Mutant generation of Mms2 and structural analysis of modeled Uev1a N-terminus docked to Uev1a-Ubc13-CHIP heterotrimer

Mahimalakshmi Mohandas

BO13M1005

A Dissertation Submitted to

Indian Institute of Technology Hyderabad

In Partial Fulfillment of the Requirements for

The Degree of Master of Technology

भारतीय प्रौद्योगिकी संस्थान हैदराबाद

Indian Institute of Technology Hyderabad

Department of Biotechnology

June, 2015

Declaration

I declare that this written submission represents my ideas in my own words, and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources that have thus not been properly cited, or from whom proper permission has not been taken when needed.

Mahimalakshmi Mohandas

BO13M1005

Approval Sheet

This thesis entitled "***Mutant generation of Mms2 and structural analysis of modeled Uev1a N-terminus docked to Uev1A-Ubc13-CHIP heterotrimer***" by Mahimalakshmi Mohandas is approved for the degree of Master of Technology from IIT Hyderabad.

N.K Raghavendra

Dr. N.K Raghavendra
Thesis Advisor
Department of Biotechnology
IIT Hyderabad

Basant Kumar Patel

Dr. Basant Kumar Patel
Department of Biotechnology
IIT Hyderabad

Anindya Roy

Dr. Anindya Roy
Department of Biotechnology
IIT Hyderabad

G. Satyanarayana

Dr. G. Satyanarayana
Department of Chemistry
IIT Hyderabad

ACKNOWLEDGEMENT

I express profound gratitude to my supervisor, Dr. N.K Raghavendra for extending his valuable guidance throughout the span of the project and for its accomplishment. I also acknowledge my sincere regards to Faculty members of the Thesis Committee: Dr. Basant Kumar Patel, Dr. Anindya Roy and Dr. G. Satyanarayana for their insightful comments and suggestions.

I would like to thank my lab members Anu George, Manne Aravind Kumar, Sreehari and my colleague Shravan Kumar for the support provided whenever it was necessary.

I also extend a special note of thanks to my Parents and Brother for the moral support that they have incessantly bestowed.

CONTENTS

1. ABBREVIATIONS

BLAST	- Basic Local Search Alignment Local Tool
BME	- β-Mercaptoethanol
CHIP	- C-Terminal peptide from Human Hsp90
DNA	- Deoxyribonucleic Acid
DOPE Score	- Discrete Optimized Protein Energy Score
E1	- Ubiquitin activating enzyme
E2	- Ubiquitin conjugating enzyme
E3	- Ubiquitin Ligase
HSP	- Heat Shock Protein
IPTG	- Isopropyl β-D-1-thiogalactopyranoside
Lys48	- Lysine 48 linkage of Ubiquitin
Lys63	- Lysine 63 linkage of Ubiquitin
Mms2	- Methyl methanosulfonate2
NCBI	- National Center for Biotechnology Information
NEMO	- Nuclear Factor-kB Essential Modulator
NMR	- Nuclear Magnetic Resonance
O.D	- Optical Density
PCNA	- Proliferating Cell Nuclear Antigen
PCR	- Polymerase Chain Reaction
PMSF	- Phenylmethylsulfonyl fluoride

Poly-Ub	- Poly-Ubiquitin chains
PRR	- Post-replicative Repair Pathway
RAD	- RADiation Sensitive
RING	- Really Interesting New Gene
SAPKs	- Stress-Activated Protein Kinases
SSBs	- Single Stranded Binding Proteins
TRAF	- Tumor Necrosis Factor (TNF) Receptor Associated Factor
Ubs	- Ubiquitins
UEV	- Ubiquitin-conjugating Enzyme Variant

2. ABSTRACT:

Part I

Mms2, a Ubiquitin-conjugating enzyme variant (UEV) protein acts a co-factor for Ubc13 in DNA damage repair. This signaling is mediated by assembly of Lys-63 linked poly-ubiquitin chains by Ubc13 and Mms2 complex. Uev1A, another UEV in conjugation with Ubc13 forms Lys-63 poly-ubiquitin chains which lead to activation of NF-kB pathway. Hence, in the presence of different cofactors, Ubc13 performs distinct biological functions. Mms2 and Uev1A were found to have a sequence similarity of more than 90%. This project aims at generating five mutants of Mms2 by changing the residues to that of Uev1A and studying their interaction with Ubc13.

Part II

Uev1a is a UEV protein that associates with Ubc13, an E2 enzyme to form Lys63 linked poly-Ub chains. Uev1a possess the core UEV domain but lack the active site cysteine rendering it catalytically inactive. The 35 residue long N-terminus is unstructured in the solution structure of Uev1a. Ubc13-Uev1a heterodimer interact with E3 ligases to ubiquitinate proteins downstream. CHIP and TRAF6 are E3 ligases possessing U box and RING domains respectively which interact with the E2 heterodimer for further ubiquitination of target proteins. There is a lack of structural data on interaction of N-terminal 35 residues of Uev1a with TRAF6 and CHIP. This project aims at understanding the possible interactions of Uev1a N-Terminal residues with the available crystal structure of Ubc13-Uev1a-CHIP after modeling the peptide and docking it onto the heterotrimer complex.

3. INTRODUCTION:

3.1 Ubiquitination:

Post-translational modifications are highly significant for regulation of protein functions. One such fundamental post-translational modification is ubiquitination which involves the covalent attachment of ubiquitin to target protein. Ubiquitin is an 8 kDa regulatory protein, with 76 amino acids terminating in a diglycine sequence (Pickart and Eddins 2004). Ubiquitin harbors seven Lysine residues, K6, K11, K27, K29, K33, K48 and K63.

Mono-ubiquitination involves formation of an isopeptide bond between ubiquitin and lysine residue of target protein. Extension of ubiquitin chains on target protein occurs via one of the seven lysine residues of ubiquitin (Clague and Sylvie 2010).

Ubiquitination involves cascade of three classes of enzymes E1, E2 and E3. E1, Ub-activating enzyme forms a thioester bond between cysteine in its active site and C-Terminus of Ub. Activated Ub then forms similar thioester-linked complex with Ub- conjugating enzyme (E2). Subsequently, Ub is covalently attached to substrate protein by E3, Ub-Ligase (E3). Although E2 and E3 are found more than one in organisms, E1 is one. The higher number of E3 in comparison to E2 can be explained by the observation that E3 acts as the primary substrate recognition factor in the ubiquitination reaction, whereas E2 is thought to be involved in the reaction largely through its association with a given E3 (Finley and Vincent 1991). Specificity of E2/E3 interaction decides the final biological destination of ubiquitin carried by a particular E2 (Pickart and Eddins 2004). Schematic representation of ubiquitination reaction is shown in Figure 1.

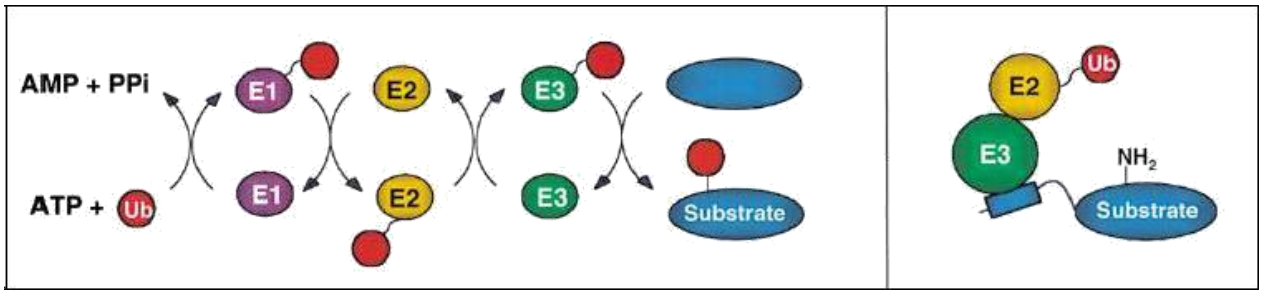


Figure1: **Schematic representation of the ubiquitination cascade**. Ubiquitin is activated by E1, then transferred to an E2 as it forms a thio-ester bond, and transferred by an E3 to a lysine residue of a substrate**.**

Specificity of Ubiquitination: The E3 binds its substrate and its specific E2 partner at separate sites. (Adapted from Pickart 2001)

Poly-ubiquitination of substrates can be linked through different lysines present in ubiquitin. Poly-Ub chains linked through Lys48-G76 isopeptide bonds are the principal signal for proteasomal proteolysis. Poly-ubiquitination through Ub Lys48 requires a chain of at least 4 Ubs linked to the target protein for proteasomal degradation via 26S proteasome (Thrower et al. 2000). Lys63 linked Ub chains can regulate kinase activation, DNA damage tolerance, signal transduction, and endocytosis (Passmore et al. 2004).

3.2 Ubc13 and its function:

UBCs are a class of structurally related proteins and they catalyze transfer of ubiquitin to substrate protein (Hochstrasser & Mark 1996). Ubc13 is an E2 enzyme that has been involved in the assembly of Lys-63 linked poly-ubiquitin chains instead of the conventional Lys-48 assembly. This enzyme in association with various partners, functions in innate immune response, DNA-damage repair, control cell cycle progression. For Ubc13 to carry out its function, UEV (Ubc Variant) is required as a co-factor (Hoffman et al. 1999) Uev1A which is homologous to Mms2 is also a Uev that acts as a cofactor for Ubc13. Ubc13 and Uev1A are required for TRAF6-mediated NF-kB activation which is triggered via bacterial lipopolysaccharide. Uev1A also mediate TRAF-6 induced NEMO polyubiquitination. Uev proteins show sequence identity yet have distinct biological functions. Ubc13, E2 carries out a function when associated with Mms2 and a different function when associated with Uev1A. Figure 3 shows the two different functions carried out by UBC13 when associated with different UEVs (Anderson et al. 2005). Pair wise sequence alignment of Mms2 and Uev1A was performed in Clustal Omega using Global Alignment method (Figure 2).

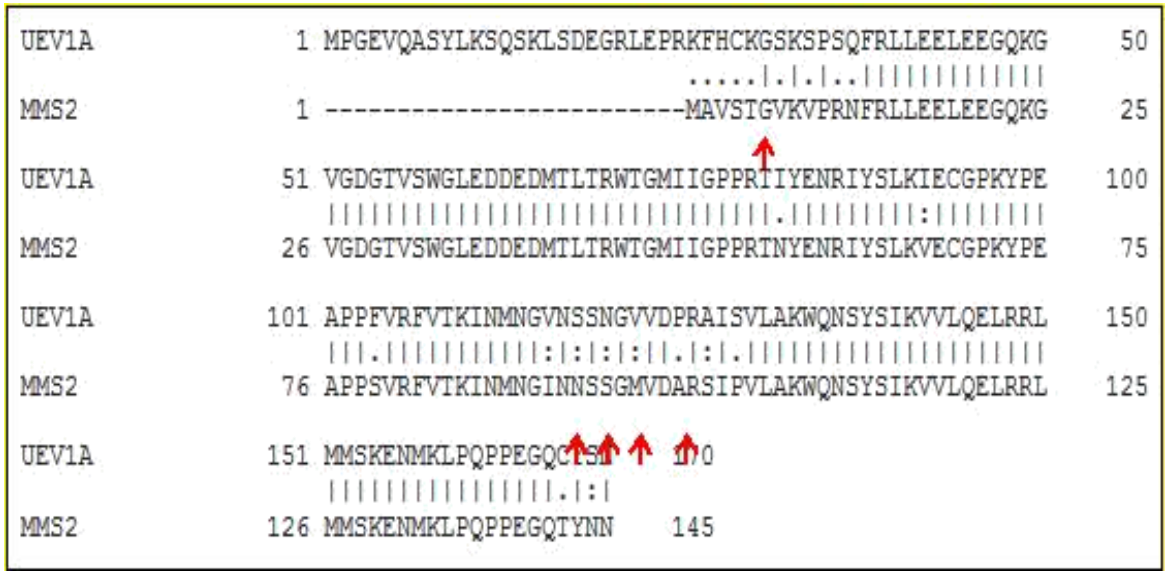


Figure2: **Amino acid sequence alignment of Uev1A and Mms2**. Residues to be mutated are indicated by red arrows. (Sequences from Uniprot and Pairwise alignment by Clustal Omega)

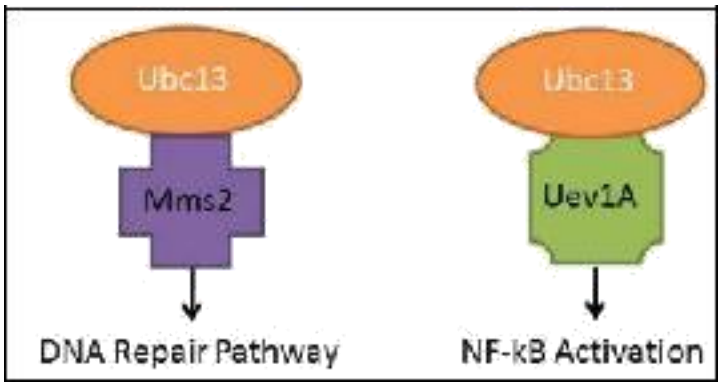


Figure3: **A working model of Ubc13–Uev functions in human cells.** When co-factor is Mms2 for Uev1A, the complex plays a role in DNA damage repair. When co-factor is UEV1A, complex confers innate immunity.

In yeast, Ubc13-Mms2 complex is recruited to DNA damage site by RING finger protein RAD5, via its RING finger domain. RAD5 acts as a bridging factor to bring Ubc13 and Mms2 into contact with the RAD6-RAD18 complex. Ubc13 and Mms2 both of which are cytosolic proteins are redistributed to nucleus on DNA damage. A model proposed in the assembly of repair complexes explains that self-association of RAD18 can promote recruitment of multiple RAD6 but heteromeric association of RAD5 and RAD18 results in the complex formation with two Ubcs: RAD6 and Ubc13-Mms2. This heteromeric complex then function in error-free repair, resulting in the formation of Lys-63 linked poly-ubiquitin chains. Homo-dimers of RAD18 could recruit RAD6 substrates independently of Ubc13-Mms2 which may lead to mono-ubiquitination or Lys-48 linked chains. The figure 4 below explains the heteromeric and homodimeric association of RAD5 and RAD18 (Ulrich et al. 2000).

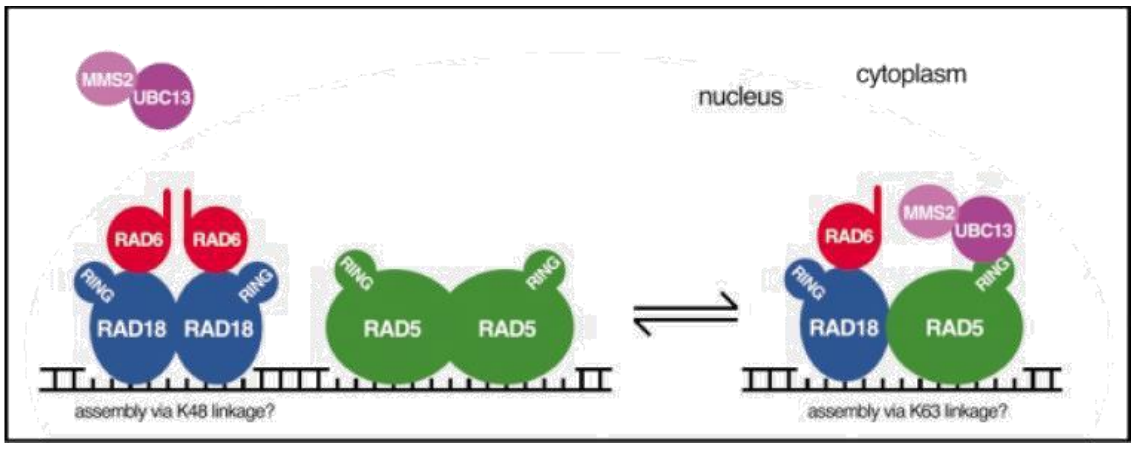


Figure 4: **Model for alternative repair complexes within the RAD6 pathway of DNA repair.**

Homodimerization of RAD18 leads to assemblies comprising RAD6 as the sole UBC, whereas heterodimerization with RAD5 recruit UBC13-MMS2 along with RAD6 for assembly of Lys-63 linkage. (Adapted from Ulrich et al.2000)

3.3 Interaction between Ubc13 and Mms2:

Crystal Structure of complex between hMms2 and hUbc13 was resolved at 1.85 Å resolutions. Their interacting interface provides a physical basis for preference of Mms2 for Ubc13 as partner. hMMS2 and hUbc13 interact through a narrow interface approximately 30 Å long and 10 Å wide. The interior formed by hydrophobic residues of Ubc13 include Phe57, Tyr34 and Leu83 which are not conserved in other E2s. Phe57 and Leu83 contribute number of van der Waal contacts which stabilize the complex. Hydrophobic residues from hMms2, Phe13 and Val7 form the hydrophobic interior of it. Charged residues, Asp38 and Glu39, in the extended L1 loop of hMms2 provide hydrogen bonds that other E2s do not possess. These two differences explain why other E2s have not been reported to recognize hUbc13. (Moraes et al. 2001)

3.4 Lys63 linked polyubiquitination:

Lys63 linked Poly-Ub chains have been found to participate in DNA Repair (Brusky et al. 2000), the NF-kB signaling pathway (Broomfield et al. 2001; Finley 2001), activation of stress-activated protein kinases (SAPKs) (Shi and Kehrl 2003), and ribosome function (Spence et al. 2000).

Lys63 linked Poly-Ub chains synthesis requires the presence of heterodimer formed by E2 enzyme (hUbc13) and its variant (UEV). UEVs are proteins that are structurally and sequentially similar to Ubcs but lack the active site cysteine residue hence catalytically inactive. Uev1a, a UEV apart from Mms2 differs from latter by an N-terminal extension (Anderson et al. 2005; Broomfield et al. 1998; Brusky et al. 2000). Despite having sequence similarity of ~90% in core domain, two UEVs function differently (Anderson et al. 2005). Ubc13- Mms2 heterodimer was found to be involved in error-free Post-replicative Repair Pathway (PRR) in Yeast (Broomfield et al. 1998). Mms2 is required in DNA damage response whereas Uev1a is involved in NF-kB signaling pathway. Hence, UEV members differentially regulate function of E2 (Ubc13) (Anderson et al. 2005).

3.5 Uev1a: Structure and Function

Solution structure of Uev1a shows the presence of core UEV domain and an unstructured 35 residue long N-Terminus which are likely to be solvent accessible. The core UEV fold is composed of N-terminal α helix, α1 (36-49), followed by a four-strand, antiparallel β sheet (β1, 56-59; β2, 69-77; β3, 86-94; and β4, 104-107) and two C-terminal α helices (α2, 140-152; and α3, 154-157), α2 forms extensive contacts with one face of the β sheet (Figure 5). Chemical shift mapping of Uev1a on titration with Ubc13 showed interacting residues around α1 helix and neighboring loop or strands. Eight residues of Uev1a are involved in interaction with hUbc13 (Hau et al. 2006). Uev1a interacts with various E3s downstream signaling pathway to achieve its function. Two such E3 include TRAF6 (Tumor Necrosis Factor (TNF) Receptor Associated Factors) and CHIP (C-Terminal peptide from Human Hsp90).

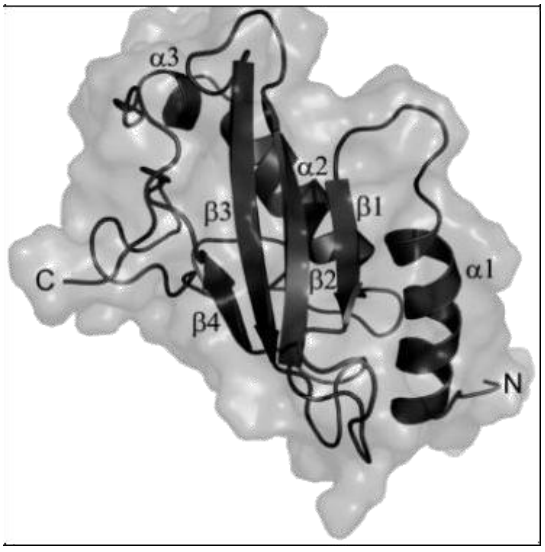


Figure 5: Ribbon representation of 33-170 residues of hUev1a. (Adapted from Hau et al. 2006).

TRAF6 is a RING-type E3 that facilitates Lys63 linked polyubiquitination.TRAF6 polyubiquitinates itself and downstream signaling proteins on specific ligand stimulation. Polyubiquitination of target proteins requires E2 (Ubc13) and its variant (Uev1a) (Pickart and Eddins 2004). The Lys63 linked Poly-Ub chains acts as anchors for recruiting IKK which phosphorylates IKB leading to its degradation and frees NF-KB for transcription (Chen et al.2009).

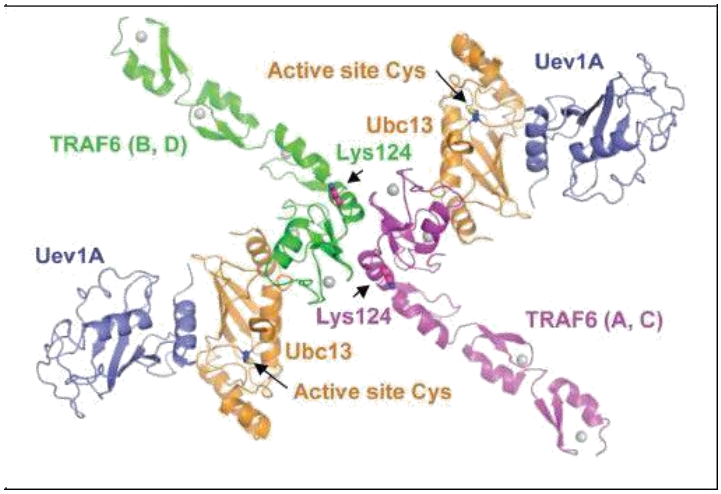


Figure 6: A dimeric TRAF6/Ubc13/Uev1A complex. (Adapted from Yin et al. 2009)

CHIP is a U box E3 Ubiquitin ligase that binds to Hsp proteins and ubiquitylates it leading to degradation of chaperone bound client proteins. CHIP associates with Ubc13-Uev1a to form Lys63 linked Poly-Ub chains for performing its function. Crystal structure of heterotrimer complex of Ubc13-Uev1a-CHIP U box domain is available. The protomers in the homodimeric CHIP protein are asymmetric making only one binding site available for Ubc for the synthesis of Poly-Ub chains (Zhang et al. 2005).

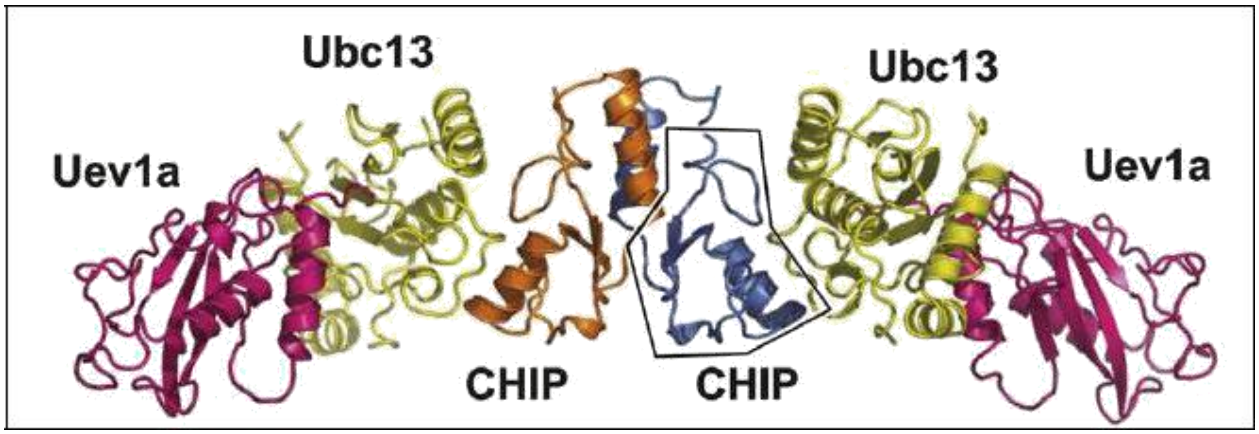


Figure 7: Homodimer of Ubc13-Uev1a-CHIP. (Adapted from Zhang et al. 2009)

NMR-based chemical shift mapping experiments of Ubc13 or Ub showed that the unstructured N-terminus tail does not interact with Ub or Ubc13. Since Uev1a interacts with E3 for performing its function, it is likely that it interacts with them (Hau et al.2006). N-terminal 35 residues are not observed in the crystal structure of Ubc13-Uev1a-CHIP U box heterotrimeric complex. Also, there is a lack of structural data on TRAF6 interaction with Ubc13-Uev1a. Interactions involving N-Terminus of Uev1a with respective E3s is therefore not completely understood structurally.

4. OBJECTIVES:

Part I

1) To generate a mutant Mms2 protein: T5K, express in bacteria and purify it.

Part II

1) To obtain sequences similar to Uev1a for template selection.

2) To Model the unstructured N-terminal 35 residues of Uev1a based on templates, to study the possible secondary structure it can adopt in heterotrimer complex.

3) Dock the modeled peptide to Uev1a, Ubc13-Uev1a heterodimer and Ubc13-Uev1a-CHIP heterotrimer to study its interactions.

5. RATIONALE

Part I

Ubc13, a ubiquitin-conjugating enzyme requires co-factor to perform its function. Ubiquitin-conjugating enzyme Variants (UEVs) are the proteins that act as co-factors. Mms2, a Uev protein conjugates with Ubc13 to synthesize Lys-63 linked poly-ubiquitin chains for signaling in DNA repair (Hoffman et al. 1999; McKenna et al. 2001). It has also been demonstrated that another UEV, Uev1A conjugates with Ubc13 to synthesize Lys-63 linked poly-ubiquitin chains that is involved in NF-kB activation. Hence, UEV family proteins in higher eukaryotes have evolved to differentially regulate E2 function (Anderson et al. 2005). UEVs, Mms2 and Uev1A share more than 90% sequence similarity. The varying amino-acids in two UEVs could regulate the activity of Ubc13 differentially and swapping these residues would aid in understanding the regulation of two distinct biological pathways.

Part II

Ubc13 is an E2 enzyme that requires presence of Uev protein to synthesize Lys63 linked Poly-Ub chains (Hoffman and Pickart 1999). TRAF6 is a RING domain possessing E3 Ligase that requires the presence of Ubc13-Uev1a for synthesis of Poly-Ub chains that leads to activation of IKK followed by NF-KB activation (Deng et al. 2000). CHIP is a U box E3 ligase that acts as a functional partner for the Ubc13-Uev1a heterodimer for chaperone bound client protein degradation. Structure of Ubc13-Uev1a-CHIP has been crystallized with only the U box domain of CHIP (Zhang et al. 2005). There is no structural information regarding interaction of CHIP and TRAF6 with the N-terminus of Uev1a. Hence, studying the interaction would aid in understanding the role of N-terminus in deciding the respective E3 ligases.

6. METHODS:

Part I

6.1 Site Directed Mutagenesis:

To generate mutant MMS2 gene, overlap PCR was performed. This was done by performing two rounds of PCR with insert specific primers and with primers introducing mutations on MMS2 gene. In first round of PCR, the set of primers used for two independent reactions were:

1) Forward primer of Insert (*MMS2*) and antisense primer for mutation and
2) Reverse primer of Insert and sense primer for mutation, respectively.

Forward primer for MMS2:

5` CGT GGA TCC ATG GCG GTC TCC ACA GGA G 3` (BamHI site incorporated) Reverse primer for MMS2:

5` CCG CTC GAG TTA ATT GTT GTA TGT TTG TCC TTC TG 3` (XhoI site incorporated)

For second round of PCR, insert specific primers were used to amplify the products of first PCR. For T5K, only one round of PCR was performed with Mutant forward primer and reverse primer for insert as mutation to be introduced was towards the start of insert.

Mutant forward primer for T5K:

5` CGT GGA TCC ATG GCG GTT TCG AAA GGA GTT AAA G 3`

Mutant Reverse Primer for T5K:

5` CCG CTC GAG TTA ATT GTT GTA TGT TTG TCC TTC TG 3`

PCR Conditions used for first cycle of denaturation was 95^{o}C for 3 mins. For Step 2, denaturation (95^{o}C, 15 secs) was followed by Annealing (50^{o}C, 15 secs) and extension (68^{o}C, 45 secs) for 25 cycles. Final Extension for 5 mins was done at 68^{o}C followed by hold at 4^{o}C.

Ligation reactions were performed to clone mutant PCR product to pRSETA vector after restriction digestion with BamHI and XhoI respectively. Screening of mutant was performed by restriction digestion with BstBI which was the restriction site gained on introducing mutation.

Transformation into DH5-α cells was followed by screening to obtain positive clones with mutant gene.

6.2 Protein Expression and Purification:

BL21 (DE3) pLysS competent cells transformed with recombinant plasmid harboring T5K mutant were allowed to grow till 0.6 O.D measured at 600nm, 37^{o}C and induced with 1mM of IPTG. Induced cells were pelleted after four hours of incubation at 37^{o}C. Cells were suspended in ice-cold lysis buffer (50mM Tris-Cl pH 8, 150mM NaCl, 20mM Imidazole, 1mM BME, 0.5% Triton X-100, 1mM PMSF), sonicated and centrifuged to obtain supernatant containing the protein of interest. The cell lysate was then added to Ni-NTA column equilibriated with ice-cold buffer (50mM Tris-Cl pH 8, 150mM NaCl, 20mM Imidazole, 1mM BME, 0.5% Triton X-100). After nutation for one hour, at 4^{o}C, flow through was collected and column was washed with ice-cold wash buffer (50mM Tris-Cl pH 8, 150mM NaCl, 20mM Imidazole) thrice the volume of column. Elutions were then collected by adding ice-cold elution buffer (50mM Tris-Cl pH 8, 150mM NaCl, 400mM Imidazole). Elution with maximum concentration of protein was dialyzed in ice-cold buffer containing 50mM Tris-Cl pH 8 and150mM NaCl for four hours on ice. Dialyzed protein was stored in equal volume of glycerol at -20 ^{o}C.

Part II

Solution structure of hUev1a (2HLW) and crystal structure of Ubc13-Uev1a-CHIP (2C2V) was obtained from PDB. Structure was visualized and analyzed using PyMol. FASTA format of hUev1a sequence was obtained from NCBI PubMed (NCBI Reference Sequence: NP_068823.2). BLAST was performed for the N-Terminal 35 residues of Uev1a to obtain templates for homology modeling. Protein blast was chosen and FASTA sequence of 35 residues was entered. Protein Data Bank protein (PDB) was chosen as database to obtain sequences that possess structural information in PDB. Sequences of hits chosen as templates were obtained from NCBI PubMed and structure from PDB. Multiple Sequence Alignment for Uev1a and template strands was performed in Clustal-W.

Discovery Studio 4.0 Client was used to build model for 35 residues of hUev1a. FASTA sequences of 35 residues of hUev1a and templates (Single Stranded Binding Proteins) was obtained and accessed in Discovery Studio.

6.3 Modeling 35 residues:

Open the FASTA sequence of hUev1a with 35 residues in Discovery Studio. In *Macromolecules* Tab, click on *Align Sequences and Structures* followed by *Align Sequence profile*. (Prior to the step, align template sequences by opening the sequences in a single FASTA file and clicking on *Align Sequences).* In the window that appears on clicking *Align Sequence Profile*, choose the pre-aligned template sequences as *Input Sequence Alignment* and choose 35 residues of hUev1a N-Terminus as *Input Sequence Set* and click on *Run.*

Open the crystal structure of the templates in a single file (Copy the structure and paste). Click on *Create Homology Models* in *Macromolecules* tab followed by *Align Sequence to Templates*. Choose the newly aligned file of template and model sequence as *Input Model Sequence* and template structures (in the same molecule window) as *Input Template Structure* and click on *Run.*

Click on *Build Homology Models* and choose final alignment sequence as *Input Sequence Alignment*, final structure output *as Input Template Structure* and 35 residue long hUev1a as Input Model Sequence. Click on *Run.*

20 Models will be generated on Job completion. To assess the validity of 3D structure, Click on *Verify Protein (Profiles-3D)* under *Create Homology Models*. Choose the model generated file as *Input Protein Molecules* and click on *Run.*

6.4 Cleaning the crystal structure of protein prior docking:

Open the crystal structure of Ubc13-Uev1a-CHIP in Discovery Studio and change the display style. Change the protein representation to *flat ribbon* and color by *Secondary type*.

Click on *Structure* in the toolbar and *Crystal Cell* in the dropdown and click on *Remove Cell*.

Click on *Macromolecules* Tab, *Protein Report* and if any ligands or water molecules found, delete them.

In the same tab, click on *Prepare Protein* followed by *Clean Protein* and *Show Clean Report*. Alternate confirmations of proteins will be reported.

This is followed by *Prepare Protein* Step. In the window that appears, change the *pH* and *Ionic strength* according to conditions mentioned in published papers. Click on *Run.*

On job completion, a file opens as *protein name_prep* and reports about missing residues etc. Validate the structure on clicking *X-ray* (or *Macromolecules*) Tab and click on Validate Protein Structure followed by *Perform Validations*, *Check Structure*.

This gives the summary of residues residing in the disallowed regions and main chain and side chain violations.

In main chain confirmation violations, if any residues lie in the Disallowed region, Click on Protocols, followed by Discover Studio > Macromolecules > Refinement > Loop Refinement > Modeller > Select residues in the disallowed region present in the loops > Run.

Modeller creates models and the one with least DOPE Score will be displayed on top of the results. This model can be validated again to check if the violations have reduced, if not perform the above process again to reduce the number of residues lying in the disallowed region.

For residues with side chain violations, Protocols, followed by Discovery studio > Macromolecules > Refinement > loop refinement > side chain refinement > Select all the residues listed in the side chain incorrect confirmation report > Run.

Model with least DOPE Score and minimized violations can then be saved as clean protein for further docking study.

6.5 Docking the modeled peptide on to refined protein complex:

Model with high Verify score has to be docked with 1) Uev1a alone, 2) Ubc13-Uev1a and 3) Ubc13-Uev1a-CHIP heterotrimer. To dock with Uev1a alone, open the heterotrimer file (2C2V) select chain H and chain T and press delete. Refine the protein as per the steps mentioned above and choose the model with least energy. For heterodimer complex, choose chain T and delete and refine the protein and choose the model with lowest energy. For heterotrimer complex, refine the complex 2C2V which has only three chains (I, H and T).

Open the refined Model structure and crystal structure of heterotrimer complex to be docked in Discovery Studio. Click on *Dock and Analyze Protein Complexes* followed by *Dock Proteins (ZDOCK)*. Select least energy model of receptor protein as *Input Receptor Protein* and the best model obtained as *Input Ligand Protein*. Choose *Angular Step Size* as 6. Expand *Filter Poses* and choose Val36 (Click on the residue and *define group* in prior) as *Receptor Binding Site Residues* and click on *Run*.

On Job completion, click on *View Results* and analyze Docked Poses by browsing through *Poses in a Cluster.*

7. RESULTS AND DISCUSSIONS:

Part I

7.1 Screening of pRA19, MMS2T5K:

DH5-α cells transformed with Ligated vector and Insert (T5K MMS2) were screened for presence of positive clones which possess recombinant plasmid. Miniprep for Plasmid Isolation was performed for 10 colonies and DNA isolated was loaded on 0.6% agarose gel. Mobility shift was detected in DNA of three colonies. Figure 8 shows the observed mobility shift for colony-4, colony-7 and colony-8 (C4, C7 and C8).

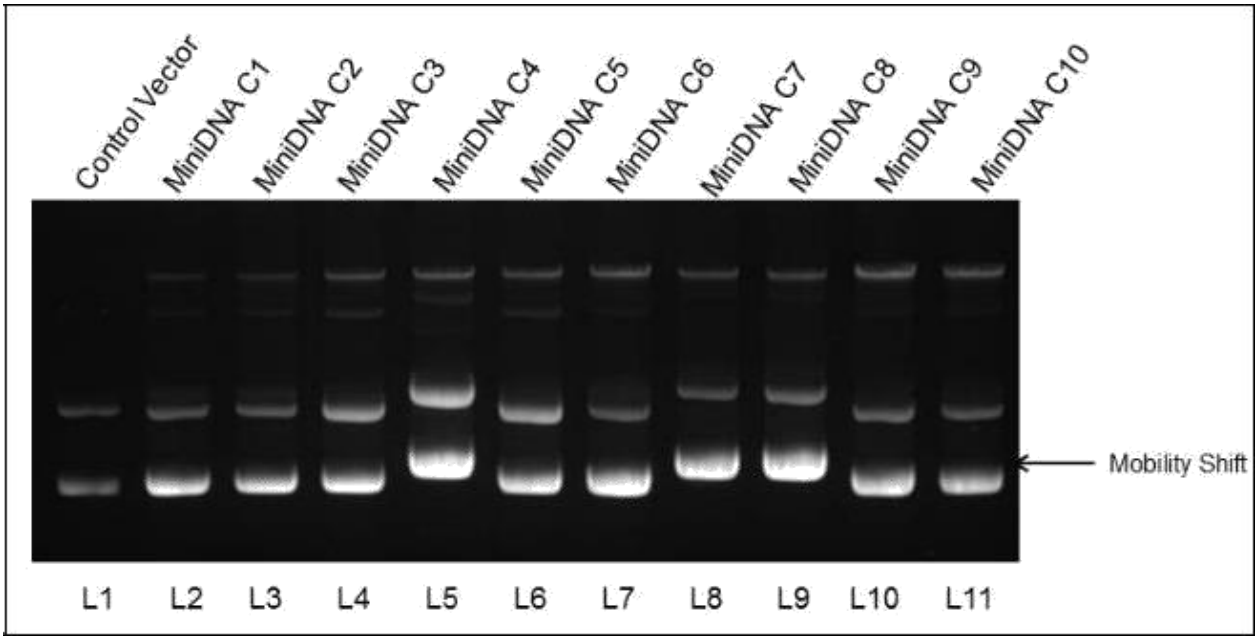


Figure 8: **Screening of Colonies for positive clone.** Isolated DNA was run on 0.6% agarose at 50V. Control Vector pRSETA was loaded on L1 and plasmid isolated from colony 1 to colony 10 by MiniPrep was loaded on L2 to L11, denoted by C1 to C10 in the figure. Band corresponding to Vector was observed in colonies except C4, C7 and C8 where mobility shift was observed.

7.2 Insert Release of pRA19

Confirmation of Insert in colonies showing mobility shift was done by digesting the DNA with restriction enzymes used to clone the insert into pRSETA vector. Restriction enzymes BamHI and XhoI were used to check the release of insert from C4, C7 and C8 DNA. After Incubation of

DNA with compatible buffer and respective enzymes for four hours, DNA was run on 0.6% agarose gel. Insert release was observed (Figure 9) and band corresponding to approximately 430bp was observed on gel.

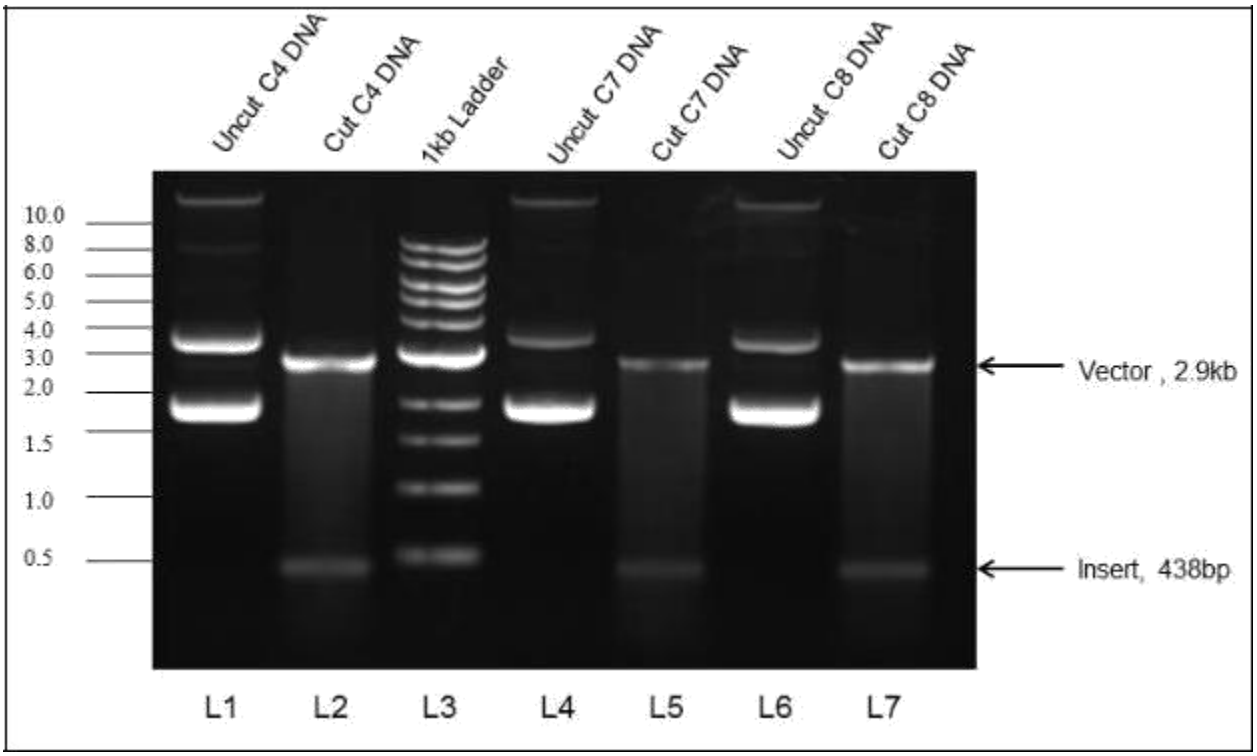


Figure 9: **Insert Release for C4, C7 and C8 DNA with BamHI and XhoI**. L1, L4 and L5 were loaded with Uncut DNA of C4, C7 and C8 respectively. Digested DNA of C4, C7 and C8 was loaded on L2, L5 and L7 respectively. 1kb DNA ladder was loaded on L3. Band size of ladder is indicated to the left of figure in kbs. Linearized vector of approximately 3kb was observed apart from Insert of 438bp.

7.3 Mutant screening of pRA19

Insert release was confirmed by digesting DNA with BamHI and XhoI. To confirm the generation of mutant site, mutation screening was performed. BstBI is a restriction site gained by pRA19 on introducing the mutation. Digestion was performed with BstBI, restriction site already present in vector. Hence, wild type MMS2 would show only single band in comparison to mutant MMS2 having two BstBI sites.

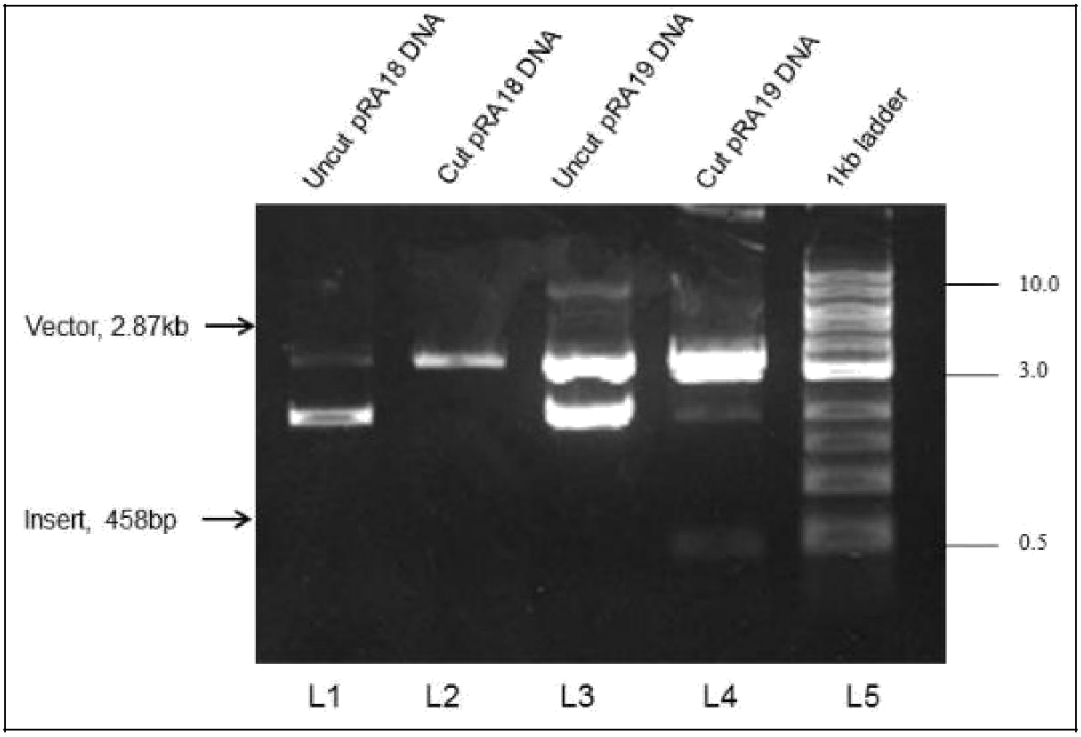


Figure 10: **Mutant Screening of pRA19 by BstBI digestion.** L1 and L3 were loaded with uncut pRA18 and pRA19 DNA respectively. L2 and L4 were loaded with BstBI digested DNA of pRA18 and pRA19 respectively. 1kb DNA ladder was loaded on L5. Band size of ladder is indicated to the left of figure in kbs. pRA18 carrying wild-type MMS2 showed only single linearized band whereas a band of 450bp was observed in pRA19 cut DNA showing presence of two BstBI sites in pRA19 mutant.

7.4 Induction with IPTG

Transformation of mutant plasmid to BL21 (DE3) pLysS was performed after confirmation of mutant site. Cells were allowed to grow overnight at 37^{o}C and this served as pre-inoculum for performing induction of 125ml culture. Uninduced and Induced cells corresponding to 0.15 O.D was loaded on 12.5% SDS PAGE to check the presence of induced band. Figure 11 shows the presence of Induced band between 25 and 20kDa.

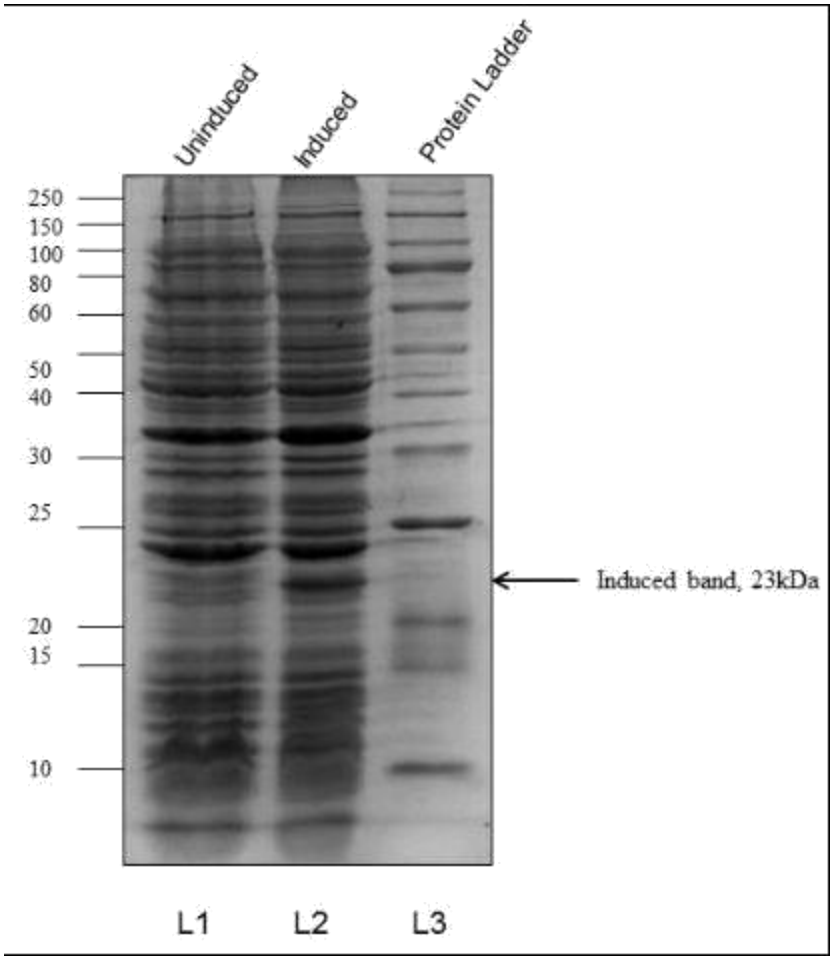


Figure 11: **Induction check for pRA19.** Cells corresponding to 0.15 O.D was loaded on 12.5% SDS PAGE to check for induced band of desired molecular weight. L1 was loaded with uninduced sample, L2 was loaded with induced sample and L3 was loaded with protein marker molecular weight indications of which are given towards right (in kDa). Induced band of approximately 23 kDa was observed in the gel image.

7.5 Protein Purification

Induced cells were pelleted and suspended in lysis buffer, sonicated and centrifuged to obtain cell pellet and lysate. Lysate is then added to Ni-NTA column equilibriated with buffer. After nutation for an hour in 4^{o}c, flow through was collected and column was washed with wash buffer thrice the volume of column. Five elutions were then collected after adding elution buffer containing 400mM imidazole. Figure 12 shows protein eluted run on 12.5% SDS PAGE along with cell pellet, Lysate, flow through, wash and protein ladder. Desired protein of 23kDa was collected in third elution in higher amounts.

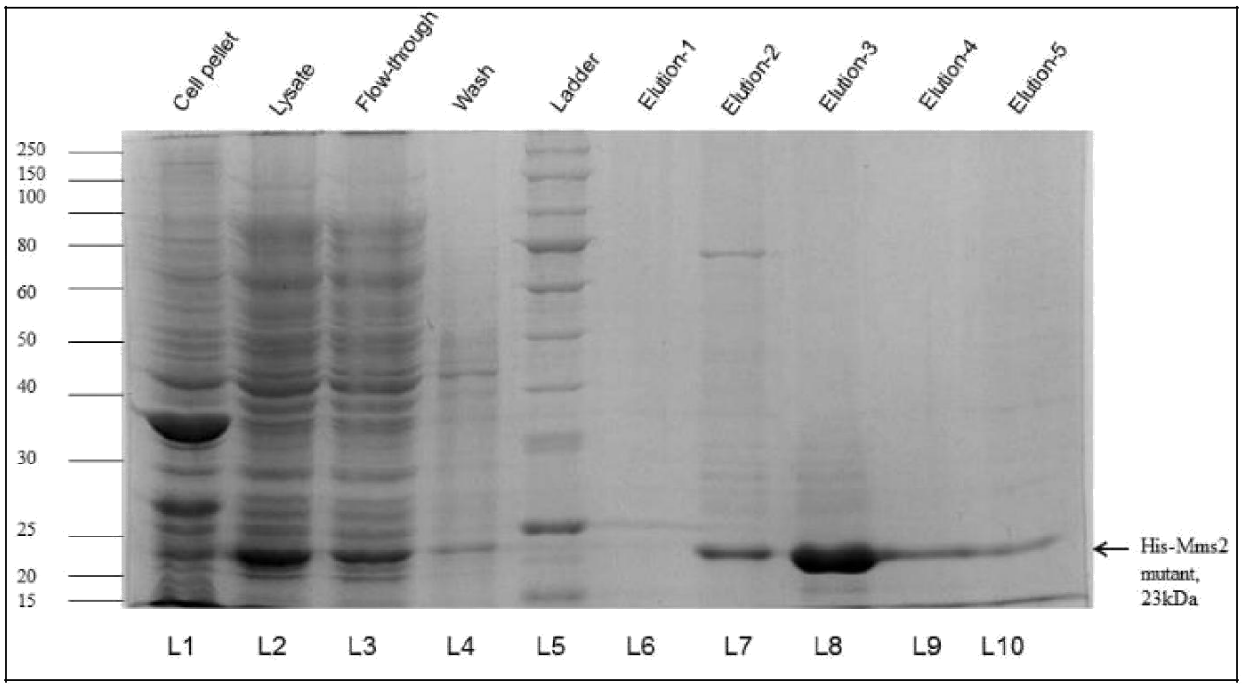


Figure 12: **Purified protein run on 10% SDS PAGE.** His-tagged MMS2 was purified by Ni-NTA affinity purification method. Cell pellet resuspended in 1X-Tris Glycine was loaded on L1, Lysate containing protein was loaded on L2, Flow through collected after nutation was collected in L3 followed by wash on L4. Protein ladder was loaded on L5 (molecular weight indications of which are shown to the right of the figure) followed by five elutions of protein. Elution 3 showed dense band of protein indicating highest amount eluted.

7.6 Screening of Wild-type MMS2, pRA18

Screening for wild-type MMS2 is performed after transforming the ligated product of Digested pRSETA vector and digested PCR amplified MMS2 insert into DH5-α. Only one round of PCR is performed for Wild-type in comparison to mutant. Plasmid isolated was run on agarose gel to screen colonies transformed with recombinant plasmid. Figure 13 shows the mobility shift observed in DNA in comparison to control pRSETA vector.

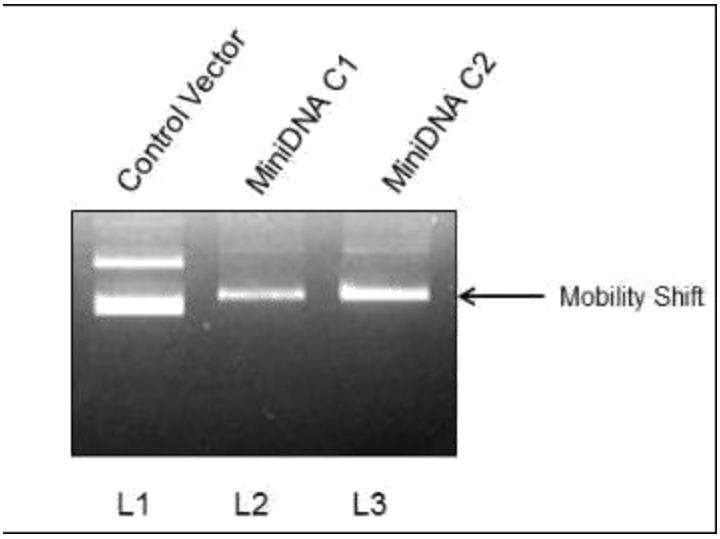


Figure 13: **Screening of Wild-type MMS2**. Plasmid isolated from colonies was loaded on agarose gel. Control Vector loaded on L1 is pRSETA followed by DNA isolated from Colony and Colony 2 in L2 and L3 respectively. Mobility shift was observed in colonies indicating presence of Insert.

7.7 Insert Release for Wild-type MMS2.

To confirm the presence of insert, DNA was digested with restriction enzymes used to clone the insert. BamHI and XhoI were used to check the release of insert of 438bp. Insert release was observed in DNA isolated confirming presence of wild-type MMS2 in recombinant plasmid (Figure 14).

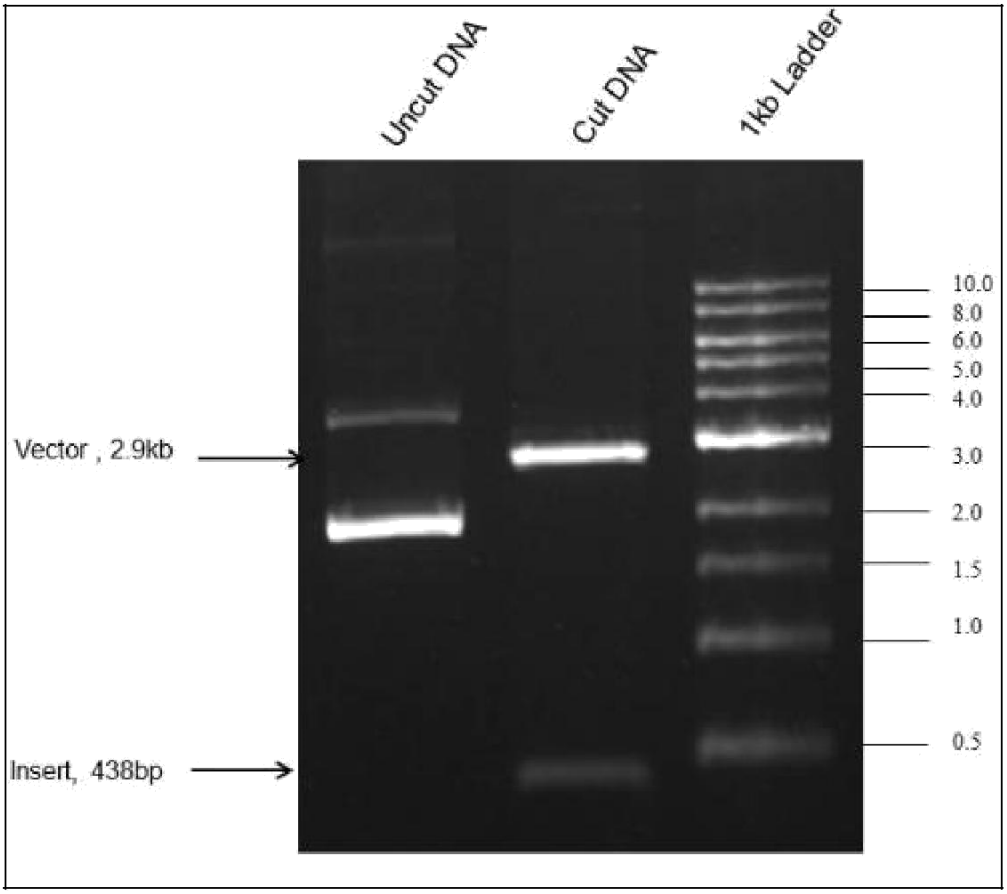


Figure 14: **Insert release of pRA18 with BamHI and XhoI.** L1 was loaded with Uncut DNA which showed mobility shift. Digested DNA was loaded on L2. 1kb DNA ladder was loaded on L3. Band size of ladder is indicated to the left of figure in kbs. Linearized vector of approximately 3kb was observed apart from Insert of 438bp.

Part II

7.8 PyMol Analysis:

Solution structure obtained from PDB was analyzed using PyMol and the length of 35 residues was measured.

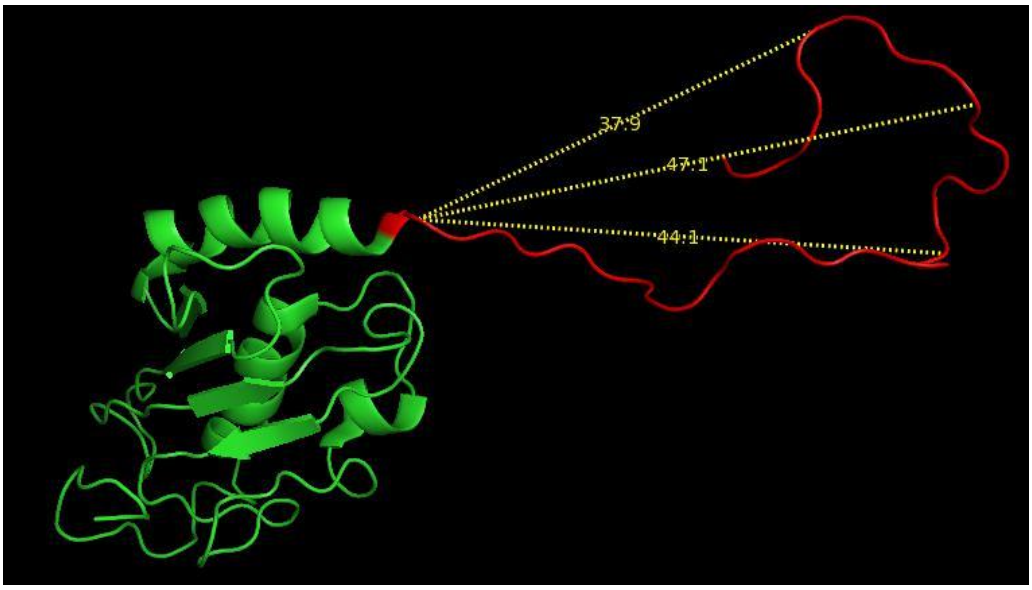


Figure 15: Solution structure of hUev1a as visualized through PyMol.

The 35 residues marked in red (Figure 15) forms an unstructured loop and the maximum distance measured from the secondary structure was 47.1 Angstrom.

Crystal Structure of Ubc13-Uev1a-CHIP was analyzed considering only one unit of the heterotrimer.

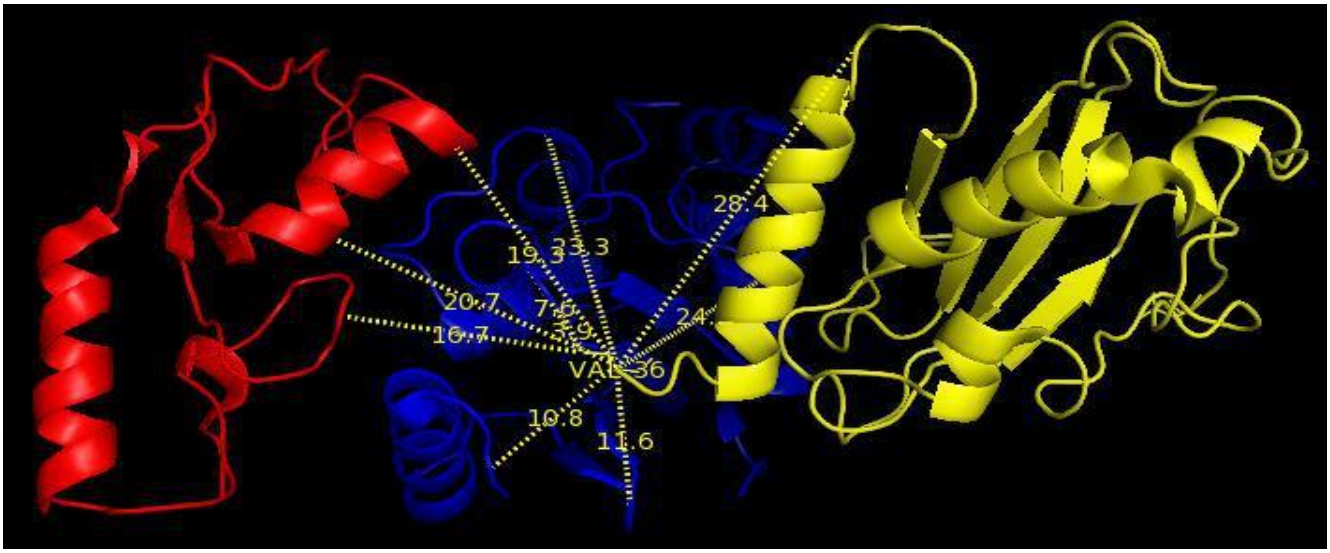


Figure 16: One Subunit of Ubc13-Uev1a-CHIP heterotrimer complex. Structure in red is the U Box domain of CHIP, Blue is Ubc13 and Yellow is Uev1a.

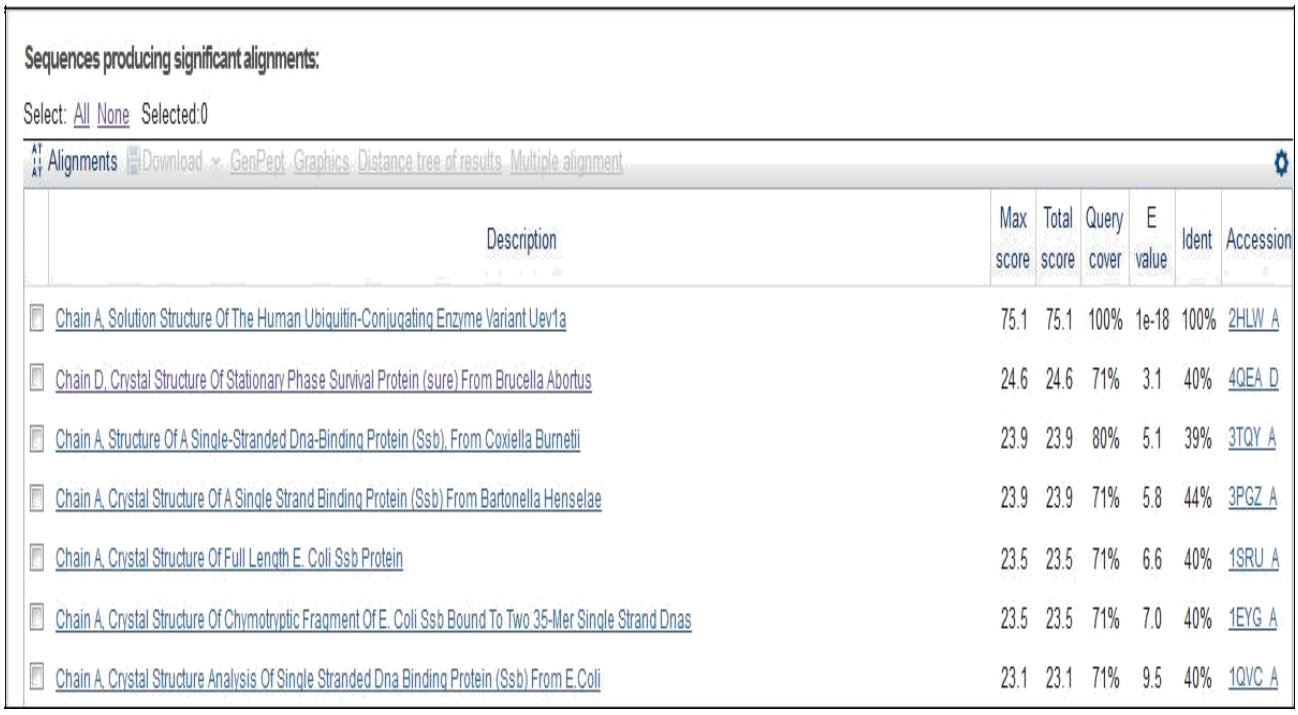
From the heterotrimer complex, it can be inferred that the maximum distance a structure (extending from Val-36) can adopt to occupy the cleft formed by the three proteins is 28 Angstrom.

Hence, there can be possibility of N-Terminal 35 residues forming a reordered secondary structure when the protein interacts with Ubc13 and CHIP. The reordered secondary structure could either interact with Ubc13 or E3 (CHIP) or fold back to its own structure.

A Model of 35 residues generated could hence be docked to the heterotrimer complex to understand the possible interactions with proteins involved.

7.9 BLAST and MSA Results:

BLAST results of 35 residue long N-Terminus showed hits as shown in the Figure 17.



Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession
Chain A, Solution Structure Of The Human Ubiquitin-Conjugating Enzyme Variant Uev1a	75.1	75.1	100%	1e-18	100%	2HLW_A
Chain D, Crystal Structure Of Stationary Phase Survival Protein (sure) From Brucella Abortus	24.6	24.6	71%	3.1	40%	4QEA_D
Chain A, Structure Of A Single-Stranded Dna-Binding Protein (Ssb), From Coxiella Burnetii	23.9	23.9	80%	5.1	39%	3TQY_A
Chain A, Crystal Structure Of A Single Strand Binding Protein (Ssb) From Bartonella Henselae	23.9	23.9	71%	5.8	44%	3PGZ_A
Chain A, Crystal Structure Of Full Length E. Coli Ssb Protein	23.5	23.5	71%	6.6	40%	1SRU_A
Chain A, Crystal Structure Of Chymotryptic Fragment Of E. Coli Ssb Bound To Two 35-Mer Single Strand Dnas	23.5	23.5	71%	7.0	40%	1EYG_A
Chain A, Crystal Structure Analysis Of Single Stranded Dna Binding Protein (Ssb) From E.Coli	23.1	23.1	71%	9.5	40%	1QVC_A

Figure 17: BLAST results of sequences showing similarity to 35 residues of Uev1a N-Terminus

Single Stranded Binding Proteins (3TQY, 3PGZ and 1SRU) showed similarity of ~60% similarity. PDB structure of the three proteins showed Helix-loop-extended beta sheet as three proteins possess structural similarity. 4QEA is a stationary phase survival protein (sure) which possesses more that 60% similarity with the sequence of interest. Structure of 4QEA included two beta strands separated by a loop. Sequences of SSBs and sure were obtained from PubMed and MSA was performed in Clustal W2. When 4QEA was excluded there was an increase in the number of similar and identical residues. Hence, MSA was performed again with Uev1a 35 residues and 3 SSB sequences that shared more than 60% similarity in BLAST.

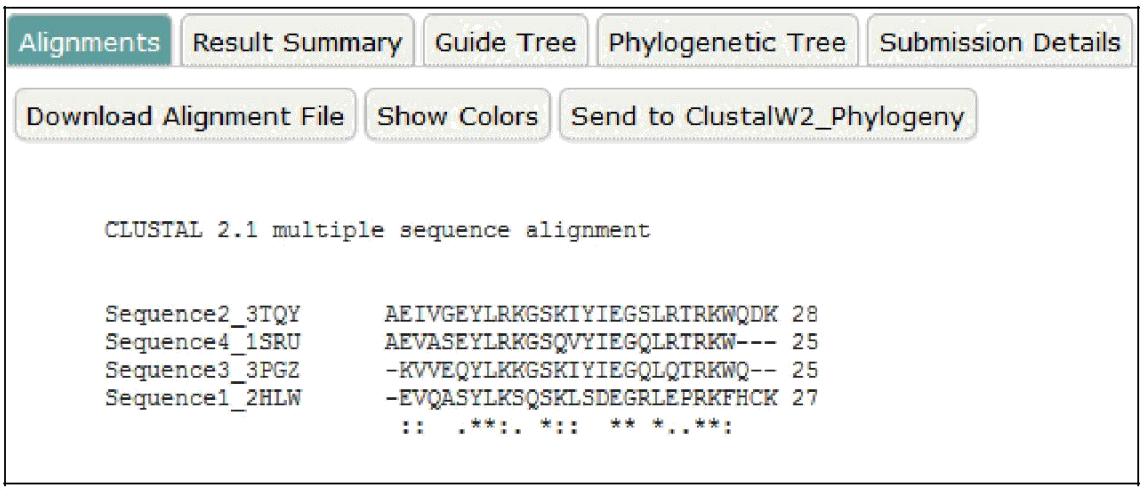


Figure 18: MSA results of sequences showing similarity to 35 residues of Uev1a N-Terminus.

Results showed 18 identical residues and 9 similar residues to 2HLW (Uev1a sequence). Hence the 3 SSB structure was chosen as template for proceeding with homology modeling.

7.10 Homology Modeling Results:

Sequence alignment of SSBs and Uev1a was performed in Discovery Studio prior to homology modeling. Figure 19 shows the alignment results.

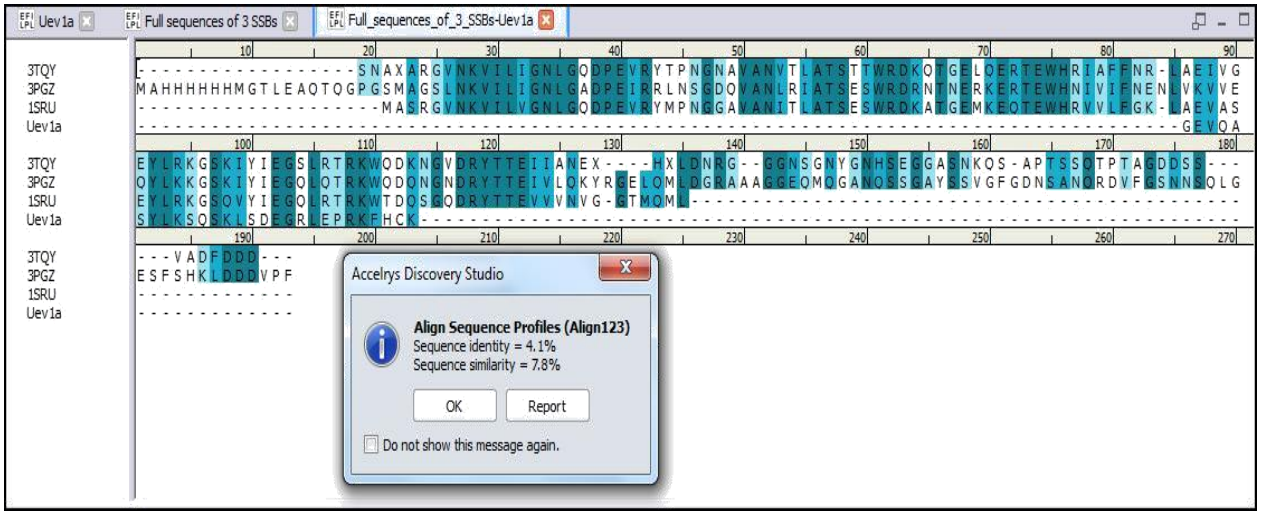


Figure 19: Sequence alignment of SSBs and Uev1a performed in Discovery Studio.

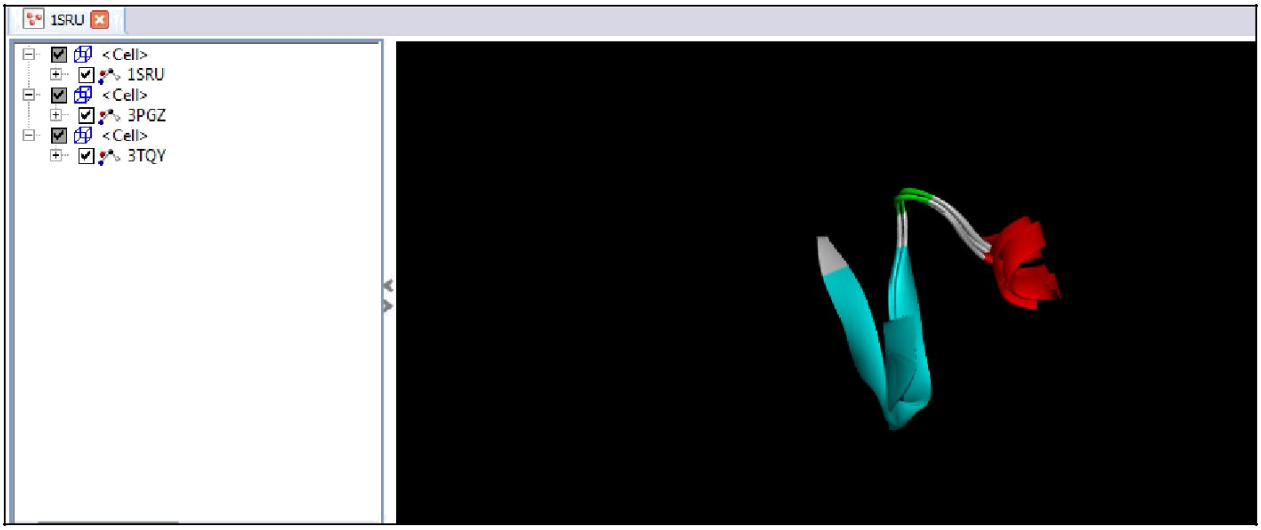


Figure 20: Structure of Three SSBs in a single molecule window

20 Models were then obtained after steps elaborated in Methods. Validation was performed and Verify score was obtained for 20 models obtained (Figure 21).

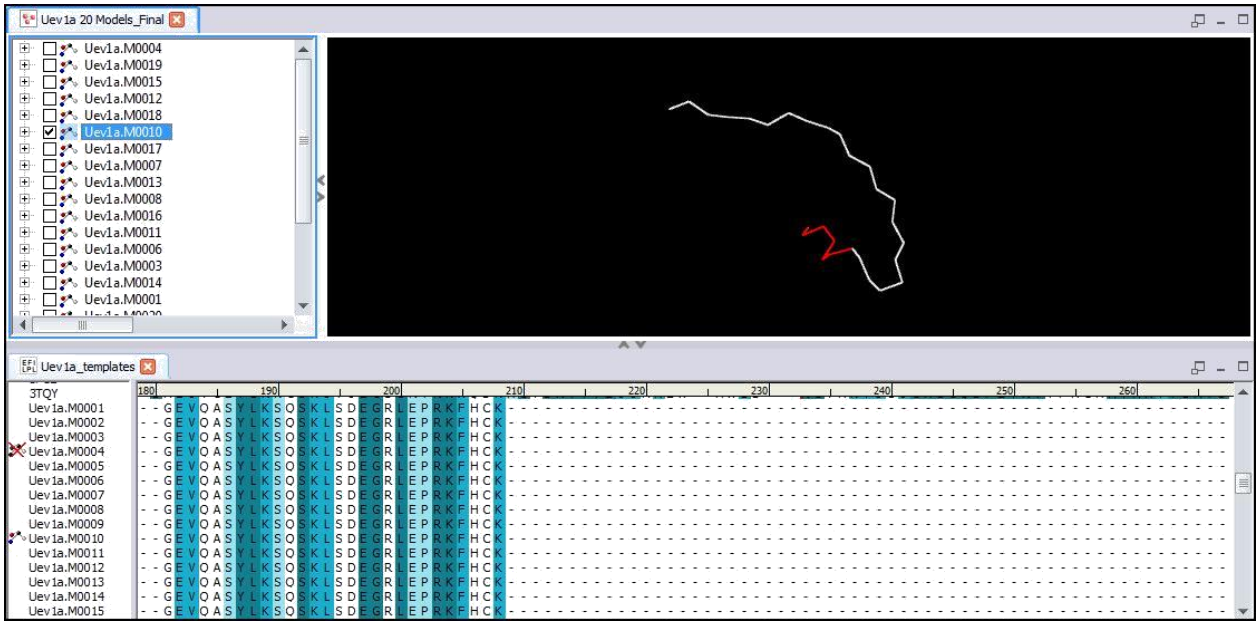


Figure 21: Structure of 20 models generated. Model 10 is shown in the window along with alignment.

Model 10 was chosen as it had the Highest Verify Score of 3.62 and Total PDF Energy of 442.68 and DOPE Score of -1067.73 (Figure 22). Model 10 possessed residues till Lys28.

	Name	PDF Total Energy	PDF Physical Energy	DOPE Score	Verify Score	Verify Expected High Score	Verify Expected Low Score
1	Uev1a.M0004	435.663	66.4133	-997.278	2.91	12.5017	5.62575
2	Uev1a.M0019	436.912	64.0227	-1,072.45	2.73	12.5017	5.62575
3	Uev1a.M0015	440.107	66.0195	-1,037.99	2.82	12.5017	5.62575
4	Uev1a.M0012	442.03	66.1203	-1,019.01	3.3	12.5017	5.62575
5	Uev1a.M0018	442.065	64.7126	-1,020.42	3.33	12.5017	5.62575
6	Uev1a.M0010	442.68	64.5884	-1,067.73	3.62	12.5017	5.62575
7	Uev1a.M0017	443.782	66.8608	-1,058.43	2.1	12.5017	5.62575
8	Uev1a.M0007	444.389	66.9207	-983.01	2.38	12.5017	5.62575
9	Uev1a.M0013	444.879	70.7875	-1,053.99	3.09	12.5017	5.62575
10	Uev1a.M0008	446.547	68.9059	-1,036.07	2.55	12.5017	5.62575
11	Uev1a.M0016	448.224	64.4409	-1,066.68	3.09	12.5017	5.62575
12	Uev1a.M0011	452.852	67.7064	-1,033.16	2.52	12.5017	5.62575
13	Uev1a.M0006	454.456	63.9538	-1,067.26	1.7	12.5017	5.62575
14	Uev1a.M0003	456.175	61.6744	-976.698	2.52	12.5017	5.62575
15	Uev1a.M0014	462.477	74.5805	-1,075.69	1.89	12.5017	5.62575
16	Uev1a.M0001	466.942	68.2642	-1,006.37	3.09	12.5017	5.62575
17	Uev1a.M0020	479.414	75.9438	-1,065.13	3.14	12.5017	5.62575
18	Uev1a.M0002	490.542	67.6916	-976.309	1.43	12.5017	5.62575
19	Uev1a.M0005	494.569	74.9952	-939.926	1.14	12.5017	5.62575
20	Uev1a.M0009	527.26	82.4733	-970.686	2.36	12.5017	5.62575

Figure 22: Verify Score of 20 models generated. Model 10 is shown in the window along with alignment.

7.11 Docking Results:

On Docking, 2000 poses were obtained for each complex. 6000 poses were analyzed and the distance between Val36 of Uev1a and Lys28 of modeled peptide was measured. The least possible distance between two residues is 10.8 Angstrom as the seven intermediate residues could form a helix of two turns (One turn length: 5.4 Angstrom). The Maximum distance of 28 Angstrom is possible if they form an extended structure like in solution structure available.

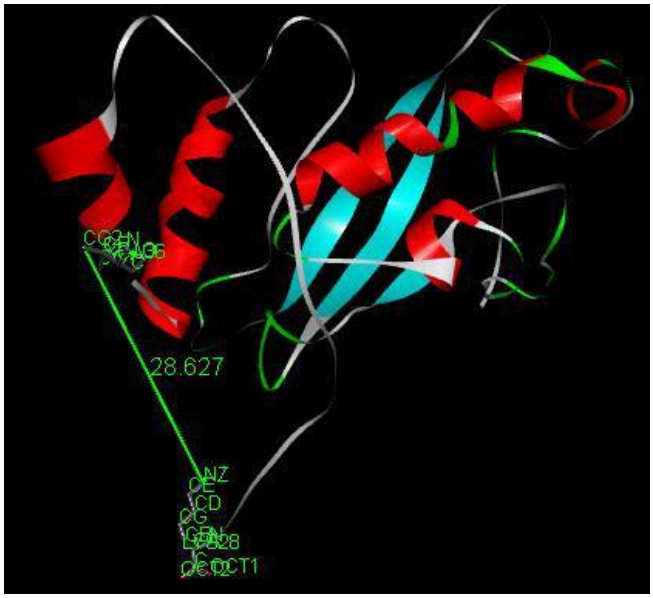


Figure 23: Model 10 docked with Uev1a alone. Pose shows the distance between Val36 of receptor and Lys28 of Model.

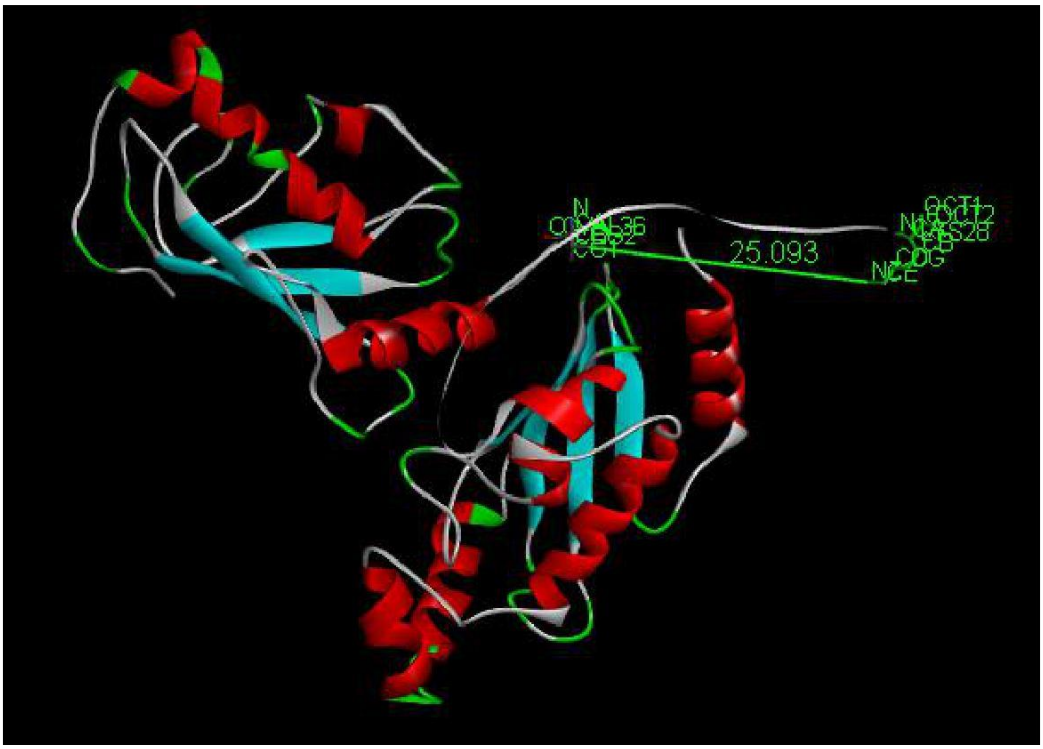


Figure 24: Model 10 docked with Ubc13-Uev1a alone. Pose shows the distance between Val36 of receptor and Lys28 of Model.

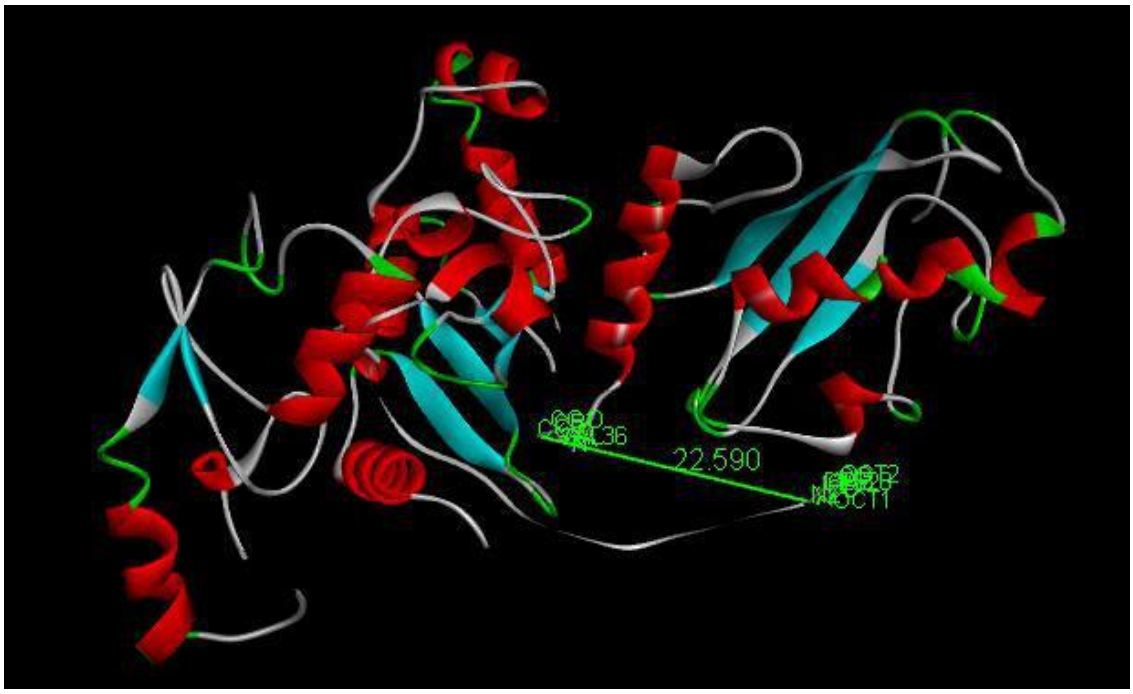


Figure 25: Model 10 docked with Ubc13-Uev1a alone. Pose shows the distance between Val36 of receptor and Lys28 of Model.

Model 10 Docked to Uev1a alone showed 787 permissible poses out of 2000.

Model 10 Docked to Ubc13-Uev1a showed 609 permissible poses out of 2000.

Model 10 Docked to Ubc13-Uev1a-CHIP Heterotrimer showed 572 permissible poses out of 2000 out of which 147 poses were not found to interact with CHIP.

Hence, 21.2% of poses analyzed were found to interact with CHIP compared to 39% that was observed when docked to Uev1a alone which reduced to 30.5% in Ubc13-Uev1a Heterotrimer complex. Hence, there are chances of N-Terminus of Uev1a interacting with E3 Ligase which may determine the final biological function.

REFERENCES:

Andersen, Parker L., et al. "Distinct regulation of Ubc13 functions by the two ubiquitin-conjugating enzyme variants Mms2 and Uev1A." The Journal of cell biology 170.5 (2005): 745-755.

Broomfield, Stacey, Todd Hryciw, and Wei Xiao. "DNA postreplication repair and mutagenesis in Saccharomyces cerevisiae." Mutation Research/DNA Repair 486.3 (2001): 167-184.

Brusky, Janna, Yu Zhu, and Wei Xiao. "UBC13, a DNA-damage-inducible gene, is a member of the error-free postreplication repair pathway in Saccharomyces cerevisiae." Current genetics 37.3 (2000): 168-174.

Clague, Michael J., and Sylvie Urbé. "Ubiquitin: same molecule, different degradation pathways." Cell 143.5 (2010): 682-685.

Deng, Li, et al. "Activation of the IκB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain." Cell 103.2 (2000): 351-361.

Finley, Daniel, and Vincent Chau. "Ubiquitination." Annual review of cell biology 7.1 (1991): 25-69.

Hau, D. Duong, et al. "Structure and Interactions of the Ubiquitin-Conjugating Enzyme Variant Human Uev1a: Implications for Enzymatic Synthesis of Polyubiquitin Chains,||." Biochemistry 45.32 (2006): 9866-9877.

Hochstrasser, Mark. "Ubiquitin-dependent protein degradation." Annual review of genetics 30.1 (1996): 405-439.

Hofmann, Roseanne M., and Cecile M. Pickart. "Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair." Cell 96.5 (1999): 645-653.

Hofmann, Roseanne M., and Cecile M. Pickart. "In vitro assembly and recognition of Lys-63 polyubiquitin chains." Journal of Biological Chemistry 276.30 (2001): 27936-27943.

McKenna, Sean, et al. "Noncovalent interaction between ubiquitin and the human DNA repair protein Mms2 is required for Ubc13-mediated polyubiquitination." Journal of Biological Chemistry 276.43 (2001): 40120-40126.

Moraes, Trevor F., et al. "Crystal structure of the human ubiquitin conjugating enzyme complex, hMms2–hUbc13." Nature Structural & Molecular Biology 8.8 (2001): 669-673.

Passmore, Loria, and David Barford. "Getting into position: the catalytic mechanisms of protein ubiquitylation." Biochem. J 379 (2004): 513-525.

Pastushok, Landon, and Wei Xiao. "DNA postreplication repair modulated by ubiquitination and sumoylation." Advances in protein chemistry 69 (2004): 279-306.

Pickart, Cecile M. "Ubiquitin enters the new millennium." Molecular cell 8.3 (2001): 499-504.

Pickart, Cecile M., and Michael J. Eddins. "Ubiquitin: structures, functions, mechanisms." Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 1695.1 (2004): 55-72.

Shi, Chong-Shan, and John H. Kehrl. "Tumor necrosis factor (TNF)-induced germinal center kinase-related (GCKR) and stress-activated protein kinase (SAPK) activation depends upon the E2/E3 complex Ubc13-Uev1A/TNF receptor-associated factor 2 (TRAF2)." Journal of Biological Chemistry 278.17 (2003): 15429-15434.

Spence, Jean, et al. "Cell cycle–regulated modification of the ribosome by a variant multiubiquitin chain." Cell 102.1 (2000): 67-76.

Thrower, Julia S., et al. "Recognition of the polyubiquitin proteolytic signal." The EMBO journal 19.1 (2000): 94-102.

Ulrich, Helle D., and Stefan Jentsch. "Two RING finger proteins mediate cooperation between ubiquitin- conjugating enzymes in DNA repair." The EMBO Journal 19.13 (2000): 3388-3397.

Wu, Zhaojia, et al. "Ubiquitin-conjugating enzyme complex Uev1A-Ubc13 promotes breast cancer metastasis through nuclear factor-κB mediated matrix metalloproteinase-1 gene regulation." Breast Cancer Research 16.4 (2014): R75.

Yin, Qian, et al. "E2 interaction and dimerization in the crystal structure of TRAF6." Nature structural & molecular biology 16.6 (2009): 658-666.

Zhang, Minghao, et al. "Chaperoned ubiquitylation—crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex." Molecular cell 20.4 (2005): 525-538.