & Vandenabeele P (2008) The mitochondrial serine protease HtrA2/Omi: an overview. *Cell Death Differ* **15**, 453–460.
- Deveraux QL & Reed JC (1999) IAP family proteins – suppressors of apoptosis. *Genes Dev* **13**, 239–252.
- Suzuki Y, Imai Y, Nakayama H, Takahashi K, Takio 'K & Takahashi R (2001) A serine protease, HtrA2, is released from the mitochondria and interacts with XIAP, inducing cell death. *Mol Cell* **8**, 613–621.
- Hegde R, Srinivasula S, Zhang Z, Wassell R, Mukattash R, Cilenti L, DuBois G, Lazebnik Y, Zervos A, Fernandes-Alnemri T *et al.* (2002) Identification of Omi/HtrA2 as a mitochondrial apoptotic serine protease that disrupts inhibitor of apoptosis protein-caspase interaction. *J Biol Chem* **277**, 432–438.
- Martins L, Iaccarino I, Tenev T, Gschmeissner S, Totty N, Lemoine N, Savopoulos J, Gray C, Creasy C, Dingwall C *et al.* (2002) The serine protease Omi/HtrA2 regulates apoptosis by binding XIAP through a reaper-like motif. *J Biol Chem* **277**, 439–444.
- Trencia A, Fiory F, Maitan M, Vito P, Barbagallo A, Perfetti A, Miele C, Ungaro P, Oriente F, Cilenti L *et al.* (2004) Omi/HtrA2 promotes cell death by binding and degrading the anti-apoptotic protein ped/pea-15. *J Biol Chem* **279**, 46566–46572.
- Cilenti L, Soundarapandian M, Kyriazis G, Stratico V, Singh S, Gupta S, Bonventre J, Alnemri ES & Zervos AS (2004) Regulation of HAX-1 anti-apoptotic protein

- by Omi/HtrA2 protease during cell death. *J Biol Chem* **279**, 50295–50301.
- 14 Li W, Srinivasula S, Chai J, Li P, Wu J, Zhang Z, Alnemri E & Shi Y (2002) Structural insights into the pro-apoptotic function of mitochondrial serine protease HtrA2/Omi. *Nat Struct Biol* **9**, 436–441.
 - 15 Bejugam PR, Kuppili RR, Singh N, Gadewal N, Chaganti LK, Sastry GM & Bose K (2013) Allosteric regulation of serine protease HtrA2 through novel non-canonical substrate binding pocket. *PLoS One* **8**, e55416.
 - 16 Chaganti LK, Kuppili RR & Bose K (2013) Intricate structural coordination and domain plasticity regulate activity of serine protease HtrA2. *FASEB J* **27**, 3054–3066.
 - 17 Yang Q, Church-Hajduk R, Ren J, Newton M & Du C (2003) Omi/HtrA2 catalytic cleavage of inhibitor of apoptosis (IAP) irreversibly inactivates IAPs and facilitates caspase activity in apoptosis. *Genes Dev* **17**, 1487–1496.
 - 18 Verhagen AM, Silke J, Ekert PG, Pakusch M, Kaufmann H, Connolly LM, Day CL, Tikoo A, Burke R, Wrobel C *et al.* (2002) HtrA2 promotes cell death through its serine protease activity and its ability to antagonize inhibitor of apoptosis proteins. *J Biol Chem* **277**, 445–454.
 - 19 Savopoulos JW, Carter PS, Turconi S, Pettman GR, Karran EH, Gray CW, Ward RV, Jenkins O & Creasy CL (2000) Expression, purification, and functional analysis of the human serine protease HtrA2. *Protein Expr Purif* **19**, 227–234.
 - 20 Gupta S (2004) The C-terminal tail of presenilin regulates Omi/HtrA2 protease activity. *J Biol Chem* **279**, 45844–45854.
 - 21 Wilkinson JC, Cepero E, Boise LH & Duckett CS (2004) Upstream regulatory role for XIAP in receptor-mediated apoptosis. *Mol Cell Biol* **24**, 7003–7014.
 - 22 Sohn J & Sauer R (2009) OMP peptides modulate the activity of DegS protease by differential binding to active and inactive conformations. *Mol Cell* **33**, 64–74.
 - 23 Mazat J & Patte J (1976) Lysine-sensitive aspartokinase of *Escherichia coli* K12. Synergy and autosynergy in an allosteric V system. *Biochemistry* **15**, 4053–4058.
 - 24 Monod J, Wyman J & Changeux JP (1965) On the nature of allosteric transitions: a plausible model. *J Mol Biol* **12**, 88–118.
 - 25 Sohn J, Grant R & Sauer R (2007) Allosteric activation of DegS, a stress sensor PDZ protease. *Cell* **131**, 572–583.
 - 26 Martins L, Turk B, Cowling V, Borg A, Jarrell E, Cantley L & Downward J (2003) Binding specificity and regulation of the serine protease and PDZ domains of HtrA2/Omi. *J Biol Chem* **278**, 49417–49427.
 - 27 Lee HJ & Zheng JJ (2010) PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun Signal* **8**, 8.
 - 28 Eckelman BP, Drag M, Snipas SJ & Salvesen GS (2008) The mechanism of peptide-binding specificity of IAP BIR domains. *Cell Death Differ* **15**, 920–928.
 - 29 Durrant JD & McCammon JA (2011) Molecular dynamics simulations and drug discovery. *BMC Biol* **9**, 71.
 - 30 Wist AD, Gu L, Riedl SJ, Shi Y & McLendon GL (2007) Structure-activity based study of the Smac-binding pocket within the BIR3 domain of XIAP. *Bioorg Med Chem* **15**, 2935–2943.
 - 31 Sun H, Nikolovska-Coleska Z, Yang CY, Qian D, Lu J, Qiu S, Bai L, Peng Y, Cai Q & Wang S (2008) Design of small-molecule peptidic and nonpeptidic Smac mimetics. *Acc Chem Res* **41**, 1264–1277.
 - 32 Hedstrom L (2002) Serine protease mechanism and specificity. *Chem Rev* **102**, 4501–4524.
 - 33 Hunkapiller MW, Forgacs MD & Richards JH (1976) Mechanism of action of serine proteases: tetrahedral intermediate and concerted proton transfer. *Biochemistry* **15**, 5581–5588.
 - 34 Gray C, Ward R, Karran E, Turconi S, Rowles A, Viglienghi D, Southan C, Barton A, Fantom K, West A *et al.* (2000) Characterization of human HtrA2, a novel serine protease involved in the mammalian cellular stress response. *Eur J Biochem* **267**, 5699–5710.
 - 35 Skorko-Glonek J, Zurawa-Janicka D, Koper T, Jarzab M, Figaj D, Glaza P & Lipinska B (2013) HtrA protease family as therapeutic targets. *Curr Pharm Des* **19**, 977–1009.
 - 36 Huber R & Bode W (1978) Structural basis of the activation and action of trypsin. *Acc Chem Res* **11**, 114–122.
 - 37 Sohn J, Grant RA & Sauer RT (2010) Allostery is an intrinsic property of the protease domain of DegS: implications for enzyme function and evolution. *J Biol Chem* **285**, 34039–34047.
 - 38 Hansen G & Hilgenfeld R (2013) Architecture and regulation of HtrA-family proteins involved in protein quality control and stress response. *Cell Mol Life Sci* **70**, 761–775.
 - 39 Gandhi PS, Chen Z, Mathews FS & Di Cera E (2008) Structural identification of the pathway of long-range communication in an allosteric enzyme. *Proc Natl Acad Sci USA* **105**, 1832–1837.
 - 40 Gunasekaran K, Ma B & Nussinov R (2004) Is allostery an intrinsic property of all dynamic proteins? *Proteins* **57**, 433–443.
 - 41 Krojer T, Pangerl K, Kurt J, Sawa J, Stingl C, Mechtler K, Huber R, Ehrmann M & Clausen T (2008) Interplay of PDZ and protease domain of DegP ensures efficient elimination of misfolded proteins. *Proc Natl Acad Sci USA* **105**, 7702–7707.
 - 42 Mohamedmohaideen N, Palaninathan S, Morin P, Williams B, Braunstein M, Tichy S, Locker J, Russell D, Jacobs WJ & Sacchettini J (2008) Structure and

- function of the virulence-associated high-temperature requirement A of *Mycobacterium tuberculosis*. *Biochemistry* **47**, 6092–6102.
- 43 Kley J, Schmidt B, Boyanov B, Stolt-Bergner P, Kirk R, Ehrmann M, Knopf R, Naveh L, Adam Z & Clausen T (2011) Structural adaptation of the plant protease Deg1 to repair photosystem II during light exposure. *Nat Struct Mol Biol* **18**, 728–731.
- 44 Urwyler S (2011) Allosteric modulation of family C G-protein-coupled receptors: from molecular insights to therapeutic perspectives. *Pharmacol Rev* **63**, 59–126.
- 45 May L, Leach K, Sexton P & Christopoulos A (2007) Allosteric modulation of G protein-coupled receptors. *Annu Rev Pharmacol Toxicol* **47**, 1–51.
- 46 Zhu ZH, Yu YP, Zheng ZL, Song Y, Xiang GS, Nelson J, Michalopoulos G & Luo JH (2010) Integrin alpha 7 interacts with high temperature requirement A2 (HtrA2) to induce prostate cancer cell death. *Am J Pathol* **177**, 1176–1186.
- 47 Ariyaratana S & Loeb DM (2007) The role of the Wilms tumour gene (WT1) in normal and malignant haematopoiesis. *Expert Rev Mol Med* **9**, 1–17.
- 48 Hartkamp J & Roberts SG (2010) HtrA2, taming the oncogenic activities of WT1. *Cell Cycle* **9**, 2508–2514.
- 49 Pozzi N, Chen R, Chen Z, Bah A & Di Cera E (2011) Rigidification of the autolysis loop enhances Na⁽⁺⁾ binding to thrombin. *Biophys Chem* **159**, 6–13.
- 50 Wolan DW, Zorn JA, Gray DC & Wells JA (2009) Small-molecule activators of a proenzyme. *Science* **326**, 853–858.
- 51 Nallamsetty S, Kapust R, Tözsér J, Cherry S, Tropea J, Copeland T & Waugh D (2004) Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease *in vivo* and *in vitro*. *Protein Expr Purif* **38**, 108–115.
- 52 Himoe A, Brandt KG, DeSa RJ & Hess GP (1969) Investigations of the chymotrypsin-catalyzed hydrolysis of specific substrates. IV. Pre-steady state kinetic approaches to the investigation of the catalytic hydrolysis of esters. *J Biol Chem* **244**, 3483–3493.
- 53 Bowers K, Chow E, Xu H, Ron O, Eastwood M, Gregersen B, Klepeis J, Kolossváry I, Moraes M, Sacerdoti F *et al.* (2006) Scalable Algorithms for Molecular Dynamics Simulations for Commodity Clusters. In *Proceedings of the ACM/IEEE Conference on Supercomputing (SC06)*, Florida.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Fig. S1. XIAP-mediated activation of HtrA2.

Fig. S2. Determination of steady-state kinetics for β -casein cleavage by HtrA2 and HtrA2 ^{Δ PDZ} in the presence of IBM peptides.

Fig. S3. Graphical representation of fluctuations in peptide bound and unbound form of HtrA2.

Table S1. Comparison of distances between the atoms of the catalytic triad in the BIR3_{pep}-bound and unbound forms of HtrA2 (Å).

Allosteric Regulation of Serine Protease HtrA2 through Novel Non-Canonical Substrate Binding Pocket

Pruthvi Raj Bejugam¹*, Raja R. Kuppili¹*, Nitu Singh¹, Nikhil Gadewal¹, Lalith K. Chaganti¹, G. Madhavi Sastry², Kakoli Bose^{1*}

1 Advanced Centre for Treatment, Research and Education in Cancer (ACTREC), Tata Memorial Centre, Kharghar, Navi Mumbai, India, **2** Schrödinger, Sanali Infopark, Banjara Hills, Hyderabad, India

Abstract

HtrA2, a trimeric proapoptotic serine protease is involved in several diseases including cancer and neurodegenerative disorders. Its unique ability to mediate apoptosis via multiple pathways makes it an important therapeutic target. In HtrA2, C-terminal PDZ domain upon substrate binding regulates its functions through coordinated conformational changes the mechanism of which is yet to be elucidated. Although allostery has been found in some of its homologs, it has not been characterized in HtrA2 so far. Here, with an *in silico* and biochemical approach we have shown that allostery does regulate HtrA2 activity. Our studies identified a novel non-canonical selective binding pocket in HtrA2 which initiates signal propagation to the distal active site through a complex allosteric mechanism. This non-classical binding pocket is unique among HtrA family proteins and thus unfolds a novel mechanism of regulation of HtrA2 activity and hence apoptosis.

Citation: Bejugam PR, Kuppili RR, Singh N, Gadewal N, Chaganti LK, et al. (2013) Allosteric Regulation of Serine Protease HtrA2 through Novel Non-Canonical Substrate Binding Pocket. PLoS ONE 8(2): e55416. doi:10.1371/journal.pone.0055416

Editor: Srinivasa M. Srinivasula, IISER-TVM, India

Received: September 6, 2012; **Accepted:** December 22, 2012; **Published:** February 14, 2013

Copyright: © 2013 Bejugam et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The project was funded by a grant provided by Department of Biotechnology (DBT), Govt. of India. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: GMS is employed by Schrodinger Inc. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials. The authors declare that they do not have any conflict of interests.

* E-mail: kbose@actrec.gov.in

† These authors contributed equally to this work.

Introduction

Multidomain proteins due to their structural complexity require different levels of regulatory mechanisms for executing cellular functions efficiently within a specified time period. Allosteric modulation of conformations is one such mechanism which often helps a protein to regulate a functional behaviour such as for an enzyme to attain an active functional state upon ligand or substrate binding. In allostery, sometimes there are large conformational changes that require significant rotations and translations of individual domains at the timescales of microsecond to millisecond. While in some other cases, minimal structural perturbation helps in propagation of the signal in an energy efficient way to the functional domain where movement is mainly restricted to the side chains, loops and linker regions and which occur within picosecond to nanosecond timescales [1]. PDZ (post-synaptic density-95/discs large/zonula occludens-1) domains that are involved in myriads of protein-protein interactions [2,3] exhibit minimal structural changes during allosteric propagation. These domains have multiple ligand docking sites and are known to possess unique dynamics that regulate conformation of the functional site from a distal region.

HtrA2 (High temperature requirement protease A2), a PDZ bearing protein, is a mitochondrial trimeric pyramidal proapoptotic serine protease with complex domain architecture whose activity is likely regulated by interdomain crosstalk and structural plasticity [4]. Mature HtrA2 comprises 325 amino acids with residues S173, D95 and H65 forming the catalytic triad which is

buried 25 Å above the base of the pyramid suggesting requirement of conformational changes for its activation. Apart from PDZ, this multidomain protein has a short N-terminal region, a serine protease domain and a non-conserved flexible linker at the PDZ-protease interface [4]. HtrA2 is involved in both caspase dependent as well as caspase independent apoptotic pathways [5,6,7]. Literature suggests it might have chaperoning functions as well and recently has been found to be associated with several neurodegenerative disorders [8,9,10]. Based on information from literature [4,11], this multitasking ability of HtrA2 can be attributed to its serine protease activity which is intricately coordinated by its unique substrate binding process, complex trimeric structure, interdomain networking and conformational plasticity. However, the unbound inactive form of the crystal structure with partially missing active site loops and flexible PDZ-protease linker has been unable to unambiguously determine the role of dynamics and allostery if any in HtrA2 activation and specificity. Therefore, to understand the molecular details of its mechanism of action, dynamics study at the substrate binding site and active site pocket becomes imperative.

HtrA2 belongs to a serine protease family that is conserved from prokaryotes to humans [12] where allostery is a common mechanism for protease activation in some of its homologs. DegS, a bacterial counterpart of HtrA2, allosterically stabilizes the active site pocket upon substrate binding at the distal PDZ domain [13]. DegP, the most extensively studied protein of the family, has a cage-like hexameric structure whose activation is regulated by allostery and oligomerization. Peptide binding to distal PDZ1

domain leads to rearrangement of the catalytic pocket into enzymatically competent form that readily oligomerizes and renders stability to the active conformation [14].

With an aim at understanding the conformational changes and structural plasticity that govern HtrA2 activity and specificity, we took an *in silico* approach to study the movements of flexible regions of the protein upon ligand binding. The PDZ domain of HtrA2 has a known hydrophobic substrate binding YIGV pocket (similar to GLGF motif) which is deeply embedded within the trimeric protein structure with P225 and V226 from the serine protease domain occupying the groove [4,15]. This structural arrangement makes it impossible for substrate protein to bind without significant conformational changes. Thus, to examine whether allosteric modulation through an alternative site is involved in substrate binding and catalysis of HtrA2, molecular dynamics simulation (MDS) approach with a bound peptide activator was used to look into the structural rearrangements that occur in nanosecond time scale. Although the information usually obtained from MDS is restricted primarily to movements in the accessible and flexible regions of a protein, it nonetheless contributes significantly towards understanding of the overall structural rearrangement and dynamics during its allosteric activation. In our study, we modelled the entire mature protease by filling in the missing regions using Prime 3.0 [16], followed by energy minimization with GRoningen MACHine for Chemical Simulation or GROMACS [17]. Identification of the putative binding site(s) on HtrA2 was done using SiteMap 2.5 [18] and the selective binding pocket (SBP) for the ligand was chosen based on optimum energy parameters. Peptides at SBP were docked from our peptide library that was generated based on available literature reports [19,20,21] and structural complementarities. MDS of the docked structures was done using Desmond 2010 [22] which provided critical information on loop and linker movements in HtrA2. These results combined with mutational and enzymology studies show that upon activator binding at the novel allosteric pocket, SBP, the linker at the PDZ-protease interface and loops L1, LA and LD around the catalytic groove undergo rearrangements in a coordinated manner so as to form an efficient active site pocket. Moreover, the PDZ domains mediate intersubunit interactions which stabilize the oxyanion hole. These observations highlight the importance of allostery which might be an important prerequisite for an active conformation of the trimeric protease.

Results

Identification of Selective Binding Pocket (SBP)

The high resolution crystal structure of HtrA2 [4] (Figure 1a) that lacked flexible loops, linkers and some N-terminal residues was the target protein for our studies. These regions were modelled and energy minimised as described under Methods section. Comparison of refined model with unrefined structure showed significant movements of the loops defining new binding sites on the protein surface. The linker at SPD-PDZ interface moved towards $\alpha 7$ of PDZ domain whereas the linker in the protease domain moved closer to the SPD-PDZ linker so as to form a groove (Figure 1b).

Among the five possible putative binding sites that were identified, Site2 or SBP (Figure 1c) that encompasses the groove generated by SPD-PDZ linker, protease and PDZ domains attained the best score (Table 1). The site score takes into account parameters such as volume, density, solvent exposure, hydrophilic and hydrophobic nature of residues and donor to acceptor ratio and hence is a comprehensive representation of the possibility of it being a binding site.

SBP has optimum volume and contacts available including maximum hydrogen donor and acceptor groups that are crucial for interacting with peptides. The size of the site is very important since the binding peptides have 6–7 residues and the site needs to be large enough to accommodate them. It also has highest hydrophobicity which makes it the best interaction site and hence used in our studies. Although sites 1 and 3 have scores closer to that of SBP, taking into account all the above-mentioned parameters, SBP was chosen for further docking and MDS studies.

Peptide Docking Show Similar Interacting Residues

Here, we have used a holistic approach in designing activator peptides where different techniques were applied in parallel so as to conduct a comprehensive search for a signature pattern that would dock at SBP. In one method, replicas for functional groups were chosen based on sequence and structural complementarities with hydrophobic SBP which were used for generating small molecular fragments. Scores obtained from docking these small molecules (Table S1) provided the framework for designing different combinations of tetrapeptides as shown in Table S2. With leads from literature and *in silico* structure-guided design, Gly and Val residues were added at N- and C-termini respectively of some peptides which subsequently increased the docking scores from -6 to -10 kcal/mol.

Similarly, two peptides previously reported in the literature as well peptides designed from the putative binding sites in pea-15 and Hax-1 also interacted well with SBP. Analysis of docking results with all these different peptides show interaction with similar residues of SBP as observed in ligplot (Figure S1). However, the control peptide KNNPNNAHQN, which has quite a few asparagine residues, is an ideal sequence to act as negative peptide for the pocket due to its stereochemical properties [19], did not bind to SBP demonstrating the specificity of designed peptides.

From the above extensive docking analysis, N216, S217, S219, E292 and E296 in SBP were found to be common for most of the peptide interactions (Figures 2a–b). Of these residues, N216, S217, S219 belong to the linker region while E292 and E296 to the PDZ domain that were either involved in hydrogen bond formation or Van der Waals interaction with the peptides. This result suggests that SBP might be the possible binding site and therefore a prospective putative allosteric site.

The role of some of these important residues in allostery if any and its subsequent effect on catalytic activity and substrate turnover was further probed by enzymology studies as described later in the text.

MDS Analyses of HtrA2 and HtrA2– Peptide Complexes

The peptides GSAWFSF was chosen for MDS studies as it gave the best XP and E-model scores (Table 2). GQYYFV has been reported to be a well known activator of HtrA2 [19] and hence used as another representative peptide for simulation studies. Moreover, the two peptides were chosen such that one is a designed peptide (GQYYFV) while the other is a part of a well-known HtrA2 binding protein Pea-15 (GSAWFSF). In addition to this, GQYYFV with docking score lesser than GSAWFSF was chosen for MDS analysis to understand whether different affinity for the substrate results in similar movements in the protease. MDS analyses of HtrA2-GQYYFV and HtrA2-GSAWFSF complexes demonstrated significant difference in conformation as well as dynamics when compared with unbound HtrA2. Visual inspection of the domain wise movements in peptide bound HtrA2 indicated large fluctuations in hinge/linker region (211–226) as shown in Figures 3a and b. Although these

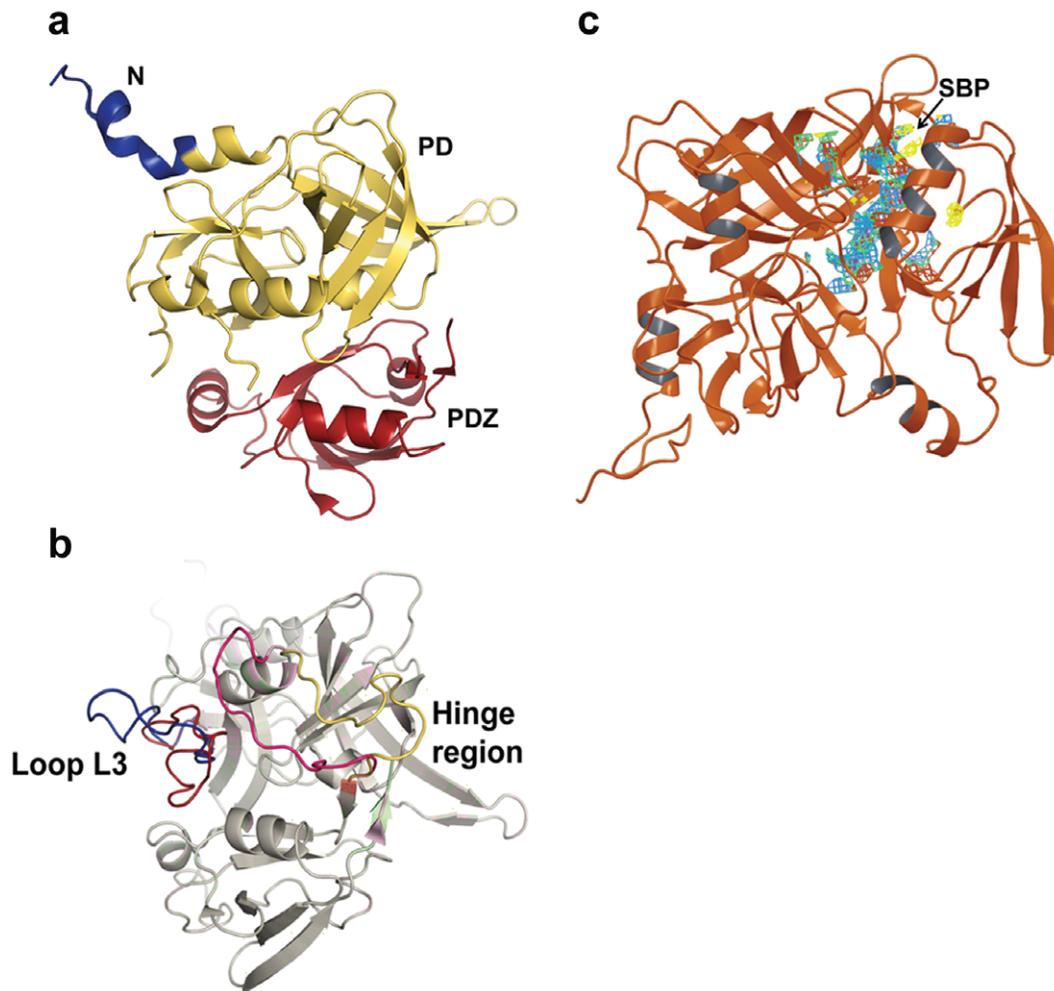


Figure 1. Ribbon model of HtrA2 structures (PDB ID: 1LCY). a. Domain organization of HtrA2 protease which comprises N-terminal region (blue), protease domain denoted as PD (yellow) and PDZ domain (red) at C-terminal end. b. Structural alignment of loop refined (light magenta) and unrefined (light green) structures of HtrA2 protein with modelled N-terminal AVPS, loop L3 (residues 142–162) and hinge region (residues 211–225) built with Prime (Schrödinger 2011). On refinement, loop L3 and hinge region are reorganized so as to define new regions at the protease and PDZ domain interface. c. Selective binding pocket (SBP) on HtrA2. The energy minimised structure of HtrA2 after modelling flexible regions in the protein is represented as a ribbon model. The binding site designated as SBP, selected on the basis of the Sitemap score and residue analyses, is located at the interface of PDZ and protease domain and shown as a multi-coloured mesh.
doi:10.1371/journal.pone.0055416.g001

movements were larger for GSAWFSF than GQYYFV bound complex, the movement pattern remained similar in these two peptides. Enhanced dynamic movement in the former complex could be attributed to the peptide length (heptameric as compared

to hexameric in the latter). Domain wise RMSD analysis of these trajectories provided quantitative output of deviations with respect to time. The trajectory graphs (Figures 3c–e) show that along the entire sequence, hinge region (211 –226) has RMSD of 2.5 Å for

Table 1. Putative binding sites in HtrA2 identified by SiteMap tool.

Site Number from SiteMap	Residues present in the site	Site score
Site 2	K214, K215, N216,S217,S219, R226, R227, Y228, I229, G230,V231,M232,M233, L234, T235, L236, S237, S239, I240, E243, H256, K262, I264,Q289, N290, A291,E292, Y295,E 296, R299, S302	1.092716
Site 1	H65, D69, R71, A89, V90, P92, D95,T324	0.957142
Site 3	N48, H65, D169, S173,K191, M232, H261,L265	0.936056
Site 4	V192, F251	0.807891
Site 5	I33,L34,D35,R36,V73,R74	0.673032

doi:10.1371/journal.pone.0055416.t001

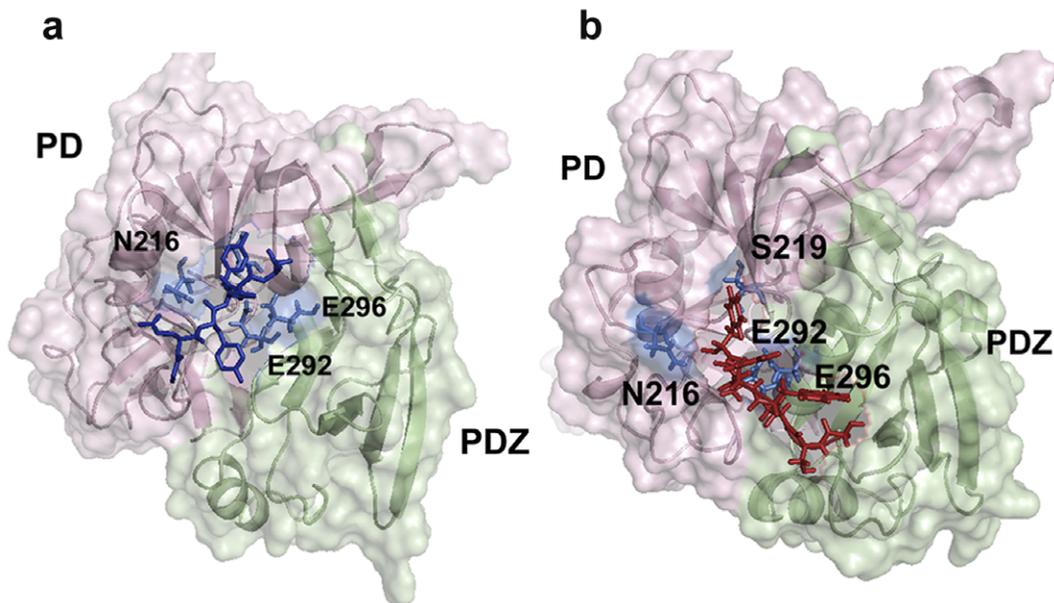


Figure 2. Representative surface structures of peptide activator docked HtrA2. a. Peptide GSAWFSF-HtrA2 complex and b. Peptide GQYYFV-HtrA2 complex. The former peptide represents putative SBP binding peptide in Pea-15 and the latter is a peptide obtained from the literature. The common interacting residues from SBP for both the peptides are labelled and are shown as blue sticks. PD denotes serine protease domain in both the Figures.

doi:10.1371/journal.pone.0055416.g002

the peptide GSAWFSF and 1.5 Å for GQYYFV from the starting unbound form.

The RMSF of these trajectories were comparable with rmsd values showing higher relative fluctuations in and around the hinge region. Representative RMSF plots for GQYYFV and GSAWFSF bound HtrA2 complexes depict these large fluctuations for residues 190–225 as shown in Figures 4b and C respectively. All structural alignment comparisons and relative fluctuation analyses post MDS emphasize distinct significant conformational change in the hinge (211–226) region upon peptide binding. In addition to this, binding of peptides led to dynamic movements in many functionally important regions distal to SBP such as helices $\alpha 5$ and $\alpha 7$ in PDZ domain.

Conformational Transitions in Flexible Regions and at the Active Site

Further detailed analyses of the effect that local subtle structural changes at SBP had on distal regions of the protease especially at the active site and its vicinity revealed the possibility of SBP being a putative allosteric site. Functional active site formation and its accessibility along with a well formed oxyanion hole are important prerequisites for the activity of an enzyme.

Structural comparison of the MD simulated peptide bound structure of HtrA2 with the unbound form show movements in different domains and linker regions. The PDZ-protease linker that covers the peptide binding groove in the PDZ domain moves away from it thus increasing its accessibility. The peptide bound HtrA2 complex show relative movements in the active site triad residues compared to the unbound form. Atomic distance analysis of both the forms revealed that distances between nitrogen (ϵ) atom of H65 and oxygen (γ) atom of S173 increased in peptide bound complexes while that between nitrogen (δ) atom of H65 and oxygen (δ) of D95 decreased when compared with the unbound HtrA2 structure (Table 3). This pattern being consistent with both

the peptides suggests that interaction of peptide activator with SBP leads to opening up of the active site cleft.

Apart from active site triad, changes were also observed in the orientation of mechanically important L1, LD and LA loops in the peptide bound complex (Figures 4d–e). Their orientations with respect to the active site determine proper oxyanion hole formation, accessibility of the active site, formation of catalytic triad and hence enzyme activity. MDS analyses for these regions showed significant deviations upon peptide binding. Structural alignment of GSAWFSF bound HtrA2 complex with the unbound form demonstrated breaking of Van der Waals contacts between loop LD and $\beta 2$ strand of protease domain which facilitates LD movement towards $\alpha 1$ of protease domain and bringing P130 of the former in proximity to A25 of the latter. Similarly, S50 in $\beta 2$ of protease domain establishes interactions with G171 of L1 (oxyanion hole residue) while breaking contacts with A132 of LD loop due to movement or tilt in the L1 loop. As a result of this reorganization, LD which was closer to L1 in the unbound HtrA2 moves sharply away from it upon peptide binding. These positional rearrangements also lead to disruption of interaction between D165 of L1 and G195 of L2 loops. All these movements coordinate to bring LD closer to the proximal region of protease domain thereby opening up the catalytic site. For GQYYFV peptide, movements of all these loops were subtle as compared to that for GSAWFSF except for the LA loop which exhibited larger deviation in the former. The other significant flexible region movement is in loop L3 which, in concert with linker region, assists in accommodating the peptide at SBP.

The relative reorientation of these loops along with catalytic triad residues seems to be assisting formation of a more open structure near the active site. However, loop L2 that harbors the specificity pocket remains mostly unchanged suggesting presence of a well formed binding pocket in the unbound form whose accessibility is limited compared to the substrate bound form. In context with trimeric HtrA2, more open conformation might be

Table 2. Peptide docking of HtrA2 and identification of interacting residues.

Peptides Used in Our study	Interacting Residues		Glide score in Kcal.mole ⁻¹
	H bond Interactions	Vdw Interactions	
PEA 15 (GSAWFSF)	Glu 292, Glu 296, Asp 293, Ile 283, Met 287	Gln 286, Ala 297, Ser 222	-10.564
Designed (VKSDSG)	Asn 216, Leu 152, Glu 296, Glu 292	Ala 89, Ile 221, ser 218,	-10.394
Designed (GRTDSV)	Glu 296, Glu 292, Asn 216, Ser 217	Asp 293	-10.037
Designed (GRDTSV)	Ser 219, Glu 292	Ser 239, Gln 286	-9.57
Designed (GRDTYV)	Asp 293, Asn 216, Ser 217, Ser 219	Glu 296, Arg 299,	-9.54
Phosphatase (PAEWTRY)	Asp 117, Ala 149, Arg 150, Lys 215, Gln 146	Pro 148, Leu 152, Lys 214, Gln 156, Val 159, Ser 239	-9.481
HAX-1 (TKPDIGV)	Glu 292, Glu 296, Ser 219, Ile 221, Arg 299	Asn 216, Ser 222	-8.486
Connexin (ARKSEWV)	Asp 293/426, Asn 290/423, Gln 156/289	Glu 292, Pro 155, Gln 289, Met 287, His 256, Glu 255, pro 238	-8.165
Presenilin (AFHQFYI)	Leu 152, Asn 216, Ser 217, Glu 292, Glu 296	Pro 155, Arg 211, ser 218, Ser 219	-8.063
IL-EBF (AGYTGfV)	Asn 216, Ser 217, Glu 292, Arg 150/, Leu 152	Ser 219, Gly 153, pro 155	-7.903
Yes Protein (ESFLTWL)	Asn 216, Leu 152, Glu 296, Asp 293, Gln 289, Ser 237	Gln 156, Pro 238, Pro 155, Ser 218, Glu 292, Gln 286	-7.722
Cathepsin SVSSIFV	Glu 296, Asn 216, Ile 283/416,	Glu 292, Leu 152, Gly 153, Ala 297	-7.524
Warts Protein Kinase (NRDLVYV)	Lys 214, Lys215, Ala 149, Glu 207, Arg150, Gln 146	Leu 152, Gln 156, Val 159	-7.321
GQYYFV⁶	Glu 292, Glu 296, Asn 216, Ile 221, Leu 152	Ser 219, Gly 153, Arg 299	-7.163
GGIRRV⁶	Glu 292, Glu 296, Asn 216, Ser 217, Ser 219	Arg 211, Gly 153	-6.785
Tuberin (EDFTEFV)	Arg 211, Asn 216, Ser 219	Ala 89, Ile 221, ser 218, Arg 299, Glu 296, Glu 292, Gly 153	-1.883
Control Peptide (KNNPNNAHQN)	Did not dock with HtrA2		

The possible residues which are involved in hydrogen bonding and Vander Waal's interactions along with Glide scores are mentioned.
doi:10.1371/journal.pone.0055416.t002

significant as it enhances the accessibility of the substrate and thereby might contribute positively toward the rate of enzyme catalysis.

Influence of SBP on HtrA2 Activity and Role of PDZ Domain

To determine whether critical SBP residues (N216, S219, E292 and E296) are important for mediating allosteric propagation in HtrA2, site directed mutagenesis to alanine were done. Mutation of a conserved YIGV residue (G230A) was also done to understand the role of canonical YIGV groove in this complex signal propagation pathway. Moreover, since the protein is found to be active in its trimeric form [4] and also that SBP encompasses a major part of PDZ, we used trimeric and monomeric HtrA2 variants, N-SPD and F16D respectively to understand the role of PDZ in intra and inter-molecular cross-talk.

To negate the role of overall conformational changes if any due to these mutations, MDS and secondary structural analyses were done on the mutant proteins. Similar active site conformations were observed in both the wildtype and mutants. Moreover, the overall secondary structure and thermal stability remained unperturbed due to the mutations (data not shown). Enzymology studies with different SBP mutants were done using β -casein, a well-established generic substrate of serine proteases [23]. β -casein has a putative SBP binding site (GPFPIIV) which has been found to interact with the similar residues at SBP by our docking studies (Table 2) and hence expected to mimic the allosteric modulation mediated by SBP binding if any. The kinetic

parameters for wild type, N-SPD domain, F16D and other mutants were determined using fluorescent β -casein (Figure 5). The catalytic efficiency (k_{cat}/K_m) for the double mutant N216A/S219A and single mutant E292A showed ~ 2.4 fold decrease in enzyme activity as compared to wild type whereas enzymatic parameters remained mostly unchanged for E296A. K_m values for the mutants were not significantly higher compared to the wild type, suggesting that the specificity pocket might be mostly intact with some subtle alterations. However, there was a marked decrease in V_{max} and in substrate turnover (k_{cat}) rates for N216A/S219A and E292A suggesting presence of a malformed oxyanion hole in the SBP mutants. These results demonstrate that N216/S219 and E292 of SBP are important for mediating allosteric activation of HtrA2 upon activator binding. This is strengthened by the observation that SBP mutants did not interact with the activating peptides as seen by isothermal calorimetric studies and a representative figure is shown in the supplementary material (Figure S3). In addition, the ligplot of the peptide showing the detailed interaction with HtrA2 is also depicted in figure S1.

In our in silico studies, YIGV has been found to be a part of the greater SBP mesh (Table 1) and since docking with small molecular fragments (~ 35 – 100 Da) showed direct binding with YIGV residues (Table S1), we wanted to understand the effect of YIGV mutation on HtrA2 activity as well. Enzymology studies with G230A demonstrated increase in K_m value compared to the wild type highlighting the involvement of YIGV in this intricate allosteric mechanism. Protein turnover rate was also much lower in G230A as compared to the wild type reiterating the importance of oxyanion hole formation upon activator binding at SBP. Thus,

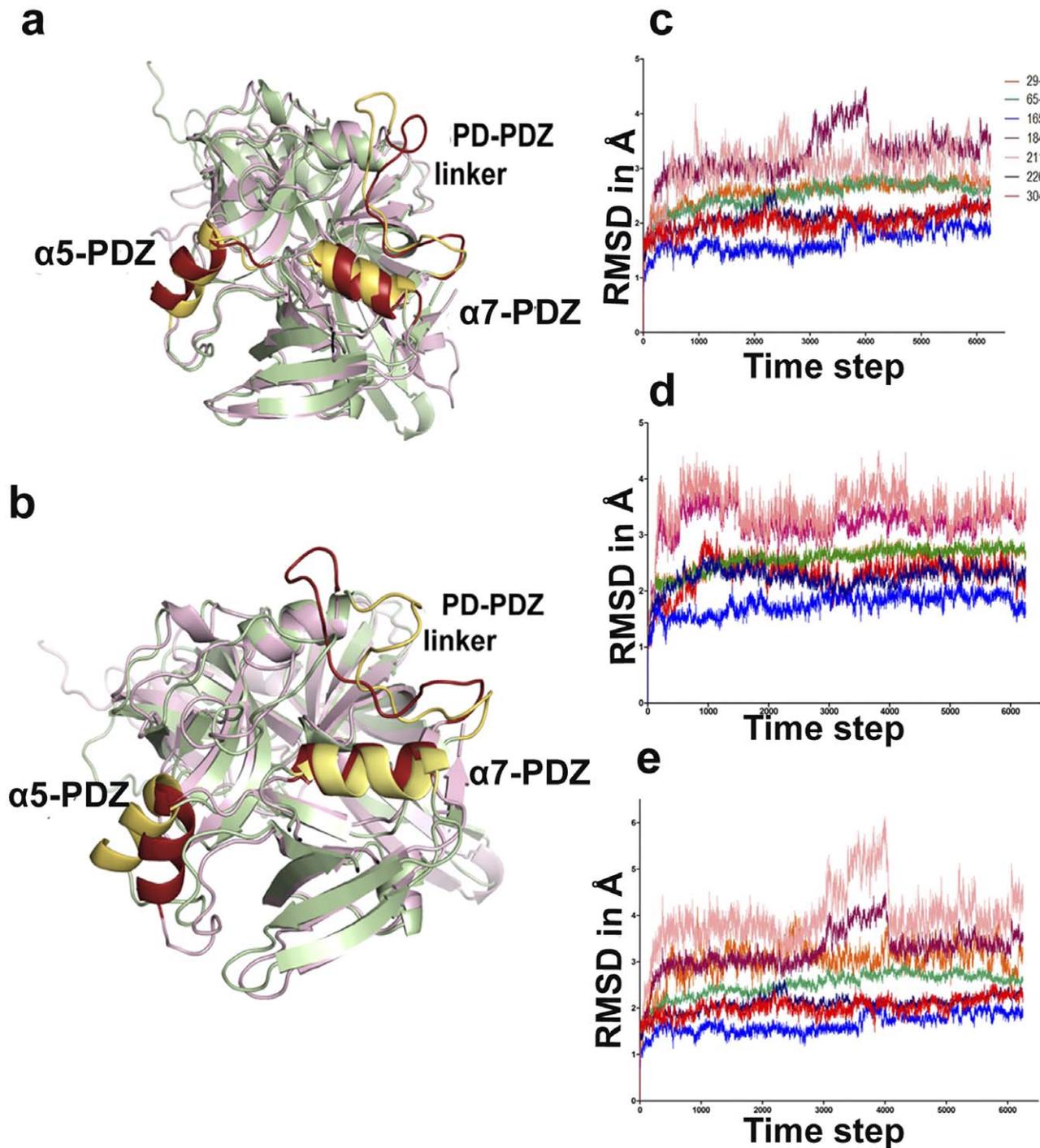


Figure 3. Domain wise conformational changes induced on peptide binding at SBP. a. The structural alignment of minimum energy structure of the peptide bound GQYYFV-HtrA2 complex (light pink) and unbound structure (green) displays orientation of the movement of the hinge region and the α -helices of PDZ. b. The structural alignment of GSAWFSF-HtrA2 complex (light pink) and unbound structure (green). Graphical representations of the RMSD for the 30 ns MDS trajectory of the following: c. HtrA2-GQYYFV complex. d. unbound HtrA2 (negative control). e. HtrA2-GSAWFSF complex. The stretch of residues selected for each set of RMSD calculations are shown on the right of panel c. doi:10.1371/journal.pone.0055416.g003

inaccessibility of the canonical PDZ binding pocket YIGV, in the trimeric protease structure might have adjoined presence of exposed SBP which is dynamically coupled to YIGV groove for efficient allosteric signal propagation to the distal active site. Direct binding of small molecules at YIGV supports this hypothesis as they could be accommodated in the classical binding groove

without requirement of any initial conformational change as it might be with the larger peptide activators.

Interestingly, although catalytic efficiency for N-SPD has been found to be 3.4 fold less as compared to the wildtype, its K_m value suggests slight increase in substrate affinity for the enzyme (Table 4). This increase in substrate affinity might be due to

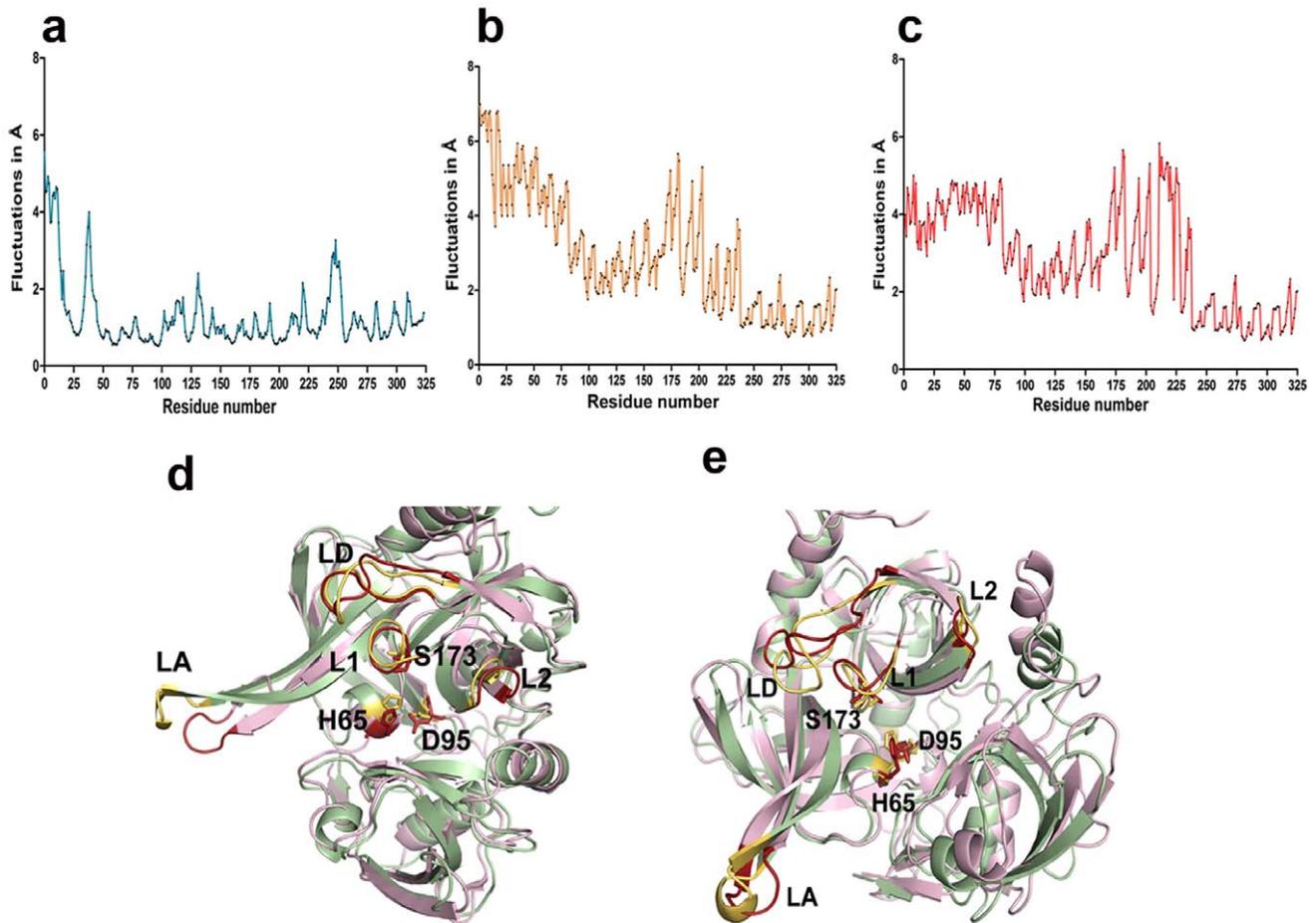


Figure 4. Graphical representation of root mean square fluctuation (RMSF) and loop movements upon peptide binding. a. MD simulation trajectory for unbound HtrA2. b. RMSF graph for GQYYFV bound HtrA2. c. RMSF graph for GSAWFSF bound HtrA2. d. Comparison of fluctuations in loops LA, L1, L2 and LD in the GQYYFV peptide bound (pink) and unbound structure (green). The loops in the bound and unbound forms are displayed in red and yellow respectively. e. Comparison of fluctuations in loops LA, L1, L2 and LD in the GSAWFSF peptide bound (pink) and unbound structure (green). The loops in the bound and unbound forms are displayed in red and yellow respectively. The catalytic triad residues are shown in both panels d. and e.
doi:10.1371/journal.pone.0055416.g004

absence of PDZ surrounding the active site region resulting in greater substrate accessibility. However in N-SPD, k_{cat} was found to be 5 fold less than that of wild type highlighting the role of PDZ in initiating conformational changes near the active site pocket as well as in the oxyanion hole so as to increase overall enzyme stability. However, in the full length monomeric mutant of HtrA2 (F16D), there is a two fold increase in K_m with significant decrease in turnover rate and hence catalytic efficiency (Table 4) which

emphasizes importance of intermolecular crosstalk between PDZ and protease domains in trimeric HtrA2 structure.

The importance of intermolecular interaction between PDZ* and SPD has also been manifested in our MD studies where structural analyses show binding of peptide activator (GQYYFV) at the SBP alters PDZ orientation and brings $\alpha 5$ helix of PDZ from one subunit in close proximity to the protease domain of the adjacent subunit. The helix moves towards LD loop of the protease domain, thereby shifting the orientation of the phenyl ring of F170 which is a part of oxyanion hole towards H65 of the catalytic triad (Figure 6a) so as to accommodate the loop. These rearrangements result in a more stable and catalytically competent HtrA2 formation with a proper oxyanion hole. Thus the full length trimeric HtrA2 is more active than trimeric N-SPD, where the activation pocket is not stable in absence of PDZ.

Table 3. Comparison of distances between atoms of the catalytic triad in the peptide bound and unbound forms of HtrA2.

Protein Complex	NE2 (His) – OG (Ser)		ND1 (His) – OD1 (Asp)	
	Bound	Unbound	Bound	Unbound
HtrA2 (GSAWFSF)5.2		4.1	2.6	2.9
HtrA2 (GQYYFV)	5.5	4.1	2.7	2.9

doi:10.1371/journal.pone.0055416.t003

Discussion

Our aim was to understand the structural dynamics that regulates activation and specificity of HtrA2. This multidomain trimeric protease has unique proapoptotic properties as it is associated with both caspase-dependent and independent cell

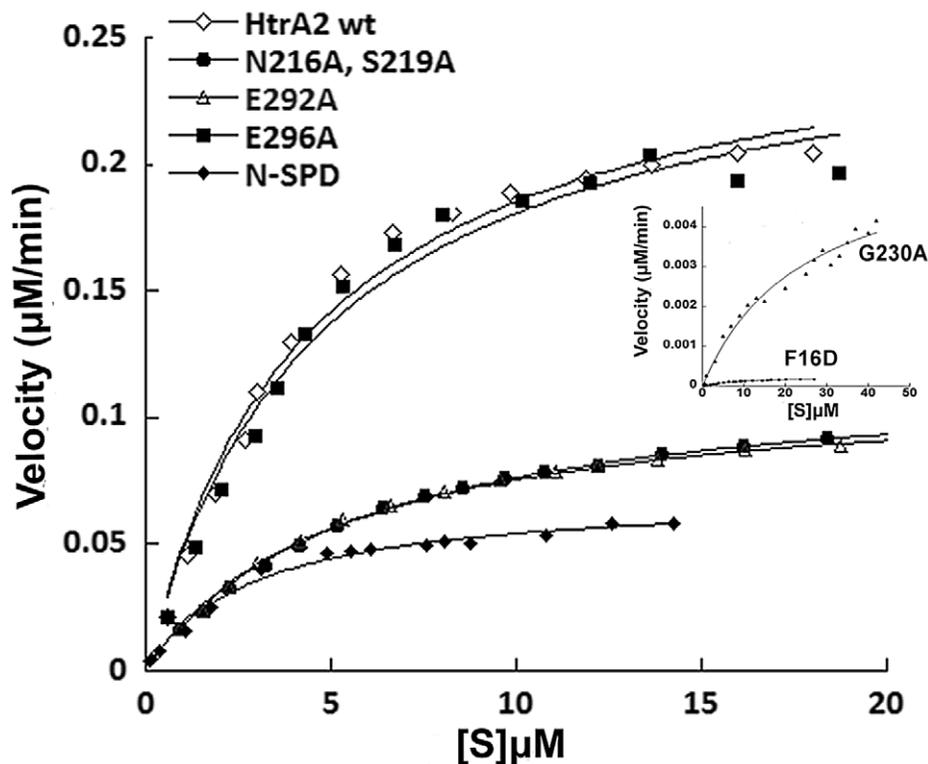


Figure 5. Steady state kinetic parameters of HtrA2. Graph representing relative activity of wild type HtrA2 and its mutants and variants with FITC labelled β -casein as the substrate. The graph for two mutants (F16D and G230A) is shown in inset. doi:10.1371/journal.pone.0055416.g005

death pathways through its serine protease activity [5,12]. Association of HtrA2 with cancer and neurodegenerative disorders makes it a promising therapeutic target. For example, over-expression of HtrA2 substrates such as IAPs and the Wilms's tumor suppressor protein WT1 in several cancers suggests modulation of HtrA2 protease activity can effectively regulate their relative levels in the cells [24,25,26,27]. Out of several approaches that can be used to regulate HtrA2 activity, allosteric modulation is one of the simplest and most efficient ways. However, modulating HtrA2 functions with desired characteristics for disease intervention will require a detailed understanding of its mode of activation and the underlying conformational plasticity that controls it.

Peptide design using site complementarity followed by MDS of the docked peptide-macromolecular complex is an extremely useful tool to study subtle conformational changes and protein dynamics. HtrA2 has a complex network of flexible loops surrounding the active site pocket and a linker at the PDZ-protease interface whose relative orientations and crosstalk with different domains might be critical in defining HtrA2 functions. With partially missing loops and the flexible linker region, the solved structure of HtrA2 [4] could not fully explain the dynamics and allostery that regulate its activity and specificity. Here, with an *in silico* and biochemical approach, we have shown that like few other HtrA family proteins, allosteric propagation does regulate HtrA2 activity.

In this study, peptide binding to SBP showed conformational changes in the distal flexible regions of HtrA2 such as the PDZ-protease interface, loops L1, LD and LA that rearrange to form a more catalytically efficient active site thus establishing the role of SBP as an allosteric site in HtrA2. A close look at and around the active site pocket shows that in the bound form, the N atom of Gly (-2 position) faces the oxyanion hole to form an H-bond whereas in the unbound form it flips in the opposite direction to form a malformed oxyanion hole [12,28]. Moreover, keeping in trend with other HtrA proteases, the phenylalanine ring of -3 position moves closer to the imidazole ring of His65 while in the unbound form, it moves outward as observed from Figures 6b-c and Movie S1. All these subtle structural rearrangements along with making and breaking of bonds at sites away from the active site might stabilize the peptide bound form such that it shifts the equilibrium toward catalysis.

Enzymology studies with β -casein that has a putative SBP binding sequence (GPFPIIV) as shown in Table 4 show significant

Table 4. Steady state kinetic parameters for HtrA2 wild type, variants and mutants with β -casein as the substrate.

HtrA2 Proteins	K_m (μM)	V_{max} (M/s)	k_{cat} (1/s)	k_{cat}/K_m (1/M.s)
Wild type	4.59	4.083×10^{-9}	0.02041	4.452×10^3
N216A, S219A	5.43	1.937×10^{-9}	0.00968	1.788×10^3
E292A	5.15	1.903×10^{-9}	0.00951	1.849×10^3
E296A	4.68	3.734×10^{-9}	0.01868	3.995×10^3
N-SPD	3.02	0.7851×10^{-9}	0.0039	1.29×10^3
F16D	9.3	4.08×10^{-12}	0.000025	0.0026×10^3
G230A	9.32	1.03×10^{-9}	0.0051	0.54×10^3

doi:10.1371/journal.pone.0055416.t004

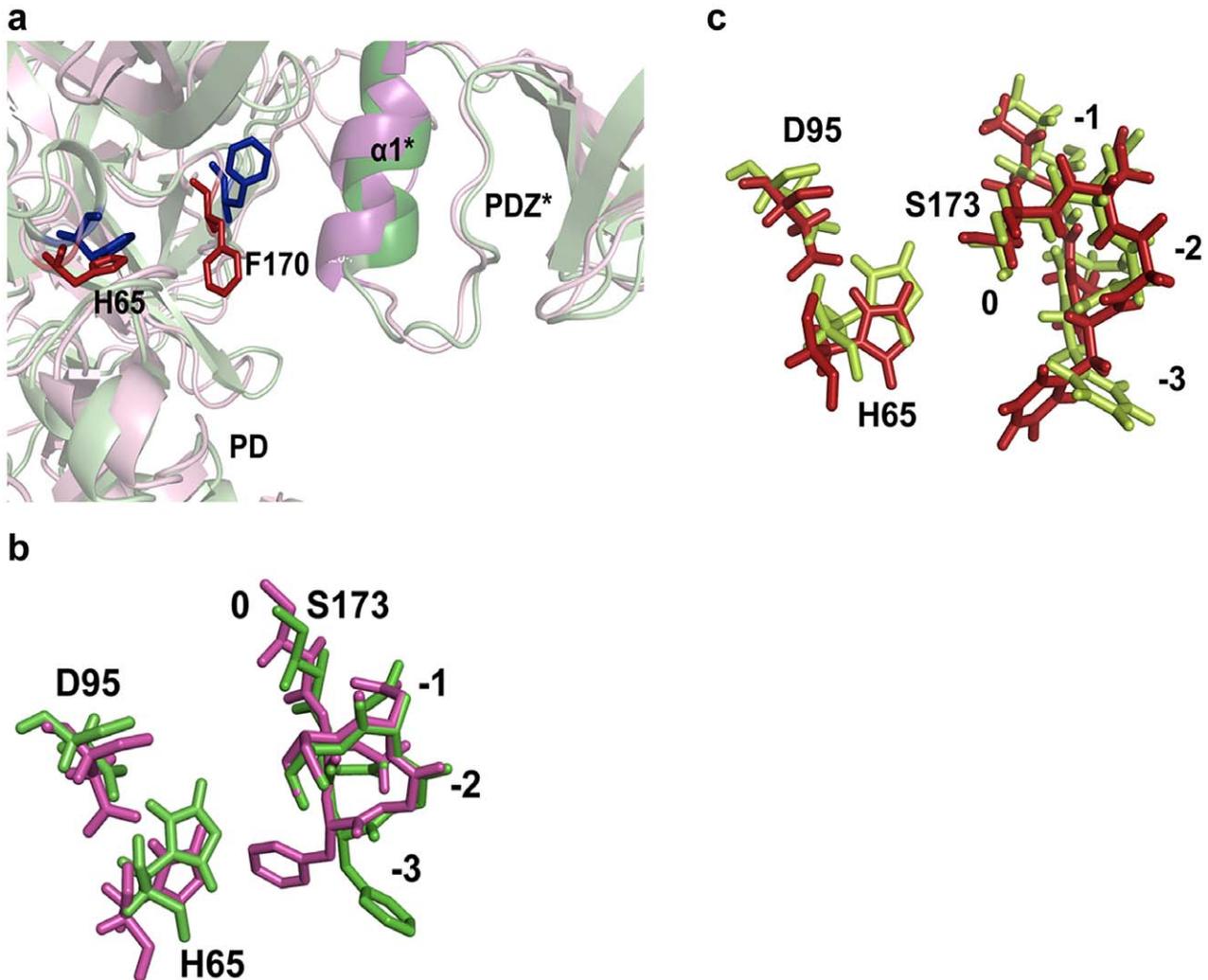


Figure 6. Structural changes at the oxyanion hole and YIGV groove upon peptide binding. a. Overlay of the oxyanion hole and catalytic triad residues represented as sticks for peptide GQYYFV bound (magenta) and unbound (green) structures. PD denotes serine protease domain of HtrA2. b. Overlay of the oxyanion hole and catalytic triad residues represented as sticks for peptide GSAWFSF bound (red) and unbound (limon green) structures. c. Role of PDZ in the formation of proper active site formation. The structural superposition of GQYYFV bound (pink) and unbound (green) structures shows $\alpha 5$ helix of PDZ of one subunit moves towards the LD loop and oxyanion hole of the adjacent subunit. The positions of the residues in the oxyanion hole are denoted as 0, -1, -2 and -3. doi:10.1371/journal.pone.0055416.g006

decrease in catalytic efficiency in SBP mutants. This observation suggests interaction of substrate protein with SBP brings about rearrangement around the active site of the enzyme by positively influencing its activity thus behaving as an allosteric regulator. The SBP mutants (N216A/S219A and E292A) show apparent decrease in V_{max} without significantly altering the apparent K_m (with L2 specificity pocket mostly unaltered) and hence follow the 'V system' of allosteric modulation [29]. In this system, both the relaxed (R) and the tensed (T) states bind the substrate at the active site with similar affinity while the peptide (activator) at SBP binds the R and T states with different affinity. This differential affinity of the peptide towards SBP along with R state stabilization shifts the equilibrium towards R state thus positively influencing its turnover rate and hence catalytic efficiency which has been observed in case of HtrA2.

In N-SPD, where the PDZ domain is absent, apparent decrease in K_m can be attributed to greater accessibility of the substrate to the active site. However, since the change in binding affinity is not

large, the specificity pocket might be mostly unaltered compared to the wild type which is confirmed through our MD studies where the loop L2 remains mostly unaltered. Interestingly, k_{cat} value in N-SPD has been found to decrease significantly which is suggestive of either a malformed oxyanion hole and/or decrease in overall protein stability which might be due to absence of supporting PDZ domain. However, similar studies with F16D (monomeric full length HtrA2 mutant) also show significant decrease in turnover rate and catalytic efficiency which accentuates the importance of intermolecular and not intramolecular PDZ-protease crosstalk in trimeric HtrA2. Our MDS supports this observation by demonstrating that in the peptide bound form of HtrA2, $\alpha 5^*$ of PDZ* moves towards LD loop of protease domain of adjacent subunit thus pushing phenyl ring of F170 of the oxyanion hole towards H65 of the catalytic triad (Figure 6a). This reorientation in the oxyanion hole makes the protease poised for catalysis as seen in other HtrA family members as well [12] thus significantly enhancing the turnover rate. Therefore, intermolecular crosstalk

stabilizes the active site and makes it catalytically competent establishing the requirement of complex trimeric architecture of the protease.

The GLGF motif (YIGV in HtrA2) is the canonical peptide binding site [2,4] in PDZ domains. However, in HtrA2, it is deeply embedded within a hydrophobic groove where the residues are intertwined with each other through several intramolecular interactions making the site highly inaccessible to the binding of peptide [4]. Thus, peptide binding to YIGV is only possible upon certain structural rearrangements at that site. Given the property of PDZ domains of having multiple docking sites and the fact that HtrA2 requires huge conformational changes for proper active site formation, we hypothesized presence of a relatively exposed pocket where peptide binding occurs prior to interaction with the buried YIGV groove. In our studies, we have found a novel surface exposed region (SBP) around PDZ domain which is easily accessible to the peptide. With an aim at understanding the allosteric mechanism in HtrA2 and whether the binding site is structurally conserved, we did a side-by-side comparison with the peptide-bound PDZ structure of its bacterial counterpart DegS that is known to exhibit allostery [30]. The structural overlay of peptide bound forms of these two proteins show striking structural similarity in the regions of binding (Figure 7a) with the GLGF groove (YIGV in HtrA2 and YIGI in DegS) oriented differently. Since the YIGV motif is buried in HtrA2 structure, its inaccessibility might be the reason for the peptide to initially bind to another relatively accessible region with similar hydrophobic milieu. However, in DegS, the YIGI groove is already exposed to accommodate the peptide easily and hence this kind of initial interaction is not required.

Our MDS studies show that peptide binding at SBP leads to subtle structural changes in the region adjoining YIGV leading to opening up of the pocket. The last β strand of PDZ domain which

lies on one side of YIGV groove moves away from it. The YIGV and the loop spanning residues 67–73 move away from each other while the loop comprising residues 263–277 of the β - α - β motif also drifts at an angle away from the YIGV making it more solvent exposed (Figure 7b). Therefore, upon SBP binding, the relative movements of the loops in vicinity of the hydrophobic YIGV pocket might confer it with the kind of exposure that is required for interaction with peptides. These observations along with our enzymology studies with SBP and YIGV mutants, led to defining a model (Figure 8) for allosteric propagation in HtrA2. The model suggests that initial binding of the peptide activator at SBP leads to structural fluctuations which result in subtle rearrangement at and around the YIGV groove (a part of greater SBP mesh as identified by Sitemap) thus exposing it. Opening up of the deeply embedded YIGV pocket makes it accessible to the substrate molecule which consequently leads to allosteric signal propagation at the active site in the serine protease domain.

This alternative non-canonical PDZ binding site though novel in HtrA family of proteins, is not unprecedented in literature. It has been observed that PDZ7 of the scaffold protein Glutamate receptor interacting protein 1 (GRIP1) has an alternative exposed hydrophobic pocket that binds its substrate GRASP-1 since the canonical binding site is deeply embedded within the protein [31]. Overlay of the PDZ from HtrA2 and PDZ7 of GRIP1 show striking structural similarity including the classical peptide binding groove and the novel non-canonical pocket (Figure S2). Thus, in these two proteins, perturbations at the alternative distal binding sites might be coupled dynamically to the classical binding groove by a complex mechanism that includes fast (ps–ns) timescale dynamics which consequently leads to allosteric signal propagation to the active site.

In the recent past, allosteric modulators have evolved into important drug targets due to several advantages they have over

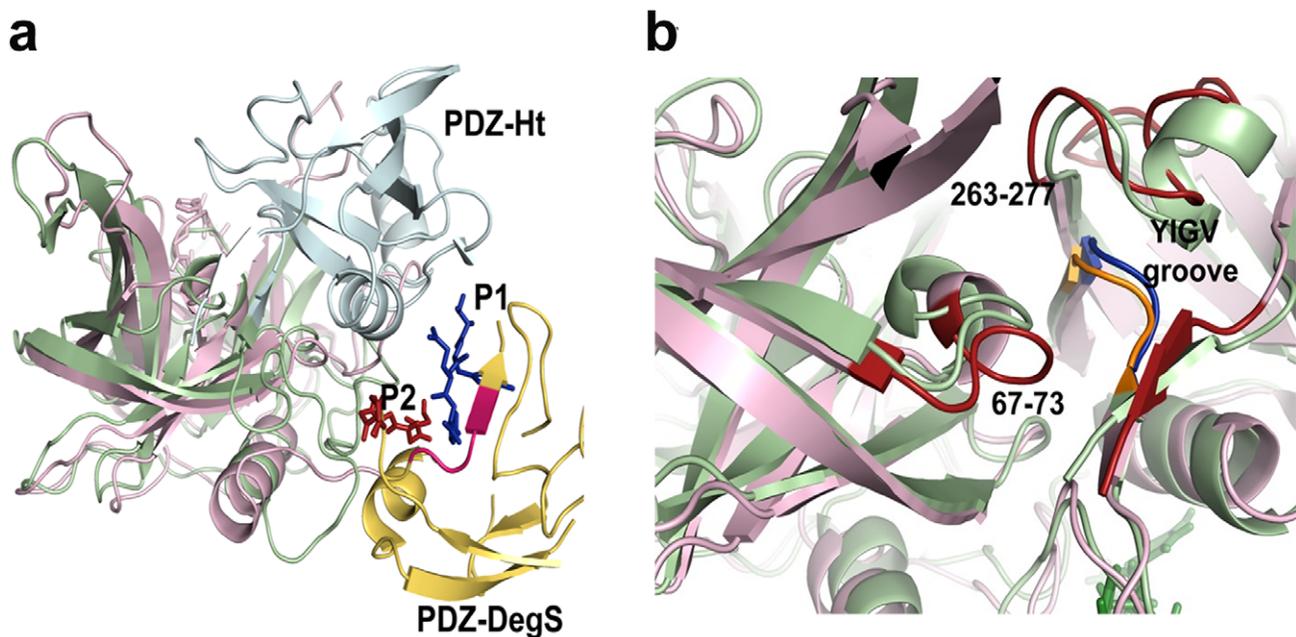


Figure 7. Structural comparison of PDZ domain orientation. a. Structural alignment of *E.coli* DegS (PDB ID: 1SOZ) and the peptide bound HtrA2 showing PDZ domains for both the proteins (represented in blue and yellow respectively) are oriented differently but the peptides, P1 (blue) and P2 (pink) represented as sticks for the respective proteins seem to bind to a structurally similar region. The GLGF substrate binding motif is exposed for DegS while buried for HtrA2 as shown in pink and blue respectively. b. Alignment of the peptide bound (pink) and unbound (green) structures at the region around the YIGV groove shows outward movement of the loops spanning residues 67–73 and 263–277 shown in red for the bound structures which leads to opening up of the YIGV groove.
doi:10.1371/journal.pone.0055416.g007

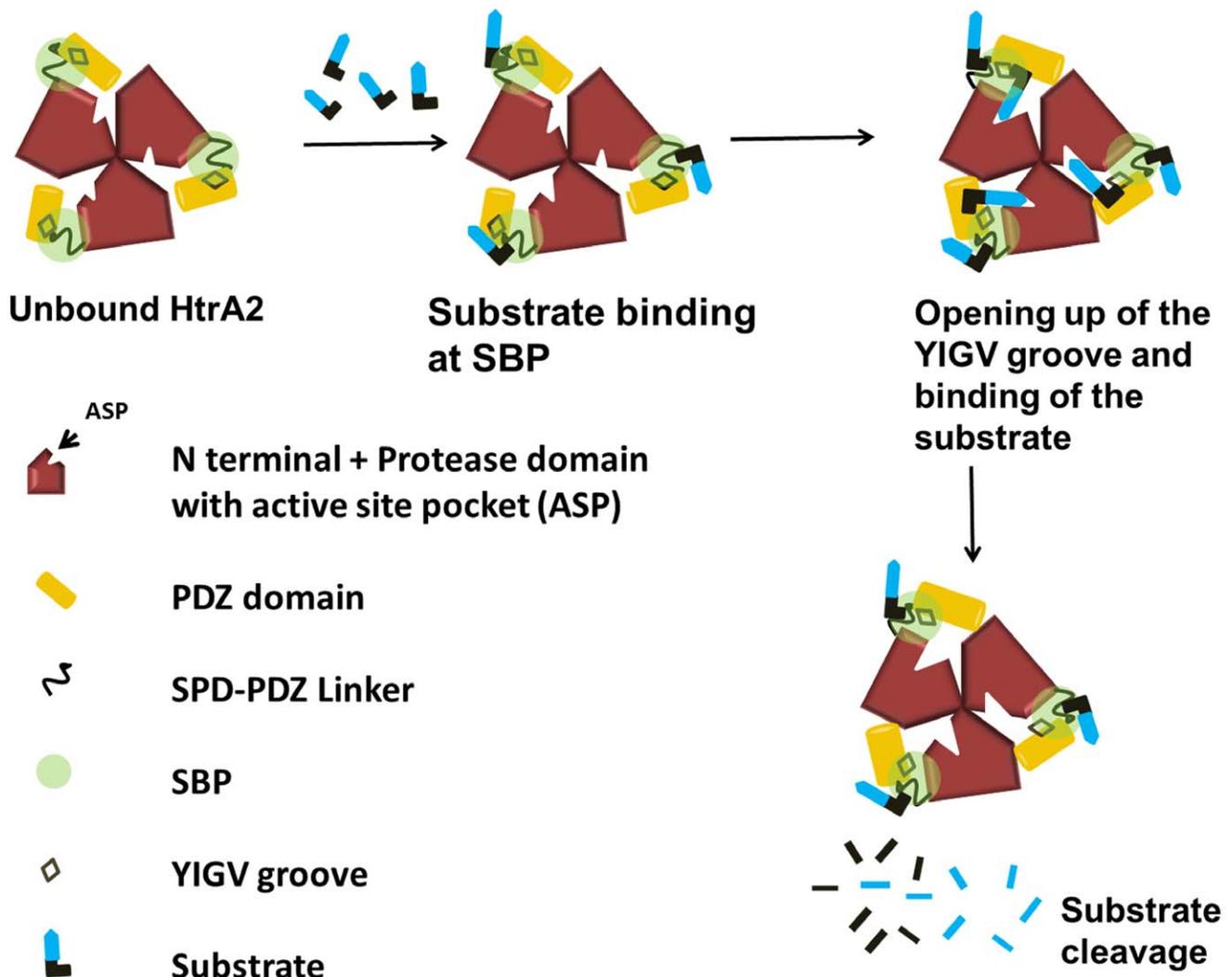


Figure 8. Allosteric model for HtrA2 protease activity. The substrate protein binds to relatively exposed part of SBP due to inaccessibility of the YIGV groove which triggers opening up of the PDZ domain. This reorientation makes the YIGV groove accessible for substrate interaction and the PDZ of a subunit moves closer to the protease domain of the adjacent subunit leading to formation of a proper active site and oxyanion hole. This complex allosteric signal propagation leads to subsequent substrate binding and catalysis at the active site pocket. Thus structural perturbations at these two distant sites (SBP and catalytic pocket) might be dynamically coupled to the canonical peptide binding groove through a complex allosteric mechanism.

doi:10.1371/journal.pone.0055416.g008

orthosteric ligands that include more diversity, less toxicity and absolute subtype selectivity [32,33]. Therefore, designing suitable SBP binding peptides or peptidomimetics of HtrA2 might be an excellent approach to modulate HtrA2 functions for devising therapeutic strategies against various diseases it is associated with.

Materials and Methods

Loop Modeling and Site Prediction

Crystal structure of HtrA2 (PDB ID: 1LCY) [4] obtained from Protein Data Bank [34] has missing N-terminal residues (AVPSP) and two flexible regions (211 RGEKKNSSSGISGSQ 225 and 149 ARDLGLPQT 157). These missing structures were modelled and loops were refined using Prime 3.0 (Schrödinger, LLC, New York, 2011), which was later subjected to molecular dynamics simulation for 5 ns with GROMACS, version 4.5.1 [17] to obtain the lowest energy structure of HtrA2. The binding sites were then predicted using SiteMap 2.5 (Schrödinger, LLC, New York, 2011).

Out of 5 pockets predicted, the site that scored the best based on its size, hydrophobic and hydrophilic characters, degree to which ligand might donate or accept hydrogen bonds and exposure to solvent was selected for further analysis. This site selective binding pocket (SBP) encompasses PDZ-protease interface with the involvement of hinge region and a part of PDZ domain (Table 1).

Peptide Designing and Molecular Docking

Based on properties of amino acids lining the binding site, fragment docking (Glide XP, Schrödinger, LLC, New York, 2011) [35] approach was used to dock 20 amino acids and 8 functional group replicas (N-methylacetamide, methanol, phenol, benzene, propane, acetate ion, methylammonium, methylguanidinium) at SBP [36]. Based on properties of the amino acids that form SBP, replicas were chosen and were used for generating fragments in combinations of four as shown in Table S1. Combine Fragment tool (Schrödinger, LLC, New York, 2011) was used to join the

fragments which were docked at SBP with three major filtering options (bond angle deviation 5 degrees, atom-atom distance 1 Å and fragment centroid distance 2.0). The set of replica functional groups that displayed the best docking scores were used to build the peptide. The amino acids Arg, Ser, Gln, Glu, Asp, Asn, Thr, Lys and their positions in tetrapeptide combination were chosen based on the functional groups they resembled. Subsequently all possible peptide combination of these amino acids with respect to their relative positions were generated. The predicted tetrameric peptides (Table S2) were selected and docked again with SBP.

In parallel, another mode of designing was used by identifying signature peptides from literature which bind HtrA2 [19]. Initially two peptides were chosen (GQYYFV and GGIRRV) and based on the sequence similarity and hydrophobicity, stretches of putative binding residues from two known binding partners of HtrA2 were identified (GPFPIIV from C-terminal region of β -casein and GSAWFSF, an internal motif of antiapoptotic Pea-15) [37]. A putative HtrA2 binding pattern was designed based on phage display library [21] which along with the earlier four sequences was used to generate all possible peptide combinations. Considering structural complementarity and three dimensional arrangements of amino acids at SBP, Gly and Val residues were added at N- and C-termini of some peptides to increase the stability of the docked complex. A 10 mer peptide having the sequence KNNPNNAHQN that does not match the consensus SBP binding peptide pattern was used as a negative control. These combinations were used for searching all possible sequences of known and potential HtrA2 binding partners [38].

All designed peptides were built *in silico* using BREED (Schrödinger, LLC, New York, 2011) and Combine Fragments tools which were then prepared for docking using LigPrep 2.5 (Schrödinger, LLC, New York, 2011). After ligand preparation, Confgen 2.3 (Schrödinger, LLC, New York, 2011) was used to generate all possible energetically minimum conformers of the designed peptides which were then docked using Glide [39,40].

In the modeled HtrA2 structure, energy minimization was done using Protein Preparation Wizard 2.2 (Epik Version 2.2, Schrödinger, LLC, New York, 2011) after addition of H-atoms. Molecular Docking was initiated by preparing Grid file (input file) which contains receptor (protein structure) and binding site information (Prime output). All three precision methods which include high throughput virtual screening (HTVS), standard precision (SP) and extra precision (XP) [35] of Glide [39,40] were used for docking these peptides on SBP. This series of docking methods were used to filter out energetically less favorable peptide conformers and get a subset of best possible peptides for further studies.

MD Simulation (MDS) and Analysis

After analyzing the docking results, best HtrA2-peptide complexes based on Glide XP score and E-model value were used for Molecular Dynamic Simulation which was performed using Desmond 2010 [22] software package. Optimized Potentials for Liquid Simulations (OPLS) [41] all-atom force field was used to analyze model stability. The protein structures were solvated with Monte Carlo simulated TIP3P [42] water model with a 10 Å buffer space from the protein edges in an orthorhombic box and the system was then neutralized by replacing water molecules with sodium and chloride counter ions. Similarly, unbound HtrA2 system was also developed as a control. Neutralization of systems was done by adding 2 Na⁺ ions in unbound HtrA2 and 4 Na⁺ ions for peptide bound complexes. The particle-mesh Ewald method (PME) [43] was used to calculate long-range electrostatic interactions with a grid spacing of 0.8 Å. Van der Waals and

short range electrostatic interactions were smoothly truncated at 9.0 Å. Nose-Hoover thermostats were utilized to maintain the constant simulation temperature and the Martina-Tobias-Klein method was used to control the pressure [44]. The equations of motion were integrated using the multistep RESPA integrator [45] with an inner time step of 2.0 fs for bonded interactions and non-bonded interactions within the short range cut-off. An outer time step of 6.0 fs was used for non-bonded interactions beyond the cut-off. These periodic boundary conditions were applied throughout the system.

These prepared systems were equilibrated with the default Desmond protocol that comprises a series of restrained minimizations and MDS. Two rounds of steepest descent minimization were performed with a maximum of 2000 steps and a harmonic restraint of 50 kcal/mol/per Å² on all solute atoms followed by a series of four MDS. The first simulation was run for 12 ps at a temperature of 10 K in the NVT (constant number of particles, volume, and temperature) ensemble with solute heavy atoms restrained with force constant of 50 kcal/mol/Å². The second simulation was similar to the first except it was run in the NPT (constant number of particles, pressure, and temperature) ensemble. A 24 ps simulation followed with the temperature raised to 300 K in the NPT ensemble and with the force constant retained. The last one was a 24 ps simulation at 300 K in the NPT ensemble with all restraints removed. This default equilibration was followed by a 5000 ps NPT simulation to equilibrate the system. A 30 ns NPT production simulation was then run and coordinates were saved in every 2 ps of time intervals.

The total trajectory of MD simulation was 30 ns. MD Simulation was analyzed using the analytical tools in the Desmond package. In MD quality analysis, potential energy of the protein as well as total energy of the entire system was calculated. The lowest potential energy conformations were then used for comparative analysis of peptide bound and unbound structures. Trajectories of peptide bound complexes and unbound HtrA2 were then compared based on their overall calculated RMSD (root mean square deviation), domain wise RMSD and RMSF (root mean square fluctuation) values and were plotted using GraphPad Prism 5.0 (GraphPad Software, San Diego, CA, USA).

Production of Recombinant HtrA2 Wild Type, its Mutants and Domains

Mature (Δ 133 HtrA2) with C-terminal his₆-tag in pET-20b (Addgene, Cambridge, MA) was expressed in *E. coli* strain BL21 (DE3) pLysS. N-SPD, comprising N-terminal and serine protease domains (residues 1–210) of HtrA2 was sub cloned into pMALc5E-TEV using appropriate primers. Point mutations were introduced into pET-20b Δ 133 HtrA2 by PCR using primer sets that included mutations for residues N216A, S219A, E292A, E296A and F16D. N-SPD clone and these mutants were confirmed by DNA sequencing. Protein expression was induced by culturing cells at 18°C for 20 h in presence of 0.2 mM isopropyl-1-thio-D-galactopyranoside. Cells were lysed by sonication and the centrifuged supernatants for HtrA2 and its mutants were incubated with pre-equilibrated nickel-IDA beads for 1 h at room temperature. Protein purification was done using Ni-affinity chromatography as described earlier [19]. Eluted protein was further purified using gel permeation chromatography. N-SPD was purified using amylose resin where the bound protein was eluted using 10 mM maltose and was subjected to TEV protease cleavage [46] to remove maltose binding protein (MBP). N-SPD was further separated from MBP by gel filtration using Superdex 75 column. All purified proteins were analyzed by SDS-PAGE for

purity. The fractions with >95% purity were stored in aliquots at -80°C until use.

FITC- β -Casein Cleavage Assay

The proteolytic activity of wild type and the mutants were determined using FITC-labelled β -casein cleavage assay [47]. Fluorescent substrate cleavage was determined by incubating 200 nM of enzymes with increasing concentration (0–25 μM) of β -casein at 37°C in cleavage buffer (20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$, pH 8.0, 100 mM NaCl, 0.1 mM DTT). Fluorescence was monitored in a multi-well plate reader (Berthold Technologies) using excitation wavelength of 485 nm and emission at 545 nm. Reaction rates v_0 ($\mu\text{M}/\text{min}$) were determined by linear regression analysis corresponding to the maximum reaction rates for individual assay condition. Assays are representative of at least three independent experiments done in triplicate. 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).

Supporting Information

Figure S1 Interaction of peptides with HtrA2. a. Ligplot for GSAWFSF with HtrA2 which represents residues involved and the nature of interactions. b. Ligplot for GQYYFV interaction pattern with HtrA2. c. Ligplot for GPFPIIV with HtrA2 which represents residues involved and the nature of interactions. d. Ligplot for SEHRRHFPNCFV peptide with HtrA2 which represents residues involved and the nature of interactions. The residues of peptides and HtrA2 involved in interaction are shown in blue and red respectively.

(TIF)

Figure S2 Comparison of SBP and allosteric pocket of GRIP-1 protein. Structural overlay of the protein GRIP-1 (green) bearing PDB ID 1M5Z and GQYYFV bound HtrA2 (pink) shows striking resemblance of the orientation of buried GLGF motif shown in yellow and blue respectively. The α helix denoted as αB (green) for GRIP-1, known to be involved in formation of allosteric pocket overlays very well with the one involved in SBP formation (orange) in GQYYFV (red sticks) - HtrA2 complex.

(TIF)

Figure S3 ITC studies for activating peptide with HtrA2 and the SBP double mutant. The peptide used was 13mer SEHRRHFPNCFV, which has similar consensus sequence as

defined for PDZ peptide groove binding substrate. The peptide was better in terms of solubility as compared to other activating peptides and binding studies were done using Isothermal titration calorimetry. The titrations were carried out using Micro Cal ITC200 (GE Healthcare) with the calorimetry cell containing 200 μl of wild type or N216A/S219A mutant HtrA2 in 20 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ buffer, 100 mM NaCl, pH 7.8. The concentration of protein was in range from 20 to 50 μM and was titrated with 1.5 μl injections of a solution containing 0.4 mM activator peptide reconstituted in the same buffer. To correct the effect of heat of dilution, a blank injection was made under identical conditions. All experiments were performed at 25°C and the data was analyzed using the manufacture provided MicroCal software with the integrated heat peaks fitted to a one site-binding model. Simulated ITC raw data for the protein with the activating peptide is represented in the upper panel and the integrated data in the lower panel. The dissociation constant was calculated to be 7.5 μM for wild type (left panel) and no significant heat change was observed for the SBP double mutant (right panel).

(TIF)

Table S1 Docking analysis of replica fragments with HtrA2. The fragments have been arranged according to their docking scores.

(DOC)

Table S2 Designed peptide fragments. Fragments of peptide combinations generated based on functional group studies have been enlisted.

(DOC)

Movie S1 Orientation of active site triad and oxyanion hole formation during MD simulation of HtrA2-peptide complex. From this visual representation of HtrA2 peptide (GSAWFSF) complex during MD simulation it can be seen that the catalytic triad residues H65, D95, S173 reorient to form an active conformation along with oxyanion hole residues (N172, G171 and F170). All the residues involved are represented as sticks. This movie shows proper active site and oxyanion hole formation.

(AVI)

Author Contributions

Conceived and designed the experiments: KB. Performed the experiments: PRB RRK NG NS LKC. Analyzed the data: KB PRB RRK GMS. Wrote the paper: KB PRB RRK.

References

- Ma B, Tsai C, Haliloglu T, Nussinov R (2011) Dynamic Allosterity: Linkers Are Not Merely Flexible. *Structure* 19: 907–917.
- Jeleni F, Oleksy A, Smietana K, Otlewski J (2003) PDZ domains common players in the cell signaling. *Acta Biochim Pol* 50: 985–1017.
- Hung A, Morgan S (2002) PDZ Domains: Structural Modules for Protein Complex Assembly. *J Biol Chem* 277: 5699–5702.
- Li W, Srinivasula S, Chai J, Li P, Wu J, et al. (2002) Structural insights into the pro-apoptotic function of mitochondrial serine protease HtrA2/Omi. *Nat Struct Biol* 9: 436–441.
- Suzuki Y, Takahashi-Niki K, Akagi T, Hashikawa T, Takahashi R (2004) Mitochondrial protease Omi/HtrA2 enhances caspase activation through multiple pathways. *Cell Death Differ* 11: 208–216.
- Martins L, Iaccarino I, Tenev T, Gschmeissner S, Totty N, et al. (2002) The serine protease Omi/HtrA2 regulates apoptosis by binding XIAP through a reaper-like motif. *J Biol Chem* 277: 439–444.
- Verhagen A, Silke J, Ekert P, Pakusch M, Kaufmann H, et al. (2002) HtrA2 promotes cell death through its serine protease activity and its ability to antagonize inhibitor of apoptosis proteins. *J Biol Chem* 277: 445–454.
- Kooistra J, Milojevic J, Melacina G, Ortega J (2009) A new function of human HtrA2 as an amyloid-beta oligomerization inhibitor. *J Alzheimers Dis* 17: 281–294.
- Johnson F, Kaplitt M (2009) Novel mitochondrial substrates of Omi indicate a new regulatory role in neurodegenerative disorders. *PLoS One* 4: e7100.
- Lin C, Chen M, Chen G, Tai C, Wu R (2011) Novel variant Prol43Ala in HTRA2 contributes to Parkinson's disease by inducing hyperphosphorylation of HTRA2 protein in mitochondria. *Hum Genet* 130: 817–827.
- Nam M, Seong Y, Park H, Choi J, Kang S, et al. (2006) The homotrimeric structure of HtrA2 is indispensable for executing its serine protease activity. *Exp Mol Med* 38: 36–43.
- Singh N, Kupplil R, Bose K (2011) The structural basis of mode of activation and functional diversity: A case study with HtrA family of serine protease. *Arch Biochem Biophys* 516: 85–96.
- Sohn J, Grant RA, Sauer RT (2010) Allostery is an intrinsic property of the protease domain of DegS: implications for enzyme function and evolution. *J Biol Chem* 285: 34039–34047.
- Merdanovic M, Mamant N, Meltzer M, Poepsel S, Auckenthaler A, et al. (2010) Determinants of structural and functional plasticity of a widely conserved protease chaperone complex. *Nat Struct Mol Biol* 17: 837–843.
- Beuming T, Farid R, Sherman W (2009) High-energy water sites determine peptide binding affinity and specificity of PDZ domains. *Protein Sci* 18: 1609–1619.

16. Sherman W, Day T, Jacobson M, Friesner R, Farid R (2006) Novel procedure for modeling ligand/receptor induced fit effects. *J Med Chem* 49: 534–553.
17. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark A, et al. (2005) GROMACS: fast, flexible, and free. *J Comput Chem* 26: 1701–1718.
18. Halgren T (2007) New method for fast and accurate binding-site identification and analysis. 69: 146–148.
19. Martins L, Turk B, Cowling V, Borg A, Jarrell E, et al. (2003) Binding specificity and regulation of the serine protease and PDZ domains of HtrA2/Omi. *J Biol Chem* 278: 49417–49427.
20. Ma S, Song E, Gao S, Tian R, Gao Y (2007) Rapid characterization of the binding property of HtrA2/Omi PDZ domain by validation screening of PDZ ligand library. *Sci China C Life Sci* 50: 412–422.
21. Zhang Y, Appleton B, Wu P, Wiesmann C, Sidhu S (2007) Structural and functional analysis of the ligand specificity of the HtrA2/Omi PDZ domain. *Protein Sci* 16: 2454–2471.
22. Bowers K, Chow E, Xu H, Ron O, Eastwood M, et al. (2006) Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters; Florida.
23. Savopoulos J, Carter P, Turconi S, Pettman G, Karran E, et al. (2000) Expression, Purification, and Functional Analysis of the Human Serine Protease HtrA2. *Protein Expr Purif* 19: 227–234.
24. Ibrahim A, Mansour I, Wilson M, Mokhtar D, Helal A, et al. (2012) Study of survivin and X-linked inhibitor of apoptosis protein (XIAP) genes in acute myeloid leukemia (AML). *Lab Hematol* 18: 1–10.
25. Nagata M, Nakayama H, Tanaka T, Yoshida R, Yoshitake Y, et al. (2011) Overexpression of cIAP2 contributes to 5-FU resistance and a poor prognosis in oral squamous cell carcinoma. *Br J Cancer* 105: 1322–1330.
26. Navakanit R, Graidist P, Lecanansakiri W, Dechsukum C (2007) Growth inhibition of breast cancer cell line MCF-7 by siRNA silencing of Wilms tumor 1 gene. *J Med Assoc Thai* 90: 2416–2421.
27. Sugiyama H (1998) Wilms tumor gene (WT1) as a new marker for the detection of minimal residual disease in leukemia. *Leuk Lymphoma* 30: 55–61.
28. Kraut J (1977) Serine proteases: structure and mechanism of catalysis. *Annu Rev Biochem* 46: 331–358.
29. Mazat J, Patte J (1976) Lysine-sensitive aspartokinase of *Escherichia coli* K12. Synergy and autosynergy in an allosteric V system. *Biochemistry* 15: 4053–4058.
30. Sohn J, Sauer R (2009) OMP peptides modulate the activity of DegS protease by differential binding to active and inactive conformations. *Mol Cell* 33: 64–74.
31. Feng W, Fan J, Jiang M, Shi Y, Zhang M (2002) PDZ7 of glutamate receptor interacting protein binds to its target via a novel hydrophobic surface area. *J Biol Chem* 277: 41140–41146.
32. Christopoulos A, May LT, Avlani VA, Sexton PM (2004) G protein-coupled receptor allostery: the promise and the problem(s). *Biochem Soc Trans* 32: 873–877.
33. May L, Leach K, Sexton P, Christopoulos A (2007) Allosteric modulation of G protein-coupled receptors. *Annu Rev Pharmacol Toxicol* 47: 1–51.
34. Berman H, Westbrook J, Feng Z, Gilliland G, Bhat T, et al. (2000) The Protein Data Bank. *Nucleic Acids Res* 28: 235–242.
35. Friesner R, Murphy R, Repasky M, Frye L, Greenwood J, et al. (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem* 49: 6177–6196.
36. Zeng J, Treutlein H (1999) A method for computational combinatorial peptide design of inhibitors of Ras protein. *Protein Eng* 12: 457–468.
37. Trencia A, Fiory F, Maitan M, Vito P, Barbagallo A, et al. (2004) Omi/HtrA2 promotes cell death by binding and degrading the anti-apoptotic protein ped/pea-15. *J Biol Chem* 279: 46566–46572.
38. Ma S, Song E, Gao S, Tian R, Gao Y (2007) Rapid characterization of the binding property of HtrA2/Omi PDZ domain by validation screening of PDZ ligand library. *Sci China C Life Sci* 50: 412–422.
39. Halgren T, Murphy R, Friesner R, Beard H, Frye L, et al. (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem* 47: 1750–1759.
40. Friesner R, Banks J, Murphy R, Halgren T, Klicic J, et al. (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* 47: 1739–1749.
41. Jorgensen W, Maxwell D, Tirado-Rives J (1996) Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc* 118: 11225–11236.
42. Jorgensen W, Chandrasekhar J, Madura J, Impey R, Klein M (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79: 926–935.
43. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J Chem Phys* 98: 10089–10093.
44. Martyna G, Tobias D, Klein M (1994) Constant pressure molecular dynamics algorithms. *J Chem Phys* 101: 4177–4189.
45. Humphreys D, Friesner R, Berne B (1994) A Multiple-Time-Step Molecular Dynamics Algorithm for Macromolecules. *J Phys Chem* 98: 6885–6892.
46. Nallamsetty S, Kapust R, Tózsér J, Cherry S, Tropea J, et al. (2004) Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro. *Protein Expr Purif* 38: 108–115.
47. Twining S (1984) Fluorescein isothiocyanate-labeled casein assay for proteolytic enzymes. *Anal Biochem* 143: 30–34.



Review

The structural basis of mode of activation and functional diversity: A case study with HtrA family of serine proteases

Nitu Singh, Raja R. Kuppili, Kakoli Bose*

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

ARTICLE INFO

Article history:

Received 14 September 2011

Available online 18 October 2011

Keywords:

HtrA
PDZ domain
Active site
Protease
Chaperone

ABSTRACT

HtrA (High temperature requirement protease A) proteins that are primarily involved in protein quality control belong to a family of serine proteases conserved from bacteria to humans. HtrAs are oligomeric proteins that share a common trimeric pyramidal architecture where each monomer comprises a serine protease domain and one or two PDZ domains. Although the overall structural integrity is well maintained and they exhibit similar mechanism of activation, subtle conformational changes and structural plasticity especially in the flexible loop regions and domain interfaces lead to differences in their active site conformation and hence in their specificity and functions.

© 2011 Elsevier Inc. All rights reserved.

Introduction

HtrA¹ proteins belong to a unique family of oligomeric serine proteases (S1, chymotrypsin family) that are conserved from prokaryotes to humans. DegP or HtrA is the first one in the family to be identified in *Escherichia coli* [1] which has a dual temperature dependent chaperone-protease activity and is active in the periplasm of the bacterium. Later on its homologs have been identified and studied in wide spectrum of organisms that include bacteria, fungi, plants, frogs, fish and mammals [2]. HtrA proteases are well-known for their structural complexity which is reflected in their multi-tasking ability. The structure of the protease comprises a serine protease domain and one or more C-terminal PDZ or protein-protein interaction domains [2]. They usually form higher order oligomers extending from ~100 kDa trimer (as in human HtrA2) to 24-mers of 1.2 MDa as in DegP [3–9]. The chief role of *E. coli* DegP is to degrade misfolded proteins in the periplasm [2]. However, fine-tuning of the structural organization leads to functional diversity amongst the family members although their overall structural integrity is well maintained. All the HtrA proteases share a common trimeric pyramidal architecture where each monomer comprises two or three major domains [10] and exhibit similar mechanism of activation. However, subtle conformational changes and variations lead to differences in their active site conformation and hence in

their specificity and functions. For example, HtrA proteins are involved in a variety of biological functions and pathogenicity such as protein quality control including photosystem II, in plants and cyanobacteria [11], unfolded protein response (UPR), cell growth, apoptosis, arthritis, cancer and metabolism of amyloid precursor proteins [12–15]. Here, in this review we look into this intriguing aspect with the focus on all HtrA proteins whose structures are currently available in the literature.

Representative HtrA family of serine proteases: structural and functional perspectives

Given their unique structural conservation despite overall low sequence identity (Fig. 1A) and involvement in critical cellular functions, plethora of research has been carried out on HtrA family of serine proteases in the past decade leading to identification and characterization of HtrA proteins in different organisms. Several members of the HtrA family belong to the same organism although they might reside in different cellular compartments as shown in Table 1. Out of these several proteins, structures of some of them have been solved which provide an in-depth understanding of their overall domain organization (Fig. 1B), structural architecture and correlation between conformational dynamics and functional diversity.

Prokaryotic HtrAs

DegP: dual chaperone-protease activity

E. coli DegP (also called HtrA or protease Do), a periplasmic heat shock protease, maintains protein homeostasis in the bacterial

* Corresponding author. Address: Bose Lab, Advanced Centre for Treatment, Research and Education in Cancer (ACTREC), Tata Memorial Centre, KS-138, Sector 22, Kharghar, Navi Mumbai 410210, India. Fax: +91 22 27405080.

E-mail address: kbose@actrec.gov.in (K. Bose).

¹ Abbreviations used: HtrA, high temperature requirement protease A; IGF, Insulin-like growth factor; UPR, unfolded protein response.

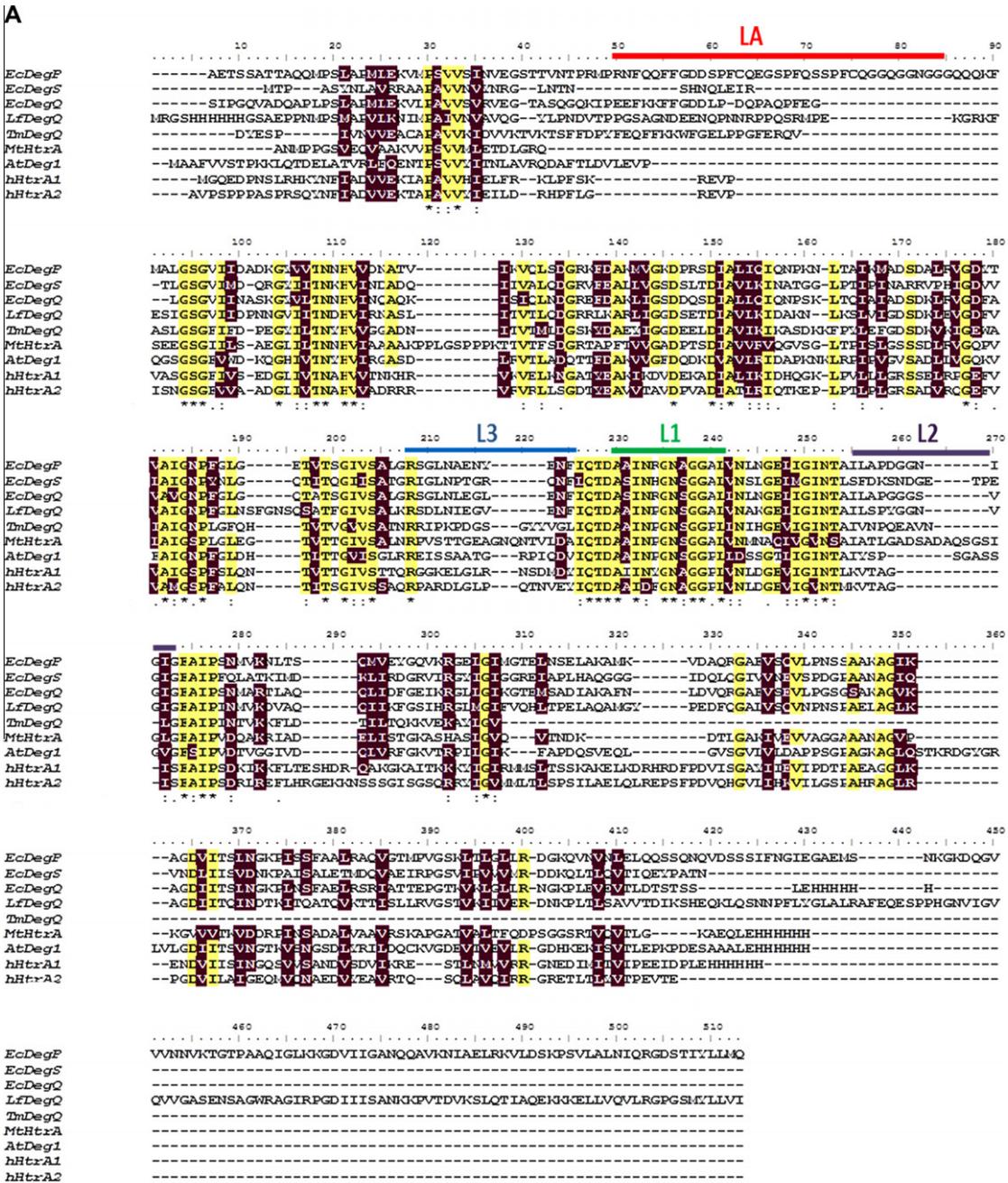


Fig. 1. (A) Multiple sequence alignment of the HtrA family of serine proteases. Identical residues across the protein family are marked in yellow and similar residues in magenta. The loops LA, L1, L2, and L3 are indicated above the respective regions. (B) Schematic representation of the domain organization of HtrA protein family. The protease domain is in green, PDZ1 and PDZ2 domains in yellow and red respectively, SS (signal sequence) is shown as blue triangles, Tm (transmembrane domain) as brown squares, IGFBP (Insulin-like growth factor binding) in orange, and KI (Kazal protease inhibitor domain) as purple pentagon. In the figure, *Ec* represents *Escherichia coli*, *Lf*: *Legionella fallonii*, *Tm*: *Thermotoga maritima*, *h*: *Homo sapiens*, *Mt*: *Mycobacterium tuberculosis* and *At*: *Arabidopsis thaliana*.

periplasm by combining chaperone and protease activity in an ATP-independent, temperature dependent manner [1,16]. DegP de-

grades or refolds misfolded and aggregated proteins, thereby protecting the bacterial envelope from their detrimental effects and

Table 1
Representative HtrA family of serine proteases.

Organism	Representative proteases	No. of amino acids	PDB ID	Sub cellular localization
<i>Escherichia coli</i>	DegP	355	1KY9	Cell envelope
	DegS	448	1TE0	Periplasmic space
	DegQ	428	3STI	Periplasmic space
<i>Cyanobacteria</i> (<i>Synechocystis</i> sp6803)	HtrA (slr1204)	452	NA	Cell outer membrane
	HhoA (hhoA, sl11679)	394	NA	Periplasm
	HhoB (hhoB, sl11427)	416	NA	Predicted to target the periplasm
<i>Thermotoga maritima</i>	HtrA	459	1L1J	Periplasmic space
<i>Legionella fallonii</i>	DegQ	451	3PV2	Periplasmic space
<i>Mycobacterium tuberculosis</i>	HtrA1 (RV1223)	528	2Z9I	Intermembrane space
	HtrA2 (RV0983)	464	NA	Periplasm
	HtrA3 (RV0125)	355	NA	Periplasm
	Deg1 (At3g27925)	439	3QO6	Attached to the luminal side of thylakoid membrane
<i>Arabidopsis thaliana</i>	Deg5 (At4g18370)	323	NA	
	Deg8 (At5g39830)	434	NA	
	Deg2 (At2g47940)	606	NA	Attached to the stromal side
	HtrA1	480	3NUM	Secreted
<i>Homo sapiens</i>	HtrA2	458	1LCY	Mitochondrion intermembrane space
	HtrA3	453	NA	Secreted
	HtrA4	445	NA	Unknown

NA – Not available.

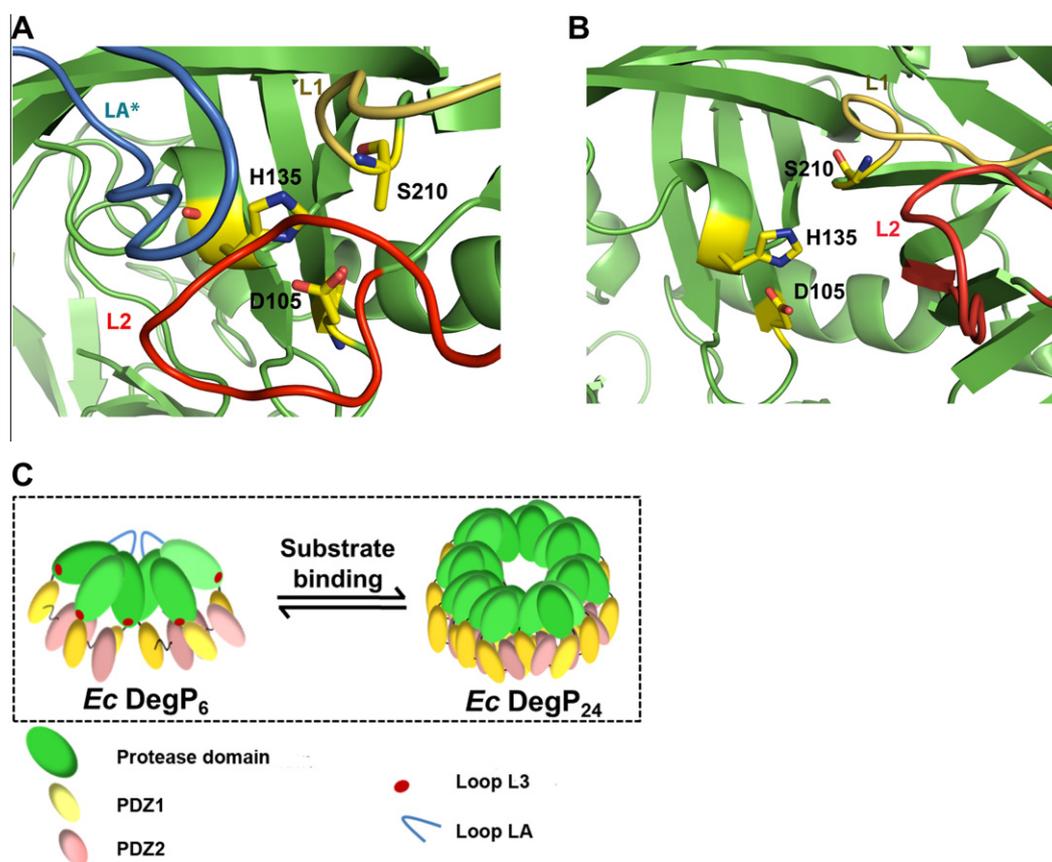


Fig. 2. Ribbon plot of the active site loop orientation in inactive DegP₆ (PDB ID: 1KY9) and active DegP₂₄ (PDB ID: 3CS0). (A) In hexameric DegP, the loops L1 (yellow), L2 (red) and neighboring subunit LA* (blue) together hinder the substrate accessibility to the catalytic triad (represented in stick mode). (B) In DegP₂₄, L1 and L2 reorient to form an active conformation with LA loop positioned away from active triad. (C) Schematic model of DegP activation. Substrate binding induces an oligomeric switch from inactive hexamer to an active state 24-mer.

plays an essential role in bacterial survival under stressful conditions [17]. While the protease activity ensures that unfolded periplasmic proteins are degraded, recent evidence suggests that the outer membrane proteins are protected in the periplasm via its chaperoning activity [15].

Architecture of DegP

The DegP monomer contains an N-terminal chymotrypsin like serine protease domain (residues 1–259) and two PDZ domains (PDZ1, residues 260–358; PDZ2, residues 359–448). The X-ray structure revealed that DegP oligomerizes into a hexameric cage

with a central chamber containing the sequestered proteolytic sites [4]. Both the top and bottom of the cage are constructed by six protease domains, whereas 12 PDZ domains generate the mobile side-walls. Three spacing pillars, each made up of two intertwined LA loops extending from opposite monomers connect this loosely bound dimer of tight trimers. The PDZ domains extend outward from the ends of the cage and interact with each other thus acting as the main gatekeepers of the inner chamber [18]. In addition, each LA loop extends beyond the pillar it is forming and protrudes into the active site of the opposite trimer where it interacts with its active site loops L1 and L2. The resulting loop triad LA^{*}-L1-L2 completely blocks the entrance to the catalytic site forcing it into a twisted inactive conformation (Fig. 2A). Therefore, the hexameric form of DegP constitutes an inactive state of the protein where substrate binding and catalysis is prevented [4].

Structural basis of functional oligomeric forms: DegP₁₂ and DegP₂₄

In 2008, two independent groups, Krojer et al. and Jiang et al. presented the structures of novel forms of DegP that oligomerize into 12-mer (DegP₁₂) and 24-mer (DegP₂₄) globular cages in the presence of substrates. In DegP₂₄, the trimers are located at the vertices of an octahedron forming a porous shell that encloses a large inner cavity with a diameter of 110 Å. Interaction between neighboring trimeric units is mediated through pairwise inter-trimer PDZ1-PDZ2 interactions. In the DegP₁₂, the four trimeric units arrange in a tetrahedral shell also enclosing a central cavity.

The cascade of events that follows when substrate binding induces change in the oligomeric state of DegP include conformational rearrangement of the linker between PDZ1 and PDZ2, destabilization of hexamer, its dissociation and finally formation of functional 12-mer or 24-mer cages (Fig. 2B). In this new conformation, the otherwise flexible PDZ domains become immobilized as PDZ1 interacts with protease loop L3. This interaction stabilizes the active site through inter-molecular interaction thus re-modeling the proteolytic site into an active state [19]. Inter-trimeric interactions in these larger oligomers are further facilitated because substrate binding induces displacement of a polar surface in PDZ1 leading to the exposure of a hydrophobic cleft in this domain. This enables it to now interact with a complementary hydrophobic patch on the surface of PDZ2 of an adjacent trimer [20].

In DegP₁₂ and DegP₂₄, the catalytic sites are in an active conformation. This is due to higher order oligomerization of the hexamer that leads to movement of loop LA away from the active site of the neighboring subunit. This consequently releases loops L1 and L2 to set up a functional proteolytic site as shown in Fig. 2C [3,5]. Once the substrate is cleaved, DegP₁₂ and DegP₂₄ cages either revert to their inactive hexameric forms or undergo auto-proteolysis stimulated by peptides resulting from substrate hydrolysis. This showcases a self-regulatory mechanism of DegP to eliminate excess of HtrA produced under stress conditions once its enzymatic activities are no longer needed by the cell [21,22].

Functional implications – antagonistic functions of inner chamber

DegP hexamer functions as a substrate filter by allowing only unfolded proteins to enter its inner cavity. These unfolded protein substrates assemble the chaperone-protease machine into a final higher order oligomeric complex of 12 or 24-mers [23]. The inner cavity has dual antagonistic functions of folding as well as proteolysis. The fate of an encapsulated substrate depends on its initial folding state as well as its propensity to adopt native conformation fast enough to escape the proteolytic machinery of DegP [5].

DegP cleaves its substrate processively by employing hold-and-cut mechanism. For this, a molecular scale is set up by the peptide binding site of PDZ1 and protease domain which exhibit strikingly similar substrate specificities. Both these sites anchor small hydrophobic residues at the C-terminus of the ligand by β-sheet aug-

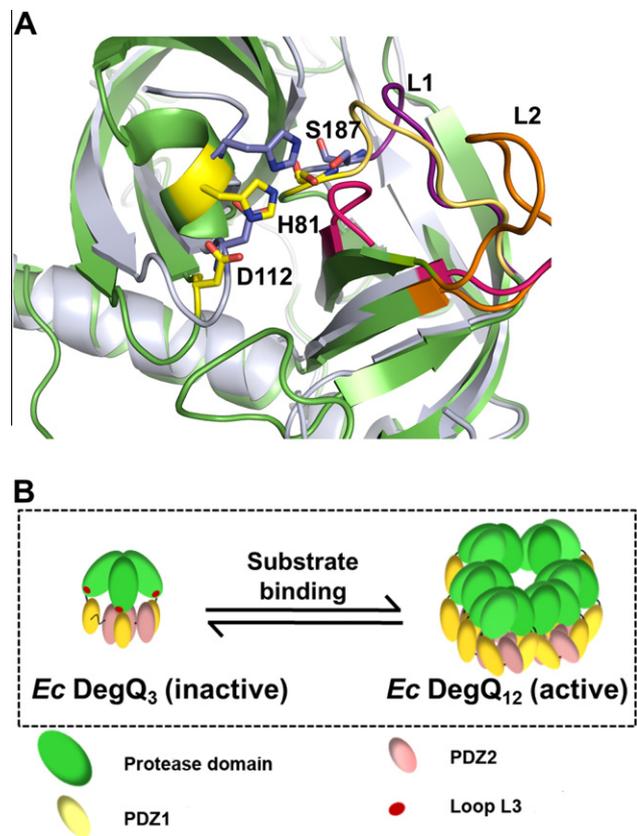


Fig. 3. (A) Comparison of the protease domain in inactive trimer versus active dodecamer of *E. coli* DegQ. Structural alignment of inactive (light blue) and active (green) protease domains highlights the conformational changes occurring in mechanically important loops L1 (yellow), L2 (orange) in inactive versus L1 (purple) and L2 (magenta) of active protease. In dodecameric form, loops L1 and L2 adopt ordered conformations as opposed to disordered L1 and L2 in the trimeric form (PDB ID. DegQ₃: 3STI and 3STJ for DegQ₁₂). (B) Activation model of *E. coli* DegQ. An equilibrium shift from a trimer to dodecamer on substrate binding renders the protease active.

mentation. This synchronization between substrate binding and anchoring enables the PDZ domains to capture peptide fragments released from the proteolytic sites via their newly generated C termini and to prepare them for the next cleavage cycle. Thus DegP undergoes allosteric activation by a ligand which also acts as its potential substrate [24].

DegQ: functional complement of DegP

Another periplasmic protease DegQ of *E. coli* also known as HhoA [25,26] shows considerable sequence similarity to and exhibits functional overlap with its homolog DegP. DegQ is about 455 amino acids long and shares 60% sequence homology with DegP. It is capable of rescuing temperature sensitive DegP⁻ strain [25]. Like DegP⁻, it has two PDZ domains and cleaves its misfolded protein substrates and thereby, is capable of functionally complementing DegP under certain conditions.

Oligomeric assembly

The catalytically active DegQ forms a higher order oligomer comprising 12 monomeric subunits [2]. It has been observed that substrate binding leads to formation of higher order oligomers in a concentration dependent manner. Lower substrate concentration leads to formation of a dodecamer, while 24-mer DegQ structure is formed with increasing substrate concentration. PDZ2 is mainly responsible for inter-trimeric contacts but is not essential for main-

taining protease activity. This is due to the observation that upon PDZ2 deletion, oligomerization of DegQ is limited to a dodecamer although the substrate cleavage activity remains unaffected [2].

Structural architecture of DegQ

Structurally DegQ is very similar to DegP. Crystal structure of DegQ(PDZ2) shows a tetrahedral hollow cage-like 400 kDa dodecameric structure where each trimeric ensemble is arranged at the four corners of the tetrahedron [2]. The planar triangular trimeric structures are connected to each other through an interaction clamp mediated by the peripheral PDZ1 domains. Structure of the DegQ variant comprising solely the protease domain reflects the inactive state of the protein which lacks the interaction clamp and hence is trimeric in nature. The greater flexibility of the loops surrounding the active site in the trimeric protein leads to improper formation of the active site pocket. However, in the dodecameric protein, the loops are held more tightly in place leading to a different conformation of the active site where the catalytic triad (His82, Asp112 and Ser187) is poised for catalysis (Fig. 3A). The loop L3 interacts with PDZ1 and orients the catalytic site for substrate binding and catalysis. Substrate binding at PDZ1 leads to the shifting of the trimer-dodecamer equilibrium towards the latter by stabilizing the higher order oligomeric structure (Fig. 3B). This leads to formation of a proper active site and enhancement of the protease activity by several folds. Thus, a disorder to order transition regulates DegQ activity [2].

Functional implications

DegQ is most active at acidic pH the maximum being around pH 5.5 suggesting it functions as a pH-sensitive protease in the bacte-

rial cell envelope. This property is required by *E. coli* to provide initial proteolytic response to pH-mediated protein misfolding [27].

DegS: tightly regulated sensory protease

DegS, periplasmic stress sensor, is a highly specific and tightly regulated sensory protease of *E. coli*. Binding of unfolded outer membrane porin (OMP) peptides to DegS initiates a proteolytic cascade that cleaves transmembrane protein (RseA) thereby activating σ^E -envelope stress response pathway [28–30].

Structural assembly

Like other members of HtrA family, DegS is a highly stable trimer having an extensive hydrophobic interaction network mediated by serine protease domains of different monomers. The protease domain consists of two perpendicular β -barrel lobes with the catalytic triad (His96, Asp126, and Ser201) located in the crevice between these lobes. This particular arrangement allows forming a fairly exposed active site assembly that is freely accessible to the periplasmic space. In contrast to the other HtrA proteases, where rigid protease domain is joined to a highly mobile PDZ domain, the relative position of the PDZ and protease domains in DegS is fixed by an extended C-terminus. This precise orientation of PDZ appears to be crucial for the regulation of its activity [9].

The crystal structure of the unliganded form of DegS shows that the proteolytic site exists in an inactive state, in which substrate binding as well as catalysis is prevented. This loss of activity is due to interaction of Ser201 with Asn94 as opposed to classical Ser201, His96 interaction for nucleophilic attack leading to a malformed oxyanion hole as shown in Fig. 4A.

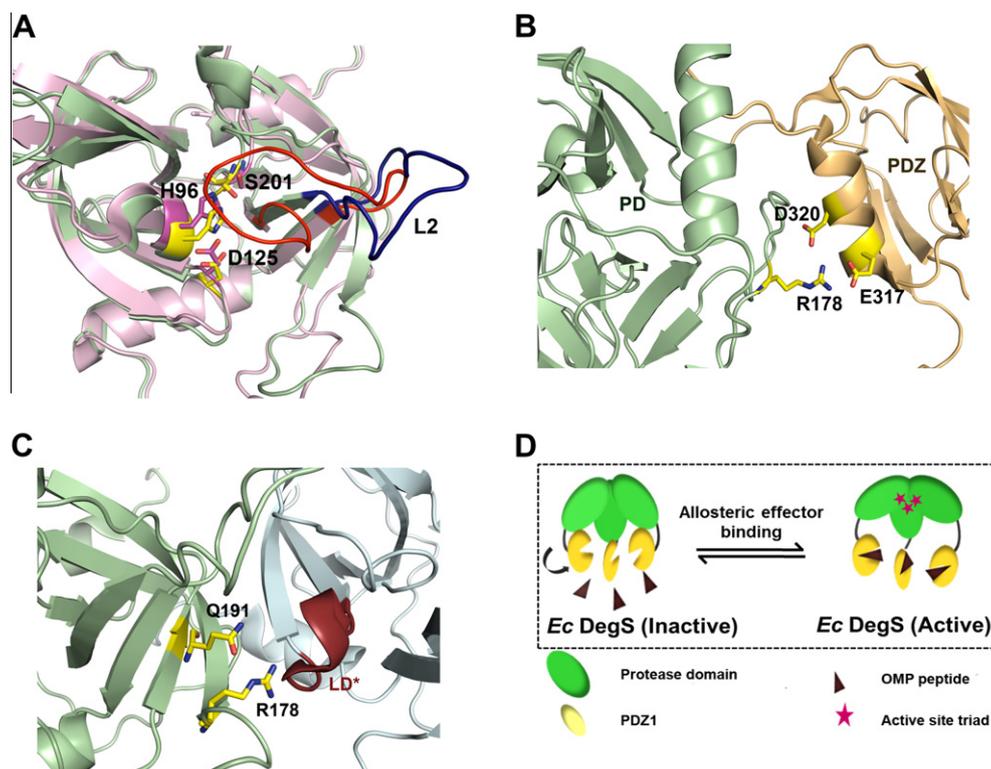


Fig. 4. (A) Structural alignment of inactive and active DegS protease domain. In inactive DegS (PDB ID: 1TE0) (light pink), the active site serine in the catalytic triad (represented by stick) is unavailable for nucleophilic attack and the loop L2 (blue) that forms the specificity pocket is highly disordered. However, an active DegS (PDB ID: 1SOZ) (green) shows properly positioned triad and substrate specificity loop L2 (red). (B) Ribbon diagram representing the intra molecular interaction between the protease and the PDZ domain of peptide free DegS (1TE0). R178 of protease domain (green) pairs with E317 and D320 of PDZ domain (light orange) to stabilize the inactive DegS conformation. (C) Inter subunit interaction in peptide bound DegS (PDB ID: 1SOZ). Residue R178 of loop L3 (represented by stick) on peptide binding, reorients itself so as to establish an interaction with loop LD* (red) of neighboring protease domain (light blue) ultimately forming an active site conformation. (D) Schematic model of DegS activation. Allosteric reversible activation of homotrimer is achieved by binding of the OMP peptides to the PDZ domain thus, relieving inhibitory contacts between PDZ and protease domains, which then induce a conformational rearrangement to form a catalytically active protease.

Mechanism of activation and functional implications

The earlier model of DegS activation proposed by Clausen and co-workers [19] suggested that in absence of protein folding stress, the PDZ domain plays an inhibitory role by directly capturing loop L3 thus preventing its interaction with activation loop LD* of the neighboring subunit. Upon peptide binding, loop L3 re-orientates itself to interact with LD* thereby remodelling the activation domain with a functional catalytic triad, oxyanion hole and substrate specificity pocket. It has been shown that the penultimate residue of the peptide contacts the protease domain to trigger formation of the functional active site. Therefore, in DegS, PDZ is not a simple protein-protein interaction domain but rather performs a novel regulatory function by sensing the misfolded OMPs in the periplasm and transferring an activation signal to the protease domain.

Recent data from Sauer's group, however, contradicted this model stating that the penultimate residue of the peptide is not the principal determinant of DegS activation but rather plays a role in defining affinity for DegS. Accordingly, they proposed that the non-functional DegS conformation in the unbound form is stabilized by the interaction between the regulatory PDZ and protease

domains [28]. The OMP peptide binding then shifts the equilibrium from an inactive/disordered state to an active/ordered state which involves Arg178 of loop L3 mediating this allosteric switch. In unliganded DegS, Arg178 of protease domain forms ion pairs with Glu317 and Asp320 of PDZ domain (Fig. 4B), but in liganded DegS, it makes necessary contacts with LD* and Gln191 forming a stable active site conformation as shown in Fig. 4C [31,32].

Thus, DegS resembles a classical allosteric enzyme where peptide binding to PDZ domain causes proteolytic activation that is reversible (Fig. 4D), allowing cells to rapidly respond to various stress in a flexible manner.

Thermotoga maritima HtrA: a DegQ homolog

Protein quality control in *Thermotoga maritima* (*Tm*), a hyperthermophilic bacterium, is undertaken alone by a DegQ homolog, HtrA that performs dual roles of a molecular chaperone and a protease in temperature dependent manner [33]. This activity of *Tm* HtrA seems to be essential for bacterial thermotolerance and for cell survival at extreme temperatures [1].

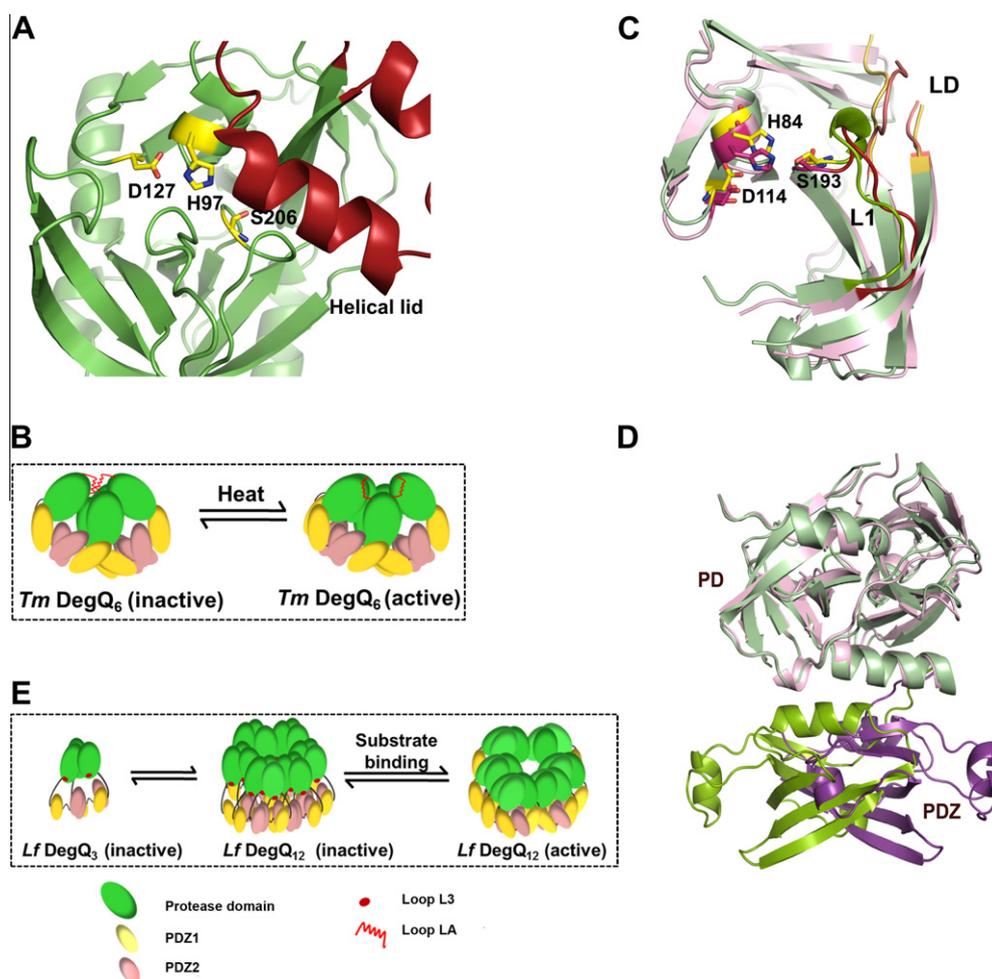


Fig. 5. (A) Ribbon diagram of *Tm* HtrA protease domain (PDB ID: 1L1J). A short helical lid (red) formed by loop LA in the protease domain (green) blocks the catalytic residues (represented by stick model) thus rendering it inactive. (B) Temperature dependent activation model of *Tm* HtrA. At low temperature, the short helix hinders the accessibility of substrate to the active site while at high temperature, helical lid flips off exposing the active site for catalysis. (C) Structural plasticity of the active site in dodecameric *Lf* DegQ. Ribbon diagram of the superimposed protease domain of inactive (light pink) and active (green) DegQ₁₂ highlighting the difference in the orientation of the activation loops L1 and LD. The catalytic triad is shown in stick model (yellow in active and magenta in inactive) in the two protease forms. (D) Superimposed structure of inactive trimeric *Lf* DegQ_{APDZ2} (PDB ID: 3PV4) and active dodecameric DegQ (PDB ID: 3PV3). The PDZ1 of the trimer (purple) is rotated by approximately 180° relative to the protease domain of the dodecamer (green). The PDZ2 domain of *Lf* DegQ₁₂ is not shown. (E) Schematic model of *Lf* DegQ activation. The rearrangement of the PDZ1 domain might lead to re-association of the trimers to form a 12-mer structure thereby trapping the substrate either during re-association or through its interaction with preassembled dodecamer.

Structural architecture and mechanism of activation

The crystal structure of *Tm* N-terminal protease domain (PD) revealed that it retains both the proteolytic and chaperoning activity [33]. The resting state of *Tm* HtrA PD is a trimer as compared to the hexameric *E. coli* HtrA. The main difference lies in the size and conformation of loop LA which in *E. coli* HtrA, interacts with the opposite subunit to form a dimer of loosely attached trimers while in *Tm* HtrA, the loop LA is very short and forms a helical lid that covers the active site rendering it inactive (Fig. 5A).

Compared to other HtrA and DegQ homologs, *Tm* HtrA has a higher sensitivity towards temperature dependent changes. At elevated temperature, movement of the helical lid exposes the active site (Fig. 5B) resulting in enhanced proteolytic activity [34]. However, little or no information is available on how the resting state of *Tm* HtrA mediates chaperone activity. Structure determination in presence of PDZ domain might shed more light on the interplay of the dual mechanism of this DegQ homolog.

Legionella fallonii, DegQ: functional complement of *E. coli* DegP

Many bacterial genome encodes only DegQ (an analog of *E. coli* HtrA-DegP) to maintain protein homeostasis in the periplasm. Like DegP, it consists of an N-terminal protease domain and two C-terminal PDZ domains (PDZ1 and PDZ2). This structural similarity puts forth a question whether DegQ could replace the resting hexameric DegP to deal with the protein folding stress in periplasmic space of the cell. Hansen and co-workers, for the first time, reported the oligomeric structure of DegQ and its variants from *Legionella fallonii* (*Lf*) and proposed the mechanistic model of activation and regulation of this homolog [35].

Structural determinants of *Lf* DegQ₃ and DegQ₁₂

The structures of the dodecameric form of *Lf* DegQ in the unliganded and liganded forms have been solved [35]. The dodecamer represents the protease-active form of *Lf* DegQ, while the chaperone-like activity is independent of DegQ₁₂ formation. Loop LA (critical for formation of higher order oligomer in *E. coli* DegP) deletion does not affect formation of the dodecamer in solution. However, a variant with deleted PDZ2 forms an inactive trimer, pinpointing the importance of PDZ2 in oligomerization. PDZ2 simultaneously interacts with three adjacent trimers via contacts to two PDZ2, one PDZ1, and one protease domain establishing a tightly interconnected stable network to form a dodecameric structure. In the unliganded dodecamer, the S1 specificity pocket is distorted with malformed oxyanion hole. However, upon ligand binding, it forms a distinctively arranged catalytic triad (S193, H84, and D114) and well-formed oxyanion hole representing the active conformation of the protease (Fig. 5C).

A small fraction of *Lf* DegQ is present as DegQ₃, the active site of which is freely accessible to the substrates for proteolysis. This could have led to uncontrolled degradation of the periplasmic proteins. However, crystal structure of DegQ Δ PDZ2, a trimer, reveals that PDZ1 is rotated 180° relative to protease domain placing it opposite to active site as shown in Fig. 5D and thereby restricting the access to the active site and its proteolytic activity [35].

Working model of activation and regulation

Hansen and co-workers presented the activation model of *Lf* DegQ based upon their crystallographic, biochemical and mutagenic data [35] which corroborated with the activation mechanism described earlier in this review for *E. coli* DegS [32] and DegP24 [5]. Briefly, in the absence of the substrates, the protease is in an inactive conformation. Upon substrate binding, disorder-to-order transition of the loops L3, LD, L2 and L1 triggers the activation cascade

forming proteolytically active protease. However, there are still several unanswered ends to this chaperone-protease complex that need further investigation. Firstly, how do DegQ₃ acts as a chaperone although it does not form a higher order oligomer as normally observed in HtrA proteins (such as DegP) with chaperoning functions? Secondly, the details of the allosteric control in activation of DegQ and the factors that determine the fate of the substrates need to be elucidated.

Mycobacterium HtrAs

Apart from Gram negative bacteria, HtrA homologs are also present in Gram positive mycobacterium where they have been shown to be important for virulence [36–38]. In *M. tuberculosis* (*Mtb*), three genes encode HtrA serine proteases: RV1223 (HtrA1), RV0983 (HtrA2), and RV0125 (HtrA3). *Mtb* HtrA1 forms integral part of the transmembrane [39] while the localization of other two proteases is still unclear.

Mtb HtrA2: a constitutively active protease

Although three HtrA proteins have been identified in *Mtb*, structure of only HtrA2 at 2.0 Å resolution has been determined while the other two are not very well characterized. Uniqueness of *Mtb* HtrA2 lies in the fact that it is constitutively active with a well formed active-site and oxyanion hole. Like *E. coli* DegP, it can act both as a chaperone and a protease [7].

Structural architecture and functional implications

Recently, the crystal structure of *Mtb* HtrA2 has been solved by Sacchettini and co-workers [7], providing the first evidence of structural and physiological significance of the protease. Like other HtrAs, the N-terminus consists of a protease domain followed by a single PDZ domain at the C-terminal end. The catalytic triad comprises Ser317-His197-Asp236 that is situated at the center of the protease domain and is freely accessible to the solvent. The structural data shows presence of two auto-proteolyzed products, one bound to the substrate binding pocket and the other to the peptide binding groove in PDZ. The latter establishes a network of non-bonded interactions with loop L3 and reorients it so as to make contacts with loop LD* of the neighboring subunit, which consequently leads to formation of proper active site (Fig. 5E). Hence, this might be one of the plausible causes of *Mtb* HtrA2 being naturally active [7]. Thus, unlike *E. coli* DegP and DegS, *Mtb* HtrA2 is a constitutively active protease [5,9].

Despite the absence of higher oligomeric forms as seen in *E. coli* DegP, *Mtb* HtrA2 has the ability to act as a chaperone protein. This might be due to the presence of a shorter LA loop as compared to that in *E. coli* DegP. It has also been observed that the deletion of HtrA2 gene leads to attenuated virulence in a mouse model of TB [7]. Hence, extensive analysis needs to be carried out to understand the auto regulatory mechanism that controls the protease and chaperone activities required for HtrA-associated virulence of *Mtb*.

Eukaryotic HtrAs

Plant protease Deg1: pH sensor protease

Plant HtrA homologs of *Arabidopsis thaliana* (Deg1, Deg2, Deg5, Deg7 and Deg8) are localized in the chloroplast and play important roles in maintenance of the photosynthetic machinery [11,40]. These HtrAs sense photodamage caused to D1 subunit of photosystem II (PS II) by reactive oxygen species generated across the thylakoid membrane during the solar energy driven electron transfer [41–43]. Thus these plant proteases form a part of the protein quality control system in the thylakoid membrane by acting as chaper-

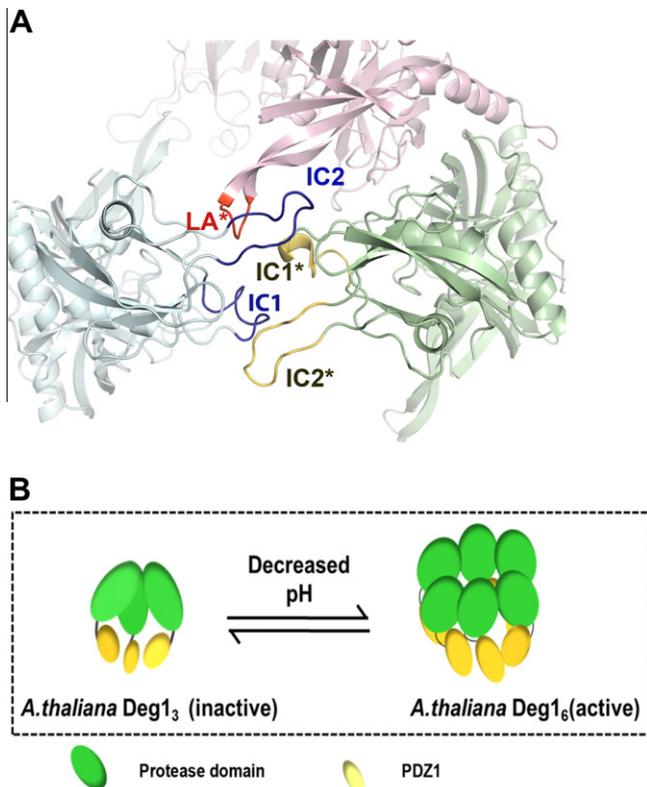


Fig. 6. (A) Hexameric structure of Deg1 (PDB ID: 3QO6). Ribbon structure of active hexameric Deg1 formed by immobilized PDZ domains through multiple inter-subunit contacts between loop LA and interaction clamps (IC) 1 and 2. The protease domain is shown in pink and PDZ domains from the same and opposite monomers are shown in light blue and green respectively, loop LA in red, IC1 and 2 in blue and yellow respectively. (B) Schematic model of Deg1 activation. pH dependent oligomeric switch converts the inactive trimer into an active hexameric form for subsequent proteolysis.

ones during assembly of PS II and as proteases during degradation of damaged photosystem centers [44,45]. Of these HtrAs, Deg1 is the best characterized member of this plant protease family. Recently Clausen and co-workers [46] solved the crystal structure of Deg1 in its hexameric form providing insights into mechanistic and physiological characteristics of this sensor protease.

Structural details and activation model

Deg1, like other HtrAs, is made up of a protease and a C-terminal PDZ domains connected by a short linker segment. The catalytic triad, which comprises His173, Asp203 and Ser282, is situated between the two perpendicular β -barrel lobes of the protease domain. The unique feature of Deg1 PDZ domain is presence of additional loop segments and interaction clamps (IC) 1 & 2 that are critical for mediating the hexameric assembly. The hexamer is stabilized through multiple interactions between protease (loop LA) and PDZ domains (IC1 and IC2) that originate from neighboring subunits of the opposing trimers thereby leading to immobilization of the otherwise flexible PDZ domain (Fig. 6A).

The activation of protease Deg1 relies on proper positioning of the PDZ domain that can further establish the intricate PDZ-L3-LD* interaction network and activate protease function. This is attained by a pH-dependent switch in the oligomeric state of Deg1 (Fig. 6B). At a lower pH (pH 6.0), the protein is more predisposed towards hexamer formation compared to when it is in pH 8.0 [46]. Hexameric form leads to subsequent enhancement of its proteolytic activity [47] thereby signifying the importance of this oligomeric state.

Functional implications

The pH-dependent studies and extensive mutational analysis along with the structural data envisioned the physiological relevance of Deg1 in photosynthetic pathway. During daylight, the photosynthetic proteins get damaged and the thylakoid membrane acidifies leading to monomer-hexamer transition of Deg1. This subsequently leads to the enhancement of Deg1 proteolytic activity which is required to maintain a quality check on the damaged proteins. On the other hand, at night, the membrane is neutral thus converting Deg1 back to its resting trimeric form [46]. Thus, Deg1 protease represents a novel pH-sensor involved in a functionally diverse role in plant photosynthetic system.

Human HtrAs

Till date four human HtrAs have been identified, HtrA1 (L56, PRSS11), [48,49], HtrA2 (Omi) [50,51], HtrA3 (PRSP) [52] and HtrA4 [53]. Among the HtrA homologs, structures of HtrA1 and HtrA2 have been solved. It has been observed that the N-terminus of all human HtrAs comprise signal peptides as well as sections that are identified as Insulin-like growth factor (IGF) binding and protease inhibitor domains [53]. While HtrA2 is the most well studied member of the protease family, which is a mitochondrial protein with proapoptotic property, lesser information is available on HtrA1 that has been found to be associated with several diseases such as arthritis, cancer, and Alzheimer's disease [54–58]. However, very little information on HtrA3 and no information on HtrA4 are currently available [59]. HtrA3 is the closest homolog of human HtrA1 and is found to be down regulated in ovarian cancer implicating its possible role as a tumor suppressor [60]. However, further studies are needed to identify its substrates and to understand the role played by its PDZ domain.

Human HtrA1: PDZ-independent protease activation

HtrA1 is the first human homolog of DegP identified [48] that shares highly similar trypsin-like serine protease domain and a C-terminal PDZ domain with its other three human counterparts (HtrA2, HtrA3 and HtrA4). It is ubiquitously expressed in normal human adult tissues with an upregulation during pregnancy suggesting its important role in placenta formation and embryogenesis [61,62].

Structural details and mechanism of action. Despite high structural conservation in protease and PDZ domains, certain additional N-terminal regulatory domains render specificity to HtrA1 functions [63]. The N-terminal region of HtrA1 comprises a sequence for secretion and two homology domains. The first domain is homologous to mac25, a gene product related to insulin-like growth factor-binding protein [64], and to follistatin, an activin-binding protein [65], while the second one is similar to bacterial serine protease domain. Along with them a Kazal-type inhibitor motif is observed as well [49].

Ehrmann and co-workers solved the crystal structure of HtrA1 in both active and inactive forms so as to understand the underlying mechanism of its mode of activation and the structural changes and dynamics involved in the process [63]. HtrA1 is a trimeric protein which resembles a flat disk. The protease domain is made up of two β -barrel lobes (β 1– β 6 and β 7– β 12) in which the active site with the catalytic triad comprising His220, Asp250 and Ser328 is embedded. Absence of activity in the inactive variant (S328A) of HtrA1 can be attributed to the disordered arrangement of the activation loops (L1, L2 and LD) (Fig. 7A) which might have negated proper formation of the active site pocket. The structure of the active variant of HtrA1 (residues 158–480) comprising the serine protease and PDZ domains has been solved with a covalently bound synthetic inhibitor peptide DPMFKLboroV to the active site

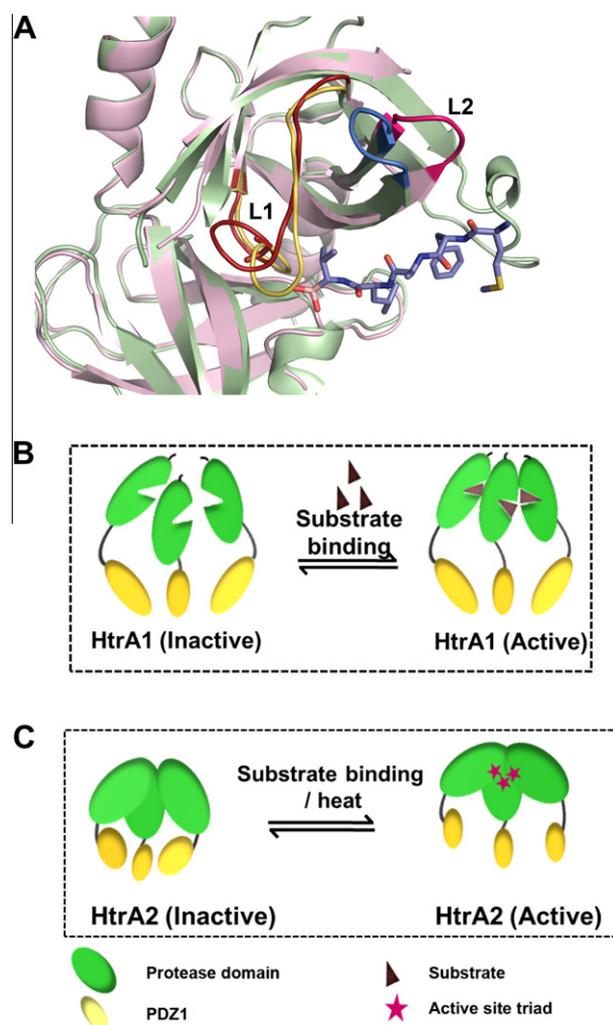


Fig. 7. (A) Structural overlay of inhibitor bound active (PDB ID: 3NZI) and unbound inactive (PDB ID: 3NUM) human HtrA1 protein. The unliganded protease domain (light pink) shows disordered activation domain loops L1 (red) and L2 (magenta) as compared to the ordered loops L1 (yellow) and L2 (blue) in the liganded protease domain (green). The inhibitor is shown as blue stick. (B) Induced fit model for HtrA1 activation. Substrate binding to protease domain re-orientates the loop L3 favoring the formation of a catalytically active conformation. (C) Model for HtrA2 activation. Substrate binding or increased temperature displaces the PDZ domain relieving its inhibitory effect on the protease domain and thus leads to its activation.

serine residue via its boronic acid group. The binding of the inhibitor molecule led to the stabilization of the active site and proper positioning of the catalytic triad. The activation and sensor loops showed significant conformational changes thus allowing formation of a well-defined activation pocket for efficient substrate binding and catalysis. Interestingly, the PDZ domain has been found to be dispensable for substrate binding and catalysis which might be due to the fact that unlike other HtrA proteins substrate directly interacts with loop L3 and rearranges the active site (Fig. 7B). Here, substrate-induced higher order oligomerization (up to 600 kDa) has also been observed. However, the functional relevance of higher order oligomerization and autoactivation of the protein needs further functional and structural analysis.

Functional implications. HtrA1 is a secreted serine protease with a variety of targets most of which are extracellular matrix proteins such as type III collagen, fibronectin and certain components of cartilage [66–68] suggesting its role in promoting degeneration of extracellular matrix components. Recent studies have identified an intracellular substrate of HtrA1 (Tuberous Schlerosis Complex 2

Protein or TSC2) that is involved in cell growth and division [69]. Decrease in HtrA1 expression in melanomas, ovarian, and lung cancers and the observation that its overexpression leads to decrease in cell proliferation suggests its probable role as a tumor suppressor [70–73]. Thus, HtrA1 plays important physiological roles and is also associated with diseases such as arthritis, cancer, familial ischemic cerebral small-vessel disease, age-related macular degeneration and Alzheimer's disease [54–58].

HtrA2: unique mitochondrial proapoptotic protease

HtrA2 (High temperature requirement protease A2) is the unique mitochondrial human protease of the HtrA family with multi-tasking ability. It shows significant sequence homology and structural similarity to its bacterial counterparts DegP and DegQ. HtrA2 has been first identified as an Inhibitor of Apoptosis (IAP) binding protein as it has an IAP recognizing tetrapeptide AVPS motif similar to human Smac/DIABLO and Drosophila death proteins Reaper, Grim, Hid and Sickie [74–76]. Later on however, Shi and co-workers [77] found that it retains its proapoptotic property even in absence of the N-terminal AVPS suggesting IAP binding is not its primary function and the tetrapeptide motif is dispensable for its proapoptotic activity. Proapoptotic property of HtrA2 is manifested through its serine protease activity in both caspase dependent and independent manner [78]. However, understanding of its protease-cum-proapoptotic properties is limited to the few numbers of substrates that have been identified so far including Pea-15, Hax-1 and WARTS kinase [79–81]. A novel substrate binding process, domain plasticity, intricate interdomain networking and complex trimeric structure are together responsible for its regulated protease activity and specificity, the mechanism of which is yet to be fully elucidated.

Structural details and mechanism of action. HtrA2/Omi is expressed as a 49-kDa proenzyme that is targeted primarily to the mitochondrial inner membrane space (IMS), [82,83] where it is attached through its N-terminal transmembrane anchor. During maturation, the first 133 amino acids from the N-terminus get cleaved, exposing the tetrapeptide IAP binding motif (AVPS) [78,82–85]. The mature HtrA2 has a short N-terminal domain (19 residues) and well defined serine protease and PDZ domains. The AVPS motif binds to IAPs and relieves its inhibitory effect on caspase 9 [86]. Shi and co-workers solved the crystal structure of mature form of inactive HtrA2 at 2 Å resolution [77]. The protein is crystallized as a pyramidal trimer (110 kDa) where each monomer comprises 7 α -helices and 19 β -strands. The catalytic triad in HtrA2 includes Ser306, His198 and Asp228 that creates a well-formed catalytic pocket which is partially reflected in the inactive mutant of the crystallized protein (Ser306 → Ala). Unlike many other homologs of HtrA2, it has a single PDZ domain which is required for initial substrate binding at the hydrophobic YIGV groove. Based on the mechanisms of activation of HtrA2 homologs and information from its high resolution structure, a model has been developed by Li et al. that suggest that substrate binding at YIGV leads to a huge conformational change at the PDZ-protease interface. This conformational change removes the inhibitory effect of PDZ from the active site thus enhancing its activity (Fig. 7C). Unfortunately, the flexible linker between the PDZ and protease domain has not been observed in the structure and therefore further studies on its dynamics are required to unambiguously prove the roles of the linker and PDZ in HtrA2 activation and specificity.

Although the crystal structure gives an overview of HtrA2's domain architecture and overall mechanism of action, some important questions still remain unanswered. It has been observed that HtrA2 is proteolytically active only in its trimeric form and the N-terminal aromatic residues are important for mediating the intermolecular oligomerization network but how the structural

rearrangement occurs in the active form upon trimerization is yet to be elucidated. Moreover, PDZ binding and trimerization alone are not the only ways of HtrA2 activation since N-terminal IAP binding leads to HtrA2 activation and IAP cleavage as well [78,87]. Thus, the structural rearrangements upon trimerization and substrate binding at the protease domain that involves the loops and the active site pocket need to be characterized. High resolution structures of the substrate-bound form of the trimeric protein as well as monomeric variant of HtrA2 need to be solved to understand the role of trimerization and PDZ domain better. Moreover, further characterization of its mode of activation upon N-terminal substrate binding will help establish a unified model of its mode of activation and specificity.

Functional implications. HtrA2 has been found to be involved in both caspase-dependent and independent apoptosis. Upon receiving apoptotic signal, HtrA2 is released in the cytosol, where it interacts with IAPs through its N-terminal tetrapeptide (AVPS) motif and relieves inhibition on upstream caspase-9. Furthermore, it cleaves several anti-apoptotic proteins such as Pea-15 and Hax-1 in a caspase-independent manner [79,80]. Apart from its proapoptotic functions, HtrA2 has been implicated in neurodegenerative disorders such as Parkinsons and Alzheimers [88]. Although, literature suggests that HtrA2 might be upregulated during unfolded protein response [89] further studies are needed to confirm its role as a mitochondrial chaperone.

HtrA: structures at a glance

Crystal structures of prokaryotic and eukaryotic HtrA family of proteins [3–5,7,9,27,28,33,35,46,63,77] have contributed tremendously toward providing molecular cues and understanding the global functioning of these PDZ containing serine proteases. Proteins of this family share 50–60% sequence similarity and have unique structural organization. However, their resting states and active forms appear to be quite different, plausibly due to the differences in their cellular localization, functions and behavior under different stress conditions.

Overall structures of the protease domains of HtrA turn out to be similar with an average r.m.s.d. of 0.611 Å (Fig. 8A). The r.m.s.d. among HtrAs mainly arise due to structural differences near the activation domain loops L1, L2, L3 and LA which are involved in the activation cascade. The basic unit of these proteases is a trimeric structure formed due to close hydrophobic packing between protease domains of different monomers [53]. However, loop LA-LA* and PDZ-PDZ* interactions result in the formation of higher multimers/oligomers.

Oligomeric property. The difference in the multimeric state of the HtrAs arises due to varied length and conformation of the loop LA. *E. coli* HtrA DegP has a long LA loop (residues 36–81) reaching the active site of the opposite trimer that forms a hexamer [9] while the other members excluding *Lf* DegQ form a stable trimer due to a short LA loop [3–5,7,9,27,28,33,35,46,63,77]. The PDZ-PDZ* interactions among the members with two PDZ domains (DegP and DegQ) determine the oligomeric nature of the proteins. *Lf* DegQ is dodecamer in solution due to PDZ2 mediated oligomerization with adjacent trimers [35], whereas the *E. coli* DegQ is a trimer that undergoes conversion to 12-mer or 24-mer complexes by peripheral PDZ1 domain-mediated contacts on either substrate binding or under acidic pH conditions [27]. *E. coli* DegP₁₂ and DegP₂₄ are formed only transiently, in presence of misfolded proteins [5], whereas *Lf* DegQ₁₂ assembles irrespective of substrate binding and represents the major oligomeric species over a broad range of conditions [35].

PDZ domain: initiator or regulator?

PDZ domains are the protein-protein interaction modules acting via the conserved GLGF carboxylate-binding motif [90–92]. HtrA family of serine proteases contain one or two C-terminal PDZ domains. Recent structural insights into HtrA proteases have revealed that PDZ domains perform a variety of functions. For example, PDZ domain of *E. coli* DegS interacts with the C-terminal signal peptides of OMP and controls the allosteric switch of cellular stress response pathway [31,32]. In DegP, two PDZ domains function as proteolytic regulators, acting as gatekeepers of the inner proteolytic-chaperone chamber [4,5]. However, in DegQ, PDZ1 and PDZ2 are not re-

Table 2

Distances between the atoms of the catalytic triad of HtrA proteins (in Å).

Protein	N ^{ε2} (His)–O ^γ (Ser)		ND1(His)–OD1(Asp)	
	Active	Inactive	Active	Inactive
<i>Ec</i> DegP	3.5	5.7	3.6	3.4
<i>Ec</i> DegS	3.1	4.4	3.5	2.6
<i>Ec</i> DegQ	2.8	3.7	3.2	3.1
<i>Lf</i> DegQ	3.4	3.5	3.3	3.0
<i>Tm</i> DegQ	–	3.5	–	3.8
<i>h</i> HtrA1	2.9	9.6	2.4	2.6
<i>At</i> Deg1	3.2	–	3.8	–

Å: Angstrom, *Ec*: *Escherichia coli*, *Lf*: *Legionella fallonii*, *Tm*: *Thermotoga maritima*, *h*: *Homo sapiens*, *At*: *Arabidopsis thaliana*.

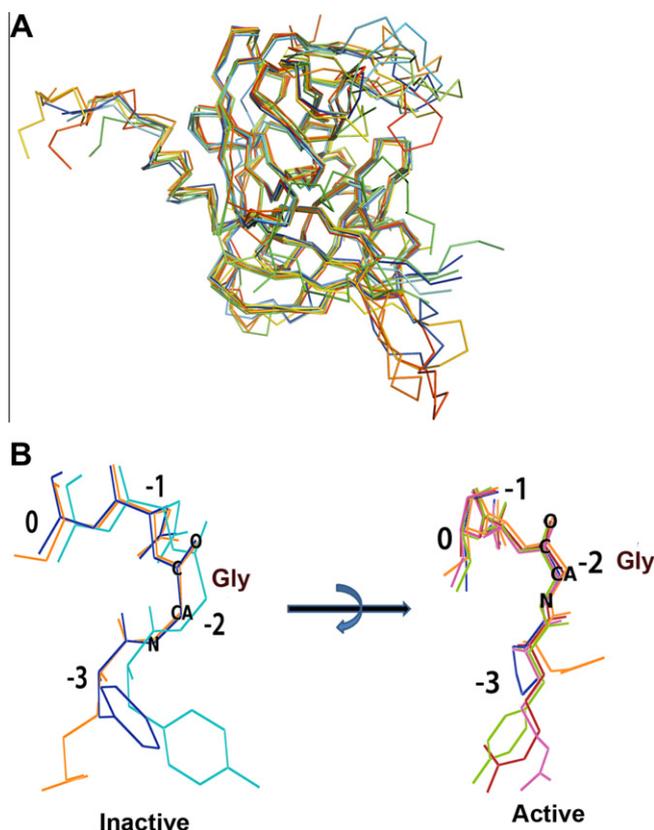


Fig. 8. (A) Structural alignment of the protease domains of HtrA proteases. The overall structure of the protease domain in the HtrA proteins is similar with an average r.m.s.d. of 0.61 Å for C α atoms. (B) Comparison between the oxyanion-hole residues of the inactive (DegP₆, HtrA2, HtrA1, and Deg1) and active forms (DegP₂₄, DegS, *E. coli* DegQ, HtrA1, and Deg1) of the HtrA proteases. The positions of residues in the oxyanion hole are denoted as 0 for active site Ser, –1, –2, and –3 for the preceding residues. Distortion of the oxyanion hole could be observed in the inactive form especially in the amide nitrogen of the Gly at –2.

lated to the regulatory activity, but are involved in forming higher order oligomers as described earlier in this review [27,35].

Till date, all well characterized eukaryotic HtrAs have been found to contain a single PDZ domain. PDZ domain of Deg1 shows the presence of an additional interaction clamp that mediates multiple interactions between protease and PDZ domains thus, resulting in the formation of a stable hexameric structure. In human HtrA1, the PDZ domain has been found to be dispensable for its proteolytic activity [63], whereas in HtrA2, the highly flexible PDZ domain might act as a regulator of proteolytic activity by blocking the active site of the trimer [77]. Overall, PDZ domains are essential modulators of this widely conserved family of HtrA serine proteases.

Active site conformation: active versus inactive proteases catalytic triad

Catalytic triad

In serine proteases, hydrolytic cleavage of the peptide bonds occurs when the catalytic triad residues are aligned in a position close enough for electron transfer from Asp to Ser through His. The serine –OH group acts as a nucleophile, attacking the carbonyl carbon of the scissile peptide bond of the substrate subsequently leading to its cleavage [93].

With the information from the structural database, we aligned the catalytic triad residues of inactive and active forms of the HtrA proteases to measure the distances between the N^{ε2} atom of His and the O^γ atom of Ser for each molecule in the asymmetric unit and are shown in Table 2. It has been observed that the distance between the N^{ε2} atom of His and the O^γ atom of Ser in active proteases is <3.5 Å whereas in the inactive form it is >3.5 Å. This might be crucial in preventing the His to function as a general base to remove the proton from active site Ser residue. Therefore, the orientation of the loops that harbor the catalytic triad might render the protease active or inactive.

Oxyanion hole

Another key factor for proteolysis by chymotrypsin like serine proteases is the stabilization of a negative charge of carbonyl oxygen in the reaction intermediate (oxyanion hole). This is achieved by hydrogen bond formation by the amide nitrogen atoms of two peptide bonds with the protease backbone [93].

The alignment of the residues forming the oxyanion hole in the active and the inactive forms of the protease is shown in Fig. 8B. In the active form, the nitrogen atom of the Gly (-2 position) acts a hydrogen donor to the putative oxyanion hole whereas in the inactive form, it faces out in the opposite direction creating a malformed oxyanion hole.

Activation mechanism

HtrA group of serine proteases form a large family with unique structural organization and overall low sequence conservation. The C-terminal PDZ domain of HtrAs is responsible for mediating its important cellular functions. The resting state of the protease has a basal activity that is subsequently triggered on activation by multiple mechanisms. Interestingly, in most of the cases, PDZ domains act as modulators of HtrA activity by binding to their substrates which subsequently lead to protease activation. This is performed either by transforming the inactive resting state to higher order oligomeric active state (Figs. 2C, 3B, 5B, E and 6B) or by allosterically transferring the signal from PDZ domain to the mechanically

important loops (Figs. 4D, 7B and C), thus further re-orienting the activation loops to form the catalytically active protease.

Concluding remarks

HtrA family of serine proteases has evolved to play a critical role in protein quality control. In bacterial system, these proteins have dual chaperone-protease activity whereas in humans they are involved in a variety of physiological functions as well as implicated in several diseases. In this review, we have carried out an overall comparison of HtrA proteins from different organisms whose structures have been solved with an aim at highlighting the importance of their unique structural signatures and conformational dynamics in functional diversity.

Although, a lot of studies have been done on bacterial DegP and DegS and some work on human HtrA2 proteins, a lot more needs to be done to elucidate their mechanism of action and mode of regulation in the cell. Understanding of the physiological roles of HtrAs has been limited by the number of substrates that have been identified till date. Therefore, extensive proteomics and biochemical studies need to be done to identify and characterize HtrA substrates and binding partners. Moreover, structures of other HtrA proteins need to be solved to develop an understanding of the structure-function relationship of the serine protease in general.

Given their involvement in various physiological and pathophysiological pathways, HtrA proteins have the potential to be valuable targets for novel therapeutic drugs.

Therefore, high resolution structures of substrate bound HtrA proteins along with functional studies will be a step forward toward structure-based design of novel molecules with desired characteristics to modulate HtrA activity.

Acknowledgment

We thank Mr. Tanmoy Bhattacharjee for his intellectual inputs and critical feedback on the manuscript.

References

- [1] B. Lipinska, O. Fayet, L. Baird, C. Georgopoulos, *J. Bacteriol.* 171 (1989) 1574–1584.
- [2] T. Clausen, M. Kaiser, R. Huber, M. Ehrmann, *Nat. Rev. Mol. Cell. Biol.* 12 (2011) 152–162.
- [3] J. Jiang, X. Zhang, Y. Chen, Y. Wu, Z. Zhou, Z. Chang, S. Sui, *Proc. Natl. Acad. Sci. USA* 105 (2008) 11939–11944.
- [4] T. Krojer, M. Garrido-Franco, R. Huber, M. Ehrmann, T. Clausen, *Nature* 416 (2002) 455–459.
- [5] T. Krojer, J. Sawa, E. Schäfer, H. Saibil, M. Ehrmann, T. Clausen, *Nature* 453 (2008) 885–890.
- [6] F. Li, P. Jeffrey, J. Yu, Y. Shi, *J. Biol. Chem.* 281 (2006) 2960–2968.
- [7] N. Mohamedmohaideen, S. Palaninathan, P. Morin, B. Williams, M. Braunstein, S. Tichy, J. Locker, D. Russell, W.J. Jacobs, J. Sacchettini, *Biochemistry* 47 (2008) 6092–6102.
- [8] Q. Shen, X. Bai, L. Chang, Y. Wu, H. Wang, S. Sui, *Proc. Natl. Acad. Sci. USA* 106 (2009) 4858–4863.
- [9] C. Wilken, K. Kitzing, R. Kurzbauer, M. Ehrmann, T. Clausen, *Cell* 117 (2004) 483–494.
- [10] D. Kim, K. Kim, *J. Biochem. Mol. Biol.* 38 (2005) 266–274.
- [11] P. Huesgen, H. Schuhmann, I. Adamska, *Res. Microbiol.* 160 (2009) 726–732.
- [12] D. Zurawa-Janicka, J. Skorko-Glonek, B. Lipinska, *Expert Opin. Ther. Targets* 14 (2010) 665–679.
- [13] Y.T.-N. Suzuki, K.T. Akagi, T. Hashikawa, R. Takahashi, *Cell Death Differ.* 11 (2004) 208–216.
- [14] I. Polur, P. Lee, J. Servais, L. Xu, Y. Li, *Histol. Histopathol.* 25 (2010) 599–608.
- [15] M. Meltzer, S. Hasenbein, N. Mamant, M. Merdanovic, S. Poepsel, P. Hauske, M. Kaiser, R. Huber, T. Krojer, T. Clausen, M. Ehrmann, *Res. Microbiol.* 160 (2009) 660–666.
- [16] C. Spiess, A. Beil, M. Ehrmann, *Cell* 97 (1999) 339–347.
- [17] J. Skórko-Glonek, A. Wawrzynów, K. Krzewski, K. Kurpierz, B. Lipińska, *Gene* 163 (1995) 47–52.
- [18] A. Jomaa, D. Damjanovic, V. Leong, R. Ghirlando, J. Iwanczyk, J. Ortega, *J. Bacteriol.* 189 (2007) 706–716.
- [19] T. Krojer, J. Sawa, R. Huber, T. Clausen, *Nat. Struct. Mol. Biol.* 17 (2010) 844–852.

- [20] M. Merdanovic, N. Mamant, M. Meltzer, S. Poepsel, A. Auckenthaler, R. Melgaard, P. Hauske, L. Nagel-Steger, A. Clarke, M. Kaiser, R. Huber, M. Ehrmann, *Nat. Struct. Mol. Biol.* 17 (2010) 837–843.
- [21] J. Skórko-Glonek, D. Zurawa, F. Tanfani, A. Scirè, A. Wawrzynów, J. Narkiewicz, E. Bertoli, B. Lipińska, *Biochim. Biophys. Acta* 1649 (2003) 171–182.
- [22] A. Jomaa, J. Iwanczyk, J. Tran, J. Ortega, *J. Bacteriol.* 191 (2009) 1924–1932.
- [23] J. Hasenbein, V. Leong, J. Ortega, *PLoS One* 6 (2011) e18944.
- [24] T. Krojer, K. Pangerl, J. Kurt, J. Sawa, C. Stingl, K. Mechtler, R. Huber, M. Ehrmann, T. Clausen, *Proc. Natl. Acad. Sci. USA* 105 (2008) 7702–7707.
- [25] H. Kolmar, P. Waller, R. Sauer, *J. Bacteriol.* 178 (1996) 5925–5929.
- [26] S. Bass, Q. Gu, A. Christen, *J. Bacteriol.* 178 (1996) 1154–1161.
- [27] J. Sawa, H. Malet, T. Krojer, F. Canellas, M. Ehrmann, T. Clausen, *J. Biol. Chem.* 286 (2011) 30680–30690.
- [28] J. Sohn, R. Grant, R. Sauer, *Cell* 131 (2007) 572–583.
- [29] N. Walsh, B. Alba, B. Bose, C. Gross, R. Sauer, *Cell* 13 (2003) 61–71.
- [30] S. Hasenbein, M. Meltzer, P. Hauske, M. Kaiser, R. Huber, T. Clausen, M. Ehrmann, *J. Mol. Biol.* 397 (2010) 957–966.
- [31] J. Sohn, R. Grant, R. Sauer, *Structure* 17 (2009) 1411–1421.
- [32] J. Sohn, R. Sauer, *Mol. Cell* 33 (2009) 64–74.
- [33] D. Kim, D. Kim, S. Ha, N. Lokanath, C. Lee, H. Hwang, K. Kim, *J. Biol. Chem.* 278 (2003) 6541–6543.
- [34] D. Kim, E. Kwon, Y. Shin, D. Kweon, K. Kim, *J. Mol. Biol.* 377 (2008) 410–420.
- [35] R. Wrase, H. Scott, R. Hilgenfeld, G. Hansen, *Proc. Natl. Acad. Sci. USA* 108 (2011) 10490–10495.
- [36] G. Cortés, B. de Astorza, V. Benedí, S. Albertí, *Infect. Immun.* 70 (2002) 4772–4776.
- [37] K. Williams, P. Oyston, N. Dorrell, S. Li, R. Titball, B. Wren, *FEMS Microbiol. Lett.* 186 (2000) 281–286.
- [38] B. Villarreal-Ramos, J. Manser, R. Collins, V. Chance, D. Eckersall, P. Jones, G. Dougan, *Microbiology* 146 (2001) 2775–2783.
- [39] Y. Xiong, M. Chalmers, F. Gao, T. Cross, A. Marshall, *J. Proteome. Res.* 4 (2005) 855–861.
- [40] H. Itzhaki, L. Naveh, M. Lindahl, M. Cook, Z. Adam, *J. Biol. Chem.* 273 (1998) 7094–7098.
- [41] J. De Las Rivas, B. Andersson, J. Barber, *FEBS Lett.* 301 (1992) 246–252.
- [42] E. Aro, S. McCaffery, J. Anderson, *Plant Physiol.* 103 (1993) 835–843.
- [43] N. Adir, H. Zer, S. Shochat, I. Ohad, *Photosynth. Res.* 76 (2003) 343–370.
- [44] E. Kapri-Pardes, L. Naveh, Z. Adam, *Plant Cell* 19 (2007) 1039–1047.
- [45] X. Sun, M. Ouyang, J. Guo, J. Ma, C. Lu, Z. Adam, L. Zhang, *Plant J.* 62 (2010) 240–249.
- [46] J. Kley, B. Schmidt, B. Boyanov, P. Stolt-Bergner, R. Kirk, M. Ehrmann, R. Knopf, L. Naveh, Z. Adam, T. Clausen, *Nat. Struct. Mol. Biol.* 18 (2011) 728–731.
- [47] Y. Chassin, E. Kapri-Pardes, G. Sinvany, T. Arad, Z. Adam, *Plant Physiol.* 130 (2002) 857–864.
- [48] J. Zumbunn, B. Trueb, *FEBS Lett.* 398 (1996) 187–192.
- [49] S. Hu, M. Carozza, M. Klein, P. Nantermet, D. Luk, R. Crowl, *J. Biol. Chem.* 273 (1998) 34406–34412.
- [50] L. Faccio, C. Fusco, A. Chen, S. Martinotti, J. Bonventre, A. Zervos, *J. Biol. Chem.* 275 (2000) 2581–2588.
- [51] C. Gray, R. Ward, E. Karran, S. Turconi, A. Rowles, D. Viglienghi, C. Southan, A. Barton, K. Fantom, A. West, J. Savopoulos, N. Hassan, H. Clinkenbeard, C. Hanning, B. Amegadzie, J. Davis, C. Dingwall, G. Livi, C. Creasy, *Eur. J. Biochem.* 267 (2000) 5699–5710.
- [52] G. Nie, A. Hampton, Y. Li, J. Findlay, L. Salamonsen, *Biochem. J.* 371 (2003) 39–48.
- [53] T. Clausen, C. Southan, M. Ehrmann, *Mol. Cell.* 10 (2002) 443–455.
- [54] J. Milner, A. Patel, A. Rowan, *Arthritis Rheum.* 58 (2008) 3644–3656.
- [55] J. Chien, M. Campioni, V. Shridhar, A. Baldi, *Curr. Cancer Drug Targets* 9 (2009) 451–468.
- [56] K. Hara, A. Shiga, T. Fukutake, H. Nozaki, A. Miyashita, A. Yokoseki, H. Kawata, A. Koyama, K. Arima, T. Takahashi, M. Ikeda, H. Shiota, M. Tamura, Y. Shimoe, M. Hirayama, T. Arisato, S. Yanagawa, A. Tanaka, I. Nakano, S. Ikeda, Y. Yoshida, T. Yamamoto, T. Ikeuchi, R. Kuwano, M. Nishizawa, S. Tsuji, O. Onodera, *N. Engl. J. Med.* 360 (2009) 1729–1739.
- [57] H. Coleman, C. Chan, F. Ferris, E. Chew, *Lancet* 372 (2008) 1835–1845.
- [58] S. Grau, A. Baldi, R. Bussani, X. Tian, R. Stefanescu, M. Przybylsk, P. Richards, S. Jones, V. Shridhar, T. Clausen, M. Ehrmann, *Proc. Natl. Acad. Sci. USA* 102 (2005) 6021–6026.
- [59] M. Bhuiyan, F. Kohji, *Curr. Drug Targets* 10 (2009) 372–383.
- [60] M. Bowden, A. Drummond, P. Fuller, L. Salamonsen, J. Findlay, G. Nie, *Mol. Cell. Endocrinol.* 327 (2010) 13–18.
- [61] A. De Luca, M. De Falco, V. Fedele, L. Cobellis, A. Mastrogiacomio, V. Laforgia, I. Tuduce, M. Campioni, D. Giraldi, M. Paggi, A. Baldi, *J. Histochem. Cytochem.* 52 (2004) 885–892.
- [62] T. Lorenzi, D. Marzioni, S. Giannubilo, A. Quaranta, C. Crescimanno, A. De Luca, A. Baldi, T. Todros, A. Tranquilli, M. Castellucci, *Placenta* 30 (2009).
- [63] L. Truebestein, A. Tennstaedt, T. Mönig, T. Krojer, F. Canellas, M. Kaiser, T. Clausen, M. Ehrmann, *Nat. Struct. Mol. Biol.* 18 (2011) 386–388.
- [64] K. Swisshelm, K. Ryan, K. Tsuchiya, R. Sager, *Proc. Natl. Acad. Sci. USA* 92 (1995) 4472–4476.
- [65] M. Kato, H. Sato, T. Tsukada, Y. Ikawa, S. Aizawa, M. Nagayoshi, *Oncogene* 12 (1996) 1361–1364.
- [66] Murwantoko, M. Yano, Y. Ueta, A. Murasaki, H. Kanda, C. Oka, M. Kawaichi, *Biochem. J.* 381 (2004) 895–904.
- [67] S. Grau, P. Richards, B. Kerr, C. Hughes, B. Catterson, A. Williams, U. Junker, S. Jones, T. Clausen, M. Ehrmann, *J. Biol. Chem.* 281 (2006) 6124–6129.
- [68] A. Tsuchiya, M. Yano, J. Tocharus, H. Kojima, M. Fukumoto, M. Kawaichi, C. Oka, *Bone* 37 (2005) 323–336.
- [69] M. Campioni, A. Severino, L. Manente, I. Tuduce, S. Toldo, M. Caraglia, S. Crispi, M. Ehrmann, X. He, J. Maguire, M. De Falco, A. De Luca, V. Shridhar, A. Baldi, *Mol. Cancer Res.* 8 (2010) 1248–1260.
- [70] A. Baldi, A. De Luca, M. Morini, T. Battista, A. Felsani, F. Baldi, C. Catricalà, A. Amantea, D. Noonan, A. Albin, P. Natali, F. Facciolo, *Oncogene* 21 (2002) 6684–6688.
- [71] J. Chien, J. Staub, S. Hu, M. Erickson-Johnson, F. Couch, D. Smith, R. Crowl, S. Kaufmann, V. Shridhar, *Oncogene* 23 (2004) 1636–1644.
- [72] V. Esposito, M. Campioni, A. De Luca, E. Spugnini, F. Baldi, R. Cassandro, A. Mancini, B. Vincenzi, A. Groeger, M. Caputi, A. Baldi, *Anticancer Res.* 26 (2006) 3455–3459.
- [73] A. Baldi, M. Mottolese, B. Vincenzi, M. Campioni, P. Mellone, M. Di Marino, V. di Crescenzo, P. Visca, S. Menegozzo, E. Spugnini, G. Citro, A. Ceribelli, A. Mirri, J. Chien, V. Shridhar, M. Ehrmann, M. Santini, F. Facciolo, *Pharmacogenomics* 9 (2008) 1069–1077.
- [74] G. Wu, J. Chai, T. Suber, J. Wu, C. Du, X. Wang, Y. Shi, *Nature* 408 (2000) 1008–1012.
- [75] Z. Liu, C. Sun, E. Olejniczak, R. Meadows, S. Betz, T. Oost, J. Herrmann, J. Wu, S. Fesik, *Nature* 408 (2000) 1004–1008.
- [76] Y. Shi, *Cell Death Differ.* 9 (2002) 93–95.
- [77] W. Li, S. Srinivasula, J. Chai, P. Li, J. Wu, Z. Zhang, E. Alnemri, Y. Shi, *Nat. Struct. Mol. Biol.* 9 (2002) 436–441.
- [78] Y. Suzuki, Y. Imai, H. Nakayama, K. Takahashi, K. Takio, R. Takahashi, *Mol. Cell.* 8 (2001) 613–621.
- [79] A. Trecia, F. Fiory, M. Maitan, P. Vito, A. Barbagallo, A. Perfetti, C. Miele, P. Ungaro, F. Oriente, L. Cilenti, A. Zervos, P. Formisano, F. Beguinot, *J. Biol. Chem.* 279 (2004) 46566–46572.
- [80] L. Cilenti, M. Soundarapandian, G. Kyriazis, V. Stratico, S. Singh, S. Gupta, J. Bonventre, E. Alnemri, Z. AS, *J. Biol. Chem.* 279 (2004) 50295–50301.
- [81] S. Kuninaka, S. Iida, T. Hara, M. Nomura, H. Naoe, T. Morisaki, M. Nitta, Y. Arima, T. Mimori, S. Yonehara, H. Saya, *Oncogene* 26 (2007) 2395–2406.
- [82] R. Hegde, S. Srinivasula, Z. Zhang, R. Wassell, R. Mukattash, L. Cilenti, G. DuBois, Y. Lazebnik, A. Zervos, T. Fernandes-Alnemri, E. Alnemri, *J. Biol. Chem.* 277 (2002) 432–438.
- [83] L. Martins, I. Iaccarino, T. Tenev, S. Gschmeissner, N. Totty, N. Lemoine, J. Savopoulos, C. Gray, C. Creasy, C. Dingwall, J. Downward, *J. Biol. Chem.* 277 (2002) 439–444.
- [84] G. van Loo, M. van Gorp, B. Depuydt, S. Srinivasula, I. Rodriguez, E. Alnemri, K. Gevaert, J. Vandekerckhove, W. Declercq, P. Vandenebeele, *Cell Death Differ.* 9 (2002) 20–26.
- [85] A. Verhagen, J. Silke, P. Ekert, M. Pakusch, H. Kaufmann, L. Connolly, C. Day, A. Tikoo, R. Burke, C. Wrobel, R. Moritz, R. Simpson, D. Vaux, *J. Biol. Chem.* 277 (2002) 445–454.
- [86] G. van Loo, X. Saelens, F. Matthijssens, P. Schotte, R. Beyaert, W. Declercq, P. Vandenebeele, *Cell Death Differ.* 9 (2002) 1207–1211.
- [87] L. Martins, B. Turk, V. Cowling, A. Borg, E. Jarrell, L. Cantley, J. Downward, *J. Biol. Chem.* 278 (2003) 49417–49427.
- [88] J. Jones, P. Datta, S. Srinivasula, W. Ji, S. Gupta, Z. Zhang, E. Davies, G. Hajnoczky, T. Saunders, M. Van Keuren, T. Fernandes-Alnemri, M. Meisler, E. Alnemri, *Nature* 425 (2003) 721–725.
- [89] L. Papa, D. Germain, *J. Cell Sci.* 124 (2011) 1396–1402.
- [90] D. Doyle, A. Lee, J. Lewis, E. Kim, M. Sheng, R. MacKinnon, *Cell* 85 (1996) 1067–1076.
- [91] B. Harris, B. Hillier, W. Lim, *Biochemistry* 40 (2001) 5921–5930.
- [92] B. Harris, W. Lim, *J. Cell Sci.* 114 (2001) 3219–3231.
- [93] J. Kraut, *Annu. Rev. Biochem.* 46 (1977) 331–358.