TAGGING PROTEINS: SUMOYLATION MECHANISM BY MOLECULAR MODELING AND MOLECULAR DYNAMICS SIMULATIONS

by

Elif Nihal Korkmaz

B.S., Chemical Engineering, Boğaziçi University, 2008

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemical Engineering

Boğaziçi University

2010

TAGGING PROTEINS: SUMOYLATION MECHANISM BY MOLECULAR MODELING AND MOLECULAR DYNAMICS SIMULATIONS

APPROVED BY:

Prof. Türkan Haliloğlu ..... (Thesis Supervisor)

Prof. Pemra Doruker Turgut .....

Assist. Prof. Nevra Özer .....

DATE OF APPROVAL: 30.07.2010

ACKNOWLEDGEMENTS

This research was funded by The Scientific and Technological Research Council of Turkey (TUBITAK), Project no: 107T382.

First of all, I would like to express my gratitude to my thesis supervisor, Prof. Türkan Haliloğlu, for her guidance, friendly encouragement and patience. I am also thankful to Prof. Pemra Doruker and Assist. Prof. Nevra Özer for their time devoted to read and comment on my thesis. I am grateful to Prof. Can Özturan, too.

I feel lucky to have worked with Dr. Ruth Nussinov. In addition to sharing her valuable knowledge in scientific discussions, I will always appreciate her kindness, hospitality and energy.

I am also thankful all my friends in PRC including Canan Dedeoğlu. Special thanks to Burcu Aykaç Fas, Ezgi Karaca and Melda Tozluoğlu for the inspiration and for sharing their knowledge.

I would also like to thank specially Arzu Uyar, E. Zerrin Bağcı, Nazli Özbek and Elif Dönmez for always being there for me whenever I need. I appreciate their patience against all my whinings. Four great person that I can trust no matter what... It feels like I have four more sisters...

Onur Selki... I feel lucky to have you in my life... With your endless and limitless love, support, tenderness and patience, I am the happiest person on earth... Every minute, every second of the times spent together are engraved in my heart and there is lots and lots of space there to record as we grow old together. I want to have you by my side forever and make the fairy tales come true...

Lastly, my beloved, warm family... I could not make it without their support, encouragement and love.

My dear aunt, I am grateful for all the chocolate reserves...

My dad, I appreciate the energy you have gathered to fetch me from the bus stop to home every evening that I worked late for the sake of this thesis...

I am thankful to my beautiful sister for everything but especially for the Latex support. I remember the days you were in the swing across the living room with the tiniest and the cutest face and hands. Although that makes me feel a little old to see you growing, I am proud to see how intelligent, sweet and tenderhearted you have become.

My warm, beloved mom, I owe you my whole life... I do not know where to start to thank you, I am grateful for everything, starting from tight hugs to the pastries that you have baked to cheer me up, everything from A to Z... Nothing would have been the same without your help, support and tendernes... You are the best mom that a human being can be. I wish I will be a such a good mom as you, at least close to you.

This thesis is dedicated to my dear parents and sister whom I owe everything.

ABSTRACT

TAGGING PROTEINS: SUMOYLATION MECHANISM BY MOLECULAR MODELING AND MOLECULAR DYNAMICS SIMULATIONS

Sumoylation, the covalent attachment of SUMO (Small Ubiquitin Like Modifier) to target proteins is a posttranslational modification that can alter intracellular localization, interactions with other proteins or result in modifications by other post-translational modifiers. Malfunctioning of sumoylation pathway is concerned with many neurological diseases, such as Huntington's disease, Parkinson's disease and more. Besides, sumoylation is a part of cancer related pathways. Like other ubiquitin like modifier (Ubl) conjugation mechanisms, the covalent attachment of SUMO to target proteins involves three classes of enzymes: Ubiquitin-like protein activating enzyme (E1), Ubiquitin-like protein conjugating enzyme (E2) and Ubiquitin-like protein ligase (E3). Sumoylation is different than other Ubl conjunction paths in a way that, the E2 ligase, Ubc9, can function with lower reaction efficiency in the absence of an E3 ligase. One of the target proteins that Ubc9 can uniquely sumoylate with high reaction efficiency is RanGAP1. This work is based on the Ubc9-SUMO-RanGAP1-RanBP2 system. Presense of RanBP2 is thought to possess an allosteric effect on Ubc9 and SUMO, thereby increasing the efficiency of the transfer of the bond from Ubc9 to RanGAP1. Molecular Dynamics (MD) simulations Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 complexes are carried out at 300 K including the isopeptide bond between Gly97 of SUMO and 524Lys of RanGAP1. To obtain coherent results, parallel simulations of both structures are carried out. Expectedly, RanBP2 behaves as an E3 ligase for the system and allosterically enhances the conformational stability of the Ubc9-SUMO-RanGAP1 structure. Coupling of SUMO and RanGAP1 is found to dependent on couplings of Loop 2 with Thr511-Leu522 and Leu555-Pro566 regions of RanGAP1.

Furthermore, the presence of RanBP2 is shown to impose a packing motion for the Loop 2 region, which is Lys30-Met36 region of Ubc9 (Karaca *et al.*, 2010). Asp33 residue is observed to maintain a unique conformational state, where Asp33 bends towards RanGAP1. In order to clarify the role of Asp33 of Ubc9, Asp33Ala mutant is designed. With the mutation, the significance of Asp33 residue, more importantly the significance of Loop 2 region is highlighted.

ÖZET

PROTEİNLERİN ETİKETLENMESİ: MOLEKÜLER MODELLEME VE MOLEKÜLER DİNAMİK SİMULASYONLARI İLE SUMOLANMA MEKANİZMASI

Sumolanma, SUMO (Küçük Ubikütın Benzeri Değistirici) proteininin hedef proteinlere kovalent bağlanması ile gerçekleşen hücre içi yerleşimi, diğer proteinler veya diğer post-translasyonel düzenleyiciler tarafından değişiklikleri sonucu ile etkileşimleri değiştirebilir bir posttranslasyonel modifikasyonudur. Sulanma patikasindaki aksaklıklar Huntington hastalığı, Parkinson hastalığı ve bunlar gibi daha birçok nörolojik hastalık ile ilintilidir. Dahası, kanser-ilintili patikalarin da bir parçasıdır. Diğer ubikütın benzeri değistirici mekanizmaları gibi, SUMO'nun hedef proteine kovalent bağlanması da üç enzim sınıfının varlığını gerektirir: ubikütın benzeri proteini aktive edici enzim (E1), ubikütın benzeri protein bağlayıcı enzim (E2) ve ubikütın benzeri protein ligazi (E3). Sumolanma diğer ubikütın benzeri proteinlerin bağlanma mekanizmalarından farklıdır, şöyle ki, E2 ligazi olan Ubc9, E3 yokluğunda da düşük verimle işlevi sürdürebilir. Ubc9'in yüksek verimle sumolayabildigi hedef proteinlerinden birisi RanGAP1'dir. Bu çalışma, Ubc9-SUMO-RanGAP1-RanBP2 sistemi üzerine dayanmaktadır. RanBP2'nun varlığı Ubc9 ve SUMO üzerinde bir allosterik etki yarattığı düşünülmektedir, bu etki sayesinde bağın Ubc9'dan RanGAP1'e transfer olma verimliliğinin arttığı bilinmektedir. Ubc9-SUMO-RanGAP1-RanBP2 ve Ubc9-SUMO-RanGAP1 kompleksleri üzerinde Moleküler Dinamik (MD) simülasyonlari 300 K sıcaklıkta ve Gly97 SUMO ve 524Lys RanGAP1 arasında izopeptit bağı da dahil olmak üzere gerçekleştirilmiştir. Tutarlı sonuçlar elde edebilmek için, her iki kompleksin de paralel simulasyonlari yapılmıştır. Beklenildiği üzere, RanBP2 Ubc9-SUMO-RanGAP1 sistemi için bir E3 ligaz olarak davranır ve alosterik olarak Ubc9-SUMO-RanGAP1

yapısının konformasyonel stabilitesini artırır. SUMO ve RanGAP1 kuplajinin, Loop 2 ve RanGAP1 bölgelerinin (Thr511-Leu522 ve Leu555-Pro566) etkileşimine bağlı olduğu görülmüştür.

Ayrıca, RanBP2 varlığının, Ubc9'nin Lys30-Met36 arasında kalan, Loop 2 bölgesinde bir katlama hareketine neden olduğu görülmüştür (Karaca *et al.*, 2010). Asp33 rezidusu, RanGAP1'e doğru eğilerek benzersiz bir konformasyonel durum oluşturur. Ubc9 üzerindeki Asp33'un rolünü netleştirmek için ise Asp33Ala mutasyonu tasarlanmıştır. Mutasyonla beraber, Asp33 rezidusunun önemi daha da önemlisi Loop 2 bölgesinin önemi vurgulanmıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic Acid
AMP	Adenosine Monohosphate
ATP	Adenosine Triphosphate
Cryo-EM	Cryogenic Electron Microscopy
Cys	Cysteine
DA	Description of abbreviation
GAFF	General Amber Force Field
Glu	Glutamic Acid
Gly	Glycine
His	Histidine
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
MD	Molecular Dynamics
Met	Methionine
mRNA	Messenger Ribonucleic acid
NLS	Nuclear Localization Signal
NMR	Nuclear Magnetic Resonance
NPC	Nuclear Pore Complex
Phe	Phenylalanine
Pro	Proline
RanBP2	Ubiquitin-conjugating enzyme
RanGAP1	RanGTPase-activating protein1
RMSD	Root mean squared deviation
SBM	SUMO Binding Motif
Ser	Serine

SUMO	Small Ubiquitin-like Modifier
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Ub	Ubiquitin
UBC	Ubiquitin Conjugating Enzyme
Ubl	Ubiquitin-like
Val	Valine
3-D	3 Dimensional
$\langle\ \rangle$	Ensemble average

1. INTRODUCTION

1.1. Biological Background

1.1.1. Ubiquitin and Ubiquitin-like Proteins Superfamily and Posttranslational Protein Modification

Posttranslational protein modifications are essential in regulating protein activity, function, stability, intra-cellular localization and interactions with various proteins or ligands. One of these modifications entails a small protein, namely ubiquitin.

Ubiquitin, which is highly conserved in all eukaryotic cells and comprises of 76 amino acid residues, can be conjugated to other proteins posttranslationally. It can exist in free form or as a part of complex protein in which ubiquitin is attached by isopeptide link at the C-terminal to the ε-amino group of the acceptor Lysine residues on the target proteins. This kind of attachment of ubiquitin to proteins is known as ubiquitination. Ubiquitination of cellular proteins targets those proteins to the proteasome-mediated degredation by the 26S proteasome which is the major protease of the cytosol and the nucleus of eukaryotes. (Jentsch and Pyrowolakis, 2000)

More generally, Ubiquitin (Ub) and Ubiquitin like (Ubl) proteins are modifiers which mediate the post-transcriptional modification of specific target proteins. The function of posttranslational modification of proteins by covalent attachment of Ubl proteins is not merely the labeling for degradation but a broader phenomenon. Post-translational modifications can regulate the activity, interactions with other proteins or the level of degradation of the modified protein. Members of the Ub/Ubl family take place in cellular processes such as cell differentiation, apoptosis, cell cycle and stress response. The Ub/Ubl proteins comprise approximately of 100 amino acids and they can be covalently bound to Lysine residues on target proteins via various mechanisms. (Capilli and Lima, 2007,2)

1.1.2. Mechanism of Ub/Ubl conjugation

Ubiquitination, i.e. ubiquitin binding to target proteins is crucial for targeting proteins to degradation. Ubiquitin conjugation is a complex processes that necessitates multiple steps which are facilitated by specific set of enzymes (Wilkinson, 1995). The process starts with activation of ubiquitin by the ubiquitin activating enzyme E1. This ATP dependent activation involves the adenylation of the C-terminus of the ubiquitin via the release of pyrophosphate which is followed by a thioester bond formation between the carboxyl group of the C-terminal and the side chain of Cysteines in E1 along with a release of adenosine mono-phosphate (AMP). After the acetylation reaction and the release of AMP, the activated ubiquitin is transferred to the conserved Cysteine residue of the ubiquitin conjugating enzyme E2 by a transacetylation reaction to form an E2:ubiquitin thioester intermediate. (Hershko *et al.*, 1983) Once ubiquitin is conjugated to E2 by a thioester bond, it is transferred to target protein. The target and the ubiquitin are bonded to each other by means of an isopeptide bond between the ε-amino group of a Lysine in the target protein and the carboxyl group of the C-terminal of the ubiquitin. In many cases this transfer of ubiquitin from E2 to target is achieved by an E3 ligase (Wilkinson, 1995). (Yeh *et al.*, 2000)

1.1.3. SUMO: Small Ubiquitin-like Modifier

Small Ubiquitin-like Modifier (SUMO) family is a member of the Ub/Ubl superfamily. SUMO is also known as Sentrin, UBL1, GMP1 and PIC1 (Boddy *et al.*, 1996; Matunis *et al.*, 1996; Okura *et al.*, 1996; Shen *et al.*, 1996; Mahajan *et al.*, 1997; Zhao, 2007). It is present in all eukaryotes and the equivalent proteins Smt3 and Pmyt3 are present in *Saccharomyces cerevisiae* and *Saccharomyces pombe* respectively. SUMO consists of 101 amino acids and it possesses 18% identical and 48% similar to ubiquitin (Okura *et al.*, 1996; Kamitani *et al.*, 1997; Mahajan *et al.*, 1997; Bayer *et al.*, 1998). The 18% of sequence homology is low but SUMO-1, which is a member of SUMO family and found in human, contains the $\beta\beta\alpha\beta\beta\alpha\beta$ fold of the Ub/Ubl superfamily and also the 3-D structure of SUMO-1 is very similar to that of ubiquitin (Bayer *et al.*, 1998; Kretz-Remy and Tanguay, 1999).

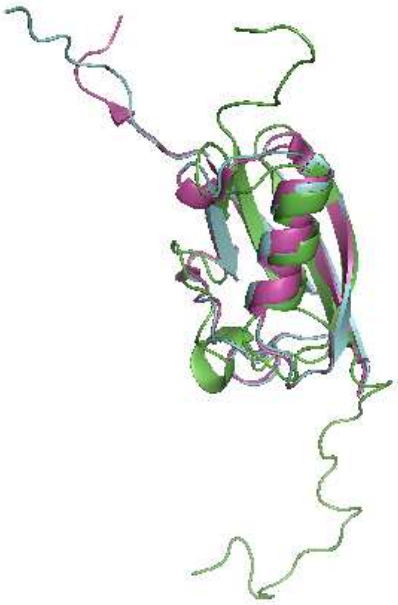

Figure 1.1. Structures of SUMO-1 (Green: PDB id 1A5R), SUMO-2 (Blue: PDB id 2IO0) and SUMO-3 (Pink: PDB id 2IO1) proteins. Structures are visualized via PyMol (DeLano, 2002).

Although the 3-D structure of SUMO-1 is very similar to that of ubiquitin, there are significant differences between the two proteins. The C-terminus of SUMO-1 must be cleaved proteolytically to have a mature terminal with Gly-Gly motif which is required for isopeptide bond formation between SUMO-1 and target proteins (Kamitani *et al.*, 1997; Yeh *et al.*, 2000). In addition to this, SUMO contains extra 21 amino acid residues at the N-terminus which brings about a highly flexible tail (Bayer *et al.*, 1998). Although the exact function of this tail is yet unknown, looking at the charged amino acid rich characteristics, it can be said that this tail is a perfect region for SUMO-specific protein-protein interactions (Bayer *et al.*, 1998). Besides, the two proteins have significantly different surface charge distribution. One distinct feature of SUMO is that it has a large negatively charged surface formed by Glu83, Glu84, Glu85 and Asp86. In addition to that a negatively charged pocket is formed by Glu11, Asp12, Glu15, Glu20 and Glu83. On the opposite site of this negatively charged regions, there is a distinctly positively charged surface consisting of Lys16, Lys17, Lys23, Lys25, Lys46 and Lys54. Out of those listed charged residues, only Glu83 and Lys25 are

conserved in ubiquitins. (Bayer *et al.*, 1998)

Unlike ubiquitination, SUMO-1 does not target proteins to the proteasome; conjugation of SUMO-1 has very diverse functions. SUMO conjugation and deconjugation regulates cellular pathways, provides shuttling of proteins between the cytosol and the nucleus, assembly and disassembly of nuclear bodies, stabilization and destabilization of proteins, activation and repression of transcriptional factors, regulation of chromosome function, genomic integrity as well as protein-protein and protein-DNA interactions (Gill 2003; Kracklauer and Schmidt, 2003; Verger *et al.*, 2003; Seeler and Dejean, 2001; Pountney *et al.*, 2003; Hilgarth *et al.*, 2004; Girdwood *et al.*, 2004; Gill, 2004; Dohmen, 2004; Müller *et al.*, 2004; Lieberman, 2004; Jacquiau *et al.*, 2005; Smallridge, 2006; Vertegaal *et al.*, 2006; Iniguez-Lluhi, 2006; Herrmann *et al.*, 2007; Zhao, 2007; Palancade and Doye, 2008).

So far there are over 100 targets known for SUMO-1 conjugation in which SUMO-1 is bonded to specified sites on target proteins with the help of E1, E2 and E3 ligases (Wolhschlegel *et al.*, 2004; Zhao *et al.*, 2004; Zhou *et al.*, 2004; Herrmann *et al.*, 2007).

In addition to SUMO-1, there are three more SUMO structures found in mammalian cells and those are named as SUMO-2, SUMO-3 and SUMO-4. The phylogenetic studies revealed that SUMO-2 and SUMO-3 genes are more closely related to each other than SUMO-1 genes (Su and Li, 2002) which is consistent by the fact that the SUMO-2 and SUMO-3 are 95% identical to each other, however, the sequence similarity of SUMO-2 and SUMO-3 with SUMO-1 is 42 per cent and 43 per cent respectively (Capilli and Lima,2007,1). Besides, there is structure similarity between the $\beta\beta\alpha\beta\beta\alpha\beta$ fold of the Ub/Ubl superfamily and the SUMO-1, SUMO-2 and SUMO-3 (Figure 1.1). In consistency with the ubiquitin, in cells SUMO has the potential to form polymeric chains via the consensus sites for sumoylation that are present in SUMO-2 and SUMO-3 (Vertegaal, 2007). Although poly-sumoylated proteins by SUMO-1 are observed *in vitro* via three N-terminal lysines (Lys7, Lys16 and Lys17) (Yang *et al.*, 2006; Pedrioli *et al.*, 2006), SUMO-1 cannot form poly-sumoylated proteins *in vivo* (Tatham *et al.*, 2001; Jentsch and Pyrowolakis, 2000; Yeh *et al.*, 2000; Vertegaal, 2007). The 48Lys

in ubiquitin which is known to be responsible in polyubiquitin chains formation is replaced by a 69Gln in SUMO-1. This difference serves as an explanation for the inability of SUMO-1 at poly-sumoylated chains formation (Bayer *et al.*, 1998; Kretz-Remy and Tanguay, 1999; Hay, 2001; Kim *et al.*, 2002). *In vitro* E3 ligases are known to enhance the SUMO polymerization (Vertegaal, 2007). Furthermore, a SUMO binding motif (SBM) in Ubc9 helps the SUMO polymerization *in vitro* (Vertegaal, 2007).

SUMO-1 and SUMO-2/-3 share the high sequence similarity and many of the same target proteins, however, the functional significance of modification by SUMO-1 and SUMO-2/-3 might be different. (Tatham *et al.*, 2001) Furthermore, the cellular localization of SUMOs are different from each other. SUMO-1/-2/-3 are highly concentrated at the nuclear membrane, nuclear bodies and cytoplasm respectively indicating that they have different functions in different compartments of the cell (Su and Li, 2002).

Modification by SUMO has role in many cellular processes such as activation of transcription factors, nuclear import and cell division. Furthermore, sumoylated proteins can change intracellular localization, their interactions with other proteins and modifications by other post translational modifiers. Besides, sumoylation has part in neurological pathways since defects in SUMO conjugation are observed to lead some neurological disorders such as Huntington's, Parkinson's and Alzheimer's diseases (Martin *et al.*, 2007; Zhao, 2007; Sarge and Sarge, 2009) along with type 1 diabetes (Li *et al.*, 2005) and familial dilated cardiomyopathy (Sarge and Sarge, 2009; Pouladi, 2009). Aside of all these, SUMO has a role in cancer related pathways, in that sense it can be said that SUMO is involved in certain types of cancer such as ovarian, breast carcinoma, melanoma and lung adenocarcinoma (Alarcon-Vargas and Ronai, 2002; Mo and Moschos, 2005; Mo *et al.*, 2005; Kim and Baek, 2006; Karamouzis *et al.*, 2006; Moschos and Mo, 2006; Cheng *et al.*, 2006; Wu and Mo, 2007; Sarge and Sarge, 2009).

1.1.4. Sumoylation Mechanism

Covalent binding of SUMO (Small Ubiquitin-like Modifier) is named as sumoylation. Sumoylation is a post-translational modification similar to ubiquitination. Although there is a high degree of similarity between SUMO and ubiquitin, their functions are very diverse. In contrast with ubiquitin, SUMO modification does not create any signal for degradation, in other words, SUMO does not take part in proteolysis. (Yeh *et al.*, 2000) Furthermore, due to findings with Mdm2 and IκBα it is suggested that SUMO prevents of the function of ubiquitin by binding to the target and thereby blocking the ubiquitin binding site *Muller* et al.*, 2001.* Similar to ubiquination, sumoylation is reversible (Matunis *et al.*, 1996; Mahajan *et al.*, 1997). It is suggested that this reversible mechanism of sumoylation might be an indication that SUMO dynamically regulates the target proteins. (Müller *et al.*, 2001) Like other Ubl conjugation mechanisms, sumoylation is concerned with three groups of enzymes: Ubiquitin-like protein activating enzyme (E1), Ubiquitin-like protein conjugating enzyme (E2) and Ubiquitin-like protein ligase (E3).

As can be expected from the high degree of similarity between SUMO and ubiquitin, the general mechanism of conjugation and the enzymes themselves are related. SUMO, which is produced as an inactive precursor, contains a C-terminal extension of four amino acids (HSTV). This extension needs to be proteolytically removed for maturation. The cleavage of SUMO produces the C-terminal Gly-Gly motif, which is conserved in Ub/Ubl family (Mahajan *et al.*, 1997; Mahajan *et al.*, 1998; Bernier-Villamor *et al.*, 2002). Next, E1 enzyme Aos1/Uba2 adenylates SUMO C-terminus and this forms a thioester bond between E1 active Cysteine and SUMO, in an ATM dependent reaction. In contrast to the ubiquitin-activating enzyme, the SUMO-E1 is a heterodimer consisting of a 40- and a 70-kDa subunit (Aos and Uba2, respectively). At the second step of sumoylation, the SUMO thioester is transferred to the active Cysteine of the E2 enzyme (Melchior, 2000). The interaction between SUMO and Ubc9 requires the ubiquitin domain and the C-terminal Gly-Gly residues of SUMO. In contrast to several E2 enzymes used in ubiquitin conjugation, Ubc9 is the single known E2 enzyme for sumoylation. (Bernier-Villamor *et al.*, 2002; Gong *et al.*, 1997)

The final step of conjugation is transfer of SUMO from active Cysteine in Ubc9 to a Lysine residue on the target protein. E3 enzymes that ensure target specificity and increase reaction efficiency usually mediate the transfer of SUMO from Ubc9 to target. In contrast to the classical E1/E2/E3 conjugation mechanism, sumoylation can also proceed without an E3 enzyme in some specific cases. It is shown that among the sumoylation of target proteins, RanGAP1, p53 and IκBα are modified without an E3 ligase *in vitro*, although the rates of reactions are considerably lower compared to E3 mediated conjunction (Melchior, 2000).

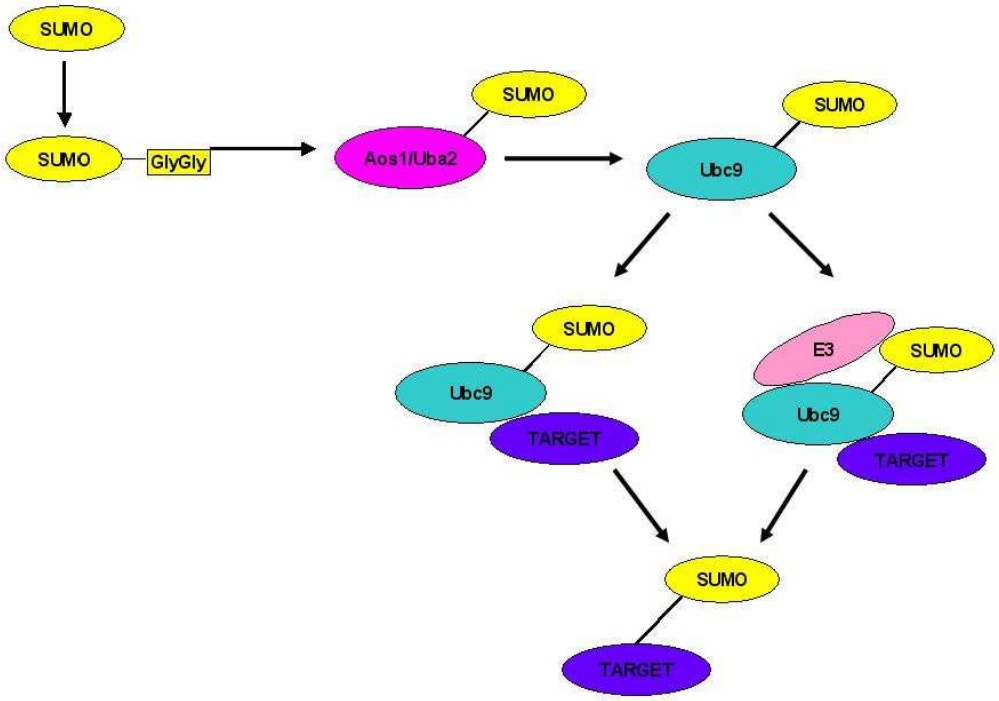

Figure 1.2. Sumoylation Mechanism. SUMO is cleaved at C-terminus. Activated SUMO is transferred on the active cysteine of E1 enzyme, Aos1/Uba2. The SUMO thioester is transferred to the active cysteine of E2, Ubc9. SUMO is transferred to the target. The conjugation can be mediated by an E3 enzyme or directly by Ubc9.

Sumoylation is different than other Ub/Ubl conjunction paths in a way that, the E2 ligase, Ubc9, can function with lower reaction efficiency in the absence of an E3 ligase. One of the target proteins that Ubc9 can uniquely sumoylate with high reaction efficiency is RanGAP1. There are several suggested models, yet the actual mechanism is not known. Presence of RanBP2 is known to possess an allosteric effect on Ubc9 and SUMO, thus increasing the efficiency of the transfer of the bond from Ubc9 to

RanGAP1.

1.1.5. E2 ligase Ubc9

E2 ligases are responsible for the transfer of activated Ub/Ubl modifier to the site of conjugation in the form of E2-Ub/Ubl thioester intermediate (Hershko *et al.*, 1983). It has been observed that the addition of SUMO conjugating enzyme Ubc9 provides the efficient transfer of thioester linkage from E1 to Ubc9, which lead to the formation of Ubc9-SUMO thioester intermediate (Desterro *et al.*, 1999). Ubc9 is the only E2 enzyme for SUMO family members and it is known that it is specific for SUMO since Ubc9 do not form thioester linkage with ubiquitin but SUMO. (Gong *et al.*, 1997; Johnson and Blobel, 1997; Desterro *et al.*, 1997; Melchior, 2000; Saitoh *et al.*, 1998; Lee *et al.*, 1998; Giraud *et al.*, 1998; Schwarz *et al.*, 1998) Besides, transacetylation of Ubc9 via SUMO binding is not associated with ubiquitin activating enzyme E1, this occurs through a specific enzymatic activity implying that SUMO conjugation to target proteins is parallel to ubiquitin machinery but by a distinct pathway (Desterro *et al.*, 1997). In that sense, Ubc9 can be considered as a SUMO conjugating enzyme rather than a ubiquitin conjugating enzyme.

Ubc9 has high sequence similarity with the ubiquitin conjugating enzymes and similar 3-D structure however the overall charge of Ubc9 is much more positive compare to other ubiquitin conjugating enzymes. (Tong *et al.*, 1997; Giraud *et al.*, 1998) The E2 enzyme, Ubc9 is a highly conserved conjugating enzyme. It has a conserved UBC (Ubiquitin Conjugating Enzyme) superfold, which is found in all of the E2 ligases of the Ub/Ubl superfamily. Among the E2 ligases there are some insertions of amino acids in to UBC domain. For Ubc9 two amino acid insertions are observed, those are Asp100 and Lys101 (Giraud *et al.*, 1998; Tatham *et al.*, 2003). Besides, there are two major binding sites for target on Ubc9; the region between Ala131 and Gln139 and the region between Lys74-Phe77 with the catalytic loop that contains the active Cys93. (Bernier-Villamor *et al.*, 2002) The first region helps Ubc9 to recognize the target, RanGAP1 and to form a stable complex with RanGAP1 by acting as an E3 ligase (Hochstrasser, 2002). Explaining how the system can work in the absence of E3, Ala131-Gln139 region

serves as an "built in E3" (Pichler *et al.*, 2004). Second region consisting of Lys74-Phe77 is the region that falls around the catalytic Cys93. The sulfydryl group of Cys93 is in close vicinity to the carboxyl groups of Glu98 and Asp127 and hydroxyl group of Thr87. The carboxyl group of Asn85 and nitrogen atom of Cys93 can form a hydrogen bond. This catalytic region is further rigidified by an extra hydrogen bond between nitrogen atom of Tyr87 and carbonyl group of Thr91. His83, Pro84, Asn85, Tyr87, Ile96, Leu97, Asp127, Pro128 and Ala129 are highly conserved implying that these residues might have role in maintaining the active Cysteine in the right conformation. (Bayer *et al.*, 1998) Out of these conserved residues, Asn85, Tyr87 and Asp127 are particularly very important. It is suggested that these residues help to suppress the pK of the lysine of the target and thus the reaction rate is enhanced (Yunus and Lima, 2006). One can observe from the 3-D structure that those three residues are in fact a part of the catalytic cleft formed by Tyr87, Asp 127, Pro128 and Ala 129 (Yunus and Lima, 2006; Bernier-Villamor *et al.*, 2002; Reverter and Lima, 2005).

Moreover, Lys101 of Ubc9 generates a positively charged lip close to the active Cysteine. This lip may distinguish between ubiquitin and SUMO by repelling the positively charged Arg74 of ubiquitin (Bayer *et al.*, 1998).

Addition to conjugating SUMO to RanGAP1, it is suggested that Ubc9 is also essential for formation of a stable ternary complex consisting of SUMO-modified RanGAP1 and RanBP2 (Zhu *et al.*, 2006). It is also pointed out that the Ubc9-RanGAP1 interactions at around Phe564-Lys526 of RanGAP1 are required for the Nuclear Pore Complex localization (Zhu *et al.*, 2006).

1.1.6. E3 ligase RanBP2

In general, the function of E3 ligases is to provide the substrate recognition and to provide high degree of specificity. In ubiquitin machinery, the role of E3 ligase is the catalysis of amide bond formation between Ub/Ubl and proteins, in other words, E3 ligase provides the transfer of ubiquitin from the E2-Ub/Ubl intermediate to the stable Ub/Ubl-protein complexes. (Hershko *et al.*, 1983) In many cases, the E3 is

comprised of a multiprotein complex which recognizes the substrate and provides the contact of substrate with the E2 which catalyzes conjugation of Ub/Ubl to the substrate. In the ubiquitination machinery, the substrate specificity is provided by the E3 ligase but in contrast, in the sumoylation machinery does not always require an E3 enzyme as Ubc9 provides the substrate specificity and thus conjugation of SUMO-1 to RanGAP1 can function without the E3 ligase, RanBP2. (Hay, 2001) The available structure of SUMO-RanGAP1-Ubc9-RanBP2 stands as a support to this idea and suggests that RanBP2 increases the efficiency of the reaction through reducing the conformational flexibility of the Ubc9-SUMO intermediate to enhance conjugation (Reverter and Lima, 2005). RanBP2 exerts its catalytic effect by changing the properties properties of Ubc9, rather than by mediating target interactions directly (Pichler *et al.*, 2002; VanDemark and Hill, 2002; Pichler *et al.*, 2004). RanGAP-1 is normally not a target for RanBP2 however the role of RanBP2 is thought to be responsible for arranging the SUMO thioester on the Ubc9 in an optimal conformation for catalytic transfer of the target RanGAP1 (Knipscheer, 2007) and prevent non-productive conformations of Ubc9-SUMO (Reverter and Lima, 2005).

Generally, ubiquitin E3 ligases can be classified in to two categories; RING and HECT. RING type E3 ligases mediate transfer of Ub/Ubl to the target through binding both to E2-Ubl and to target proteins. HECT type E3 ligases are the so-called true enzyme implying that they first transfer the Ub/Ubl on themselves from the E2 enzyme, forming a thioester bond, and then transfer the Ub/Ubl onto the target (Pichler *et al.*, 2002; Reverter and Lima, 2005). However, RanBP2 is neither HECT nor a RING type E3 ligase. It is a protein taking roles at the NPC and is known to form stable complex with a SUMO target, RanGAP1. Although RanBP2 is an E3 ligase for the sumoylation machinery, it does not belong to any of those two categories. It is separate class of enzyme. (Geiss-Friedlander and Melchior, 2007) This difference indicates that RanBP2 catalytic activity is different from ubiquitin E3s. (Ferrier, 2002)

RanBP2 is a large nucleoporin which is localized in the cytoplasmic fibrils of the nuclear pore for protein import substrate docking before the translocation. (Melchior, 1995; Pichler *et al.*, 2002) A fragment of RanBP2 between residues 2633-2761, con-

sisting of IR1-M-IR2 domains, is sufficient for E3 activity in vivo and in vitro. More specifically, IR1-M, M-IR2 and IR1-M-IR2 constructs are also functional as an E3 ligase and IR1-M is the catalytic core domain (Pichler *et al.*, 2002; Pichler *et al.*, 2004; Reverter and Lima, 2005; Tatham *et al.*, 2005). The IR domain provides E3 ligase activity whereas the M domain increases the affinity of Ubc9 binding. Although being catalytically active, the IR-1 domain lacks regions that provide target specificity which in our case achieved by Ubc9 (Pichler *et al.*, 2002; Pichler *et al.*, 2004). Although the internal repeat (IR1/2) domains of RanBP2 has no sequence similarity with known E3 enzymes of the ubiquitin family; however it functions as an E3 ligase for SUMO as it enhances the sumoylation of target RanGAP1. RanBP2 can be said to provide substrate specificity as it does not enhance or only a few of the SUMO targets. The IR domain can form a stable complex with SUMO, RanGAP1 and Ubc9. (Saitoh *et al.*, 1997; Saitoh *et al.*, 1998; Matunis *et al.*, 1998; Pichler *et al.*, 2002) It is identified that the V/I-X-V/I-V/I SUMO binding motif (SBM) is found in all proteins that can bind SUMO (Song *et al.*, 2004; Hecker *et al.*, 2006; Kerscher, 2007). The arrangement of the SBM is not important; SBM can also be arranged as V/I-V/I-X-V/I (Song *et al.*, 2005; Kerscher, 2007). In our case the RanBP2 has this motif between residues 2632-2635 (2632Val-2633Leu-2634Ile-2635Leu) (Song *et al.*, 2004). This region is important as it provides the E3 ligase activity in sumoylation machinery.

RanBP2-bound RanGAP1 is found to be required for nuclear protein import. Besides, sumoylated RanGAP1 forms a stable complex with RanBP2 at the NPC. At this point, it is interesting to underline that RanBP2 functions as a docking factor in transport as well as an E3 ligase in sumoylation (Pichler *et al.*, 2002). Among all E3s, this multiple functioning is unique, implying that the catalytic activity of RanBP2 can be different than the usual E3 enzymatic activity. RanBP2 interact with the sumoylated RanGAP1 at the NPC via the SBM.

Ubc9 and RanBP2 interaction sites reside on the opposite site of active Cysteine93 of Ubc9. The interacting residues are not conserved among different species considering the E2-E3 interactions (Tatham *et al.*, 2003). Important residues for the Ubc9 and RanBP2 binding and E3 activity are listed as: Leu 2651, Leu 2653, Phe2657, Phe2658

(Pichler *et al.*, 2004).

1.1.7. Sumoylation Target RanGAP1

Ran is a nuclear Ras-like GTPase and it is necessary for mRNA processing, maintenance of structural integrity of nuclei, cell cycle control and the transport of proteins and ribonucleoproteins into and out of the nuclear pore complex (NPC). (Gŏrlich and Mattaj 1996; Rush *et al.*, 1996; Sazer, 1996; Zhang *et al.*, 2002) Like all other Ras-related GTPases, Ran functions as molecular switch between GTP and GDP bound conformations. RanGAP1 is the 70-kD Ran-GTPase-activating protein which functions as a key regulator of the Ran GTP/GDP cycle.

RanGAP1 is found in the cytoplasmic periphery of the NPC at high concentrations when it is bound to RanBP2. Unmodified RanGAP1 resides in the cytosol but modification of RanGAP1 by SUMO is known to be essential for the translocation of RanGAP1 to the cytoplasmic fibers of the NPC. Conjugation of RanGAP1 with SUMO forms a metabolically stable complex that targets to the NPC. It is indicated that essential targeting information of RanGAP1 to the NPC is in the tail of RanGAP1 (Mahajan *et al.*, 1998). Experimental studies (Matunis *et al.*, 1996) proved this by showing that the unmodified 70-kD form of RanGAP1 is highly cytosolic, but the modified 90-kD form is found on the cytoplasmic fibers of the NPC where it associates with the 358-kD Ran-GTP-binding protein RanBP2. (Matunis *et al.*, 1996; Mahajan *et al.*, 1997; Saitoh *et al.*, 1997; Kamitani *et al.*, 1997; Matunis *et al.*, 1998; Zhang *et al.*, 2002; Pichler *et al.*, 2002; Pichler and Melchior, 2002) Cytosolic RanGAP1 is not required for nuclear protein import, however association of RanGAP1 with RanBP2 is found to be required for nuclear protein import (Mahajan *et al.*, 1997; Mahajan *et al.*, 1998; Zhang *et al.*, 2002; Pichler and Melchior, 2002; Melchior *et al.*, 2003).

It is shown that the binding of SUMO to RanGAP1 do not change the structure nor the dynamics of the either of the proteins except the site of the isopeptide linkage. Hence, RanBP2 is capable of recognizing the binary structure of SUMO:RanGAP1 at the nuclear envelope once they are connected via an isopeptide linkage. (Macauley *et*

al., 2004) SUMO modification creates an E3 binding site in the C-terminal domain of RanGAP1 which targets the RanGAP1 to the NPC. Interestingly, amino acids 541-589 in the C-terminal domain of RanGAP1 are found to be sheltering a nuclear localization signal (NLS). This C-terminal region is thought to have regions that associate with recognition of RanGAP1 by E2 and E3 ligases. Transport of proteins in and out of the nucleus involves regulation of processes by intrinsic target signals such as the NLS and the nuclear export signals in shuttle proteins and structural complexes functioning at the NPC as well. Considering the nuclear localization signal together with the presence of nine leucine rich domains which are potential nuclear export signals in the N-terminal, it is concluded that RanGAP1 may transport back and forth between nucleus and the cytoplasm (Matunis *et al.*, 1998; Mahajan *et al.*, 1998).

Among the over hundred targets for SUMO, RanGAP1 is the major substrate. A consensus motif of 4 amino acids, Ψ-K-X-D/E (Ψ represents a hydrophobic residue, K is the SUMO acceptor Lysine, X is any amino acid and D/E is an acidic residue) is found in nearly all SUMO targets (Yeh *et al.*, 2000; Watts, 2004; Johnson, 2004; Duda and Schulman, 2005; Heun, 2007; Geiss-Friedlander and Melchior, 2007; Martin *et al.*, 2007). It has been suggested that the first and the third residues of this binding motif are preferably aromatic residues (Schwamborn *et al.*, 2008). The SUMO consensus motif on RanGAP1 is LKSE is found between the residues 525-528. (Bernier Villamor *et al.*, 2002; Martin *et al.*, 2007)

Unlike the ubiquitination machinery, the substrate specificity in sumoylation machinery might be provided by recognition of this conserved motif by the Ubc9-SUMO-1 thioester intermediate (Hay, 2001). In that sense, the consensus motif is a major determinant of Ubc9 binding and SUMO-1 modification (Sampson *et al.*, 2001). However, presence of the sumoylation motif is not always sufficient for sumoylation to occur (Wilkinson, 2008). Secondary structure is appears as another major determinant. A protein having high probability sumoylation sites but if the binding sites are hidden deep inside the core of the protein structure, the sumoylation is unlikely to occur unless a conformational change happens (Wilkinson, 2008). At this point, Ubc9 binding may modify the relatively unstructured regions of the proteins allowing binding to lysine

residues to SUMO (Wilkinson, 2008). Besides that, for Ubc9 it has been suggested that the binding of Ubc9 to the consensus motif might be reason that sumoylation can function in the absence of E3 since this novel motif provides specificity for Ubc9 binding and SUMO-1 modification (Sampson *et al.*, 2001). In sumoylation, the C-terminal Glycine 97 of SUMO-1 makes an isopeptide bond to ε-amino group of the 526Lys of RanGAP1 on that conserved motif (Mahajan *et al.*, 1998). The fact that mutating RanGAP1 at 526Lys blocks the SUMO-binding activity, 526Lys is the only site that sumoylation can occur (Mahajan *et al.*, 1998).

Aside of the conserved motif, RanGAP1 has an additional contact area for Ubc9 which provides higher efficiency of conjugation than other substrates. This region resides between 511-522 and 555-566 on RanGAP1. (Bernier Villamor *et al.*, 2002) Residues between 420 and 471 on RanGAP1 are important for its interactions with the E3 ligase RanBP2. (Wilson and Rangasamy, 2001)

1.1.8. Interaction Summary

1.1.8.1. Interactions of SUMO. The E2 enzyme for sumoylation machinery; Ubc9 can form both covalently (specific binding) and non-covalently (non-specific binding) bound complexes with SUMO (Capili and Lima, 2007,1; Reverter and Lima, 2005). In the non-specifically bounded Ubc9-SUMO complex, the structure shows direct interactions between seven Ubc9 residues and eleven SUMO residues within a cut off distance of 3.8 A. These residues can be listed as follows: Arg13, Arg17, Lys18, Phe22, Gly23, Val27, and Lys49 and Lys25, Gly28, Gln29, Arg63, Glu67, Gly81, Glu83, Asp86, Val87, Glu89, and Tyr91, respectively for Ubc9 and SUMO. On the contrary, an NMR chemical shift perturbations study revealed 20 SUMO-1 residues and 25 Ubc9 residues. Those are (Ile27- Ser32, Val38, Arg63-Ile71 and Met82-Val90 regions of SUMO-1 and Ser7, Ala10, Lys14, His20, Phe22, Val27, Thr35, Asn37, Ala44, Lys49, Leu60, Met62, Leu63, Glu99, Ile109, Lys110, Ile112, Leu113, and Ser158 regions of Ubc9 (Capili and Lima, 2007b; Liu *et al.*, 1999; Tatham *et al.*, 2003).

The SUMO non-specific binding sites of Ubc9 overlap with the interaction sites of

Ubc9 with E3. However, the specific and non-specific binding sites to SUMO on Ubc9 resides on different surfaces of Ubc9, implying that Ubc9 can interact with two SUMO molecules simultaneously while one being covalently bound to the active cysteine, the other being non specifically bounded. Moreover, it is pointed out that the non-specific binding of SUMO-Ubc9 may have roles in poly-sumoylation (Capili and Lima, 2007,1).

The covalent binding, which is also referred to as specific binding occurs via the transfer of SUMO from the E1 enzyme (Aos1/Uba2) to Ubc9. In the binary complex, the C-terminal Gly97 of SUMO is attached to the Cys93 of Ubc9. In the complex, several Ubc9 residues are in contact distance with SUMO. Glu122 of Ubc9 makes a hydrogen bond with Arg63 of SUMO. Arg104 of Ubc9 is in contact with Gln92 of SUMO and Gln111 of Ubc9 is in contact with the backbone atoms of Gln29 of SUMO. Thr95 and Gly96 of SUMO make contacts with Ser95 of Ubc9. Furthermore, Arg85 of Ubc9 is in contact with C- terminal di-Gly (96Gly-97Gly) motif of SUMO (Reverter and Lima, 2005).

1.1.8.2. Interactions of RanGAP1. In sumoylation machinery target recognition is partially achieved by the E2 enzyme Ubc9. Ubc9 can recognize the Ψ-K-X-D/E consensus SUMO conjugation motif of the target. A binary complex of RanGAP1 and Ubc9 reveals that Ubc9 directs the target recognition (Bernier-Villamor *et al.*, 2002). The first residue of the consensus motif Ψ which is a hydrophobic residue corresponds to Leu523 on RanGAP1. Leu523 has van der Waals contacts with the residues Pro128, Ala129, Gln130 and Ala131 on Ubc9. The second residue of the motif, K is Lys524 on RanGAP1 which is the acceptor lysine of SUMO conjunction. The acceptor Lysine resides in a grove formed by Asp127, Pro128, Ala129 and Tyr87 residues of Ubc9. Besides, Asp127 and Lys524 are within a hydrogen bonding radius. Lys526 also resides within a hydrogen bonding distance to the active cysteine Cys93 of Ubc9, where it can attack the thioester bond formed between Gly97 of SUMO and Cys93 of Ubc9. The third residue of the consensus motif is not conserved and corresponds to the Ser525 of RanGAP1. The last residue of the consensus motif should be an Aspartic acid or a Glutamic acid residue; which corresponds to Glu526 on RanGAP1. This residue is

within the hydrogen bonding radius with Ser89, Thr91 and Lys74 of Ubc9. Besides, van der Waals interactions are present between Glu526 of RanGAP1 and Tyr87 of Ubc9 (Bernier-Villamor *et al.*, 2002). The interaction sites of RanGAP1 with Ubc9 complex is not limited to the sumoylation motif. Additionally, two helices of RanGAP1, Ser509-Met520 and Lys553-Thr564, interact with the Ala131-Asn140 region of Ubc9.

Residues Asn85, Tyr87 and Asp127 of Ubc9 are regarded as having catalytical roles but not structural roles via biochemical and mutational analuysis. Asp100 and Lys101 of Ubc9 are in such an orientation that provides interactions with a target in enough proximity. In mutational studies involving Asp100 and Lys101 of Ubc9, the SUMO binding to targets is reduced by 2.5 fold (Tatham *et al.*, 2003; Yunus and Lima, 2006). Moreover, the His-Pro-Asn (HPN) motif that corresponds to the residues 83-85 of Ubc9, and Glu-Pro-Asn (EPN) that corresponds to the residues 122-124 of Ubc9 are suggested to have structural significance. The HPN motif maintains hydrogen-bonding networks through orienting the SUMO C-terminal Gly-Gly motif. The EPN motif resides closer to the SUMO contact regions (Bernier-Villamor *et al.*, 2002; Wu *et al.*, 2003).

Mutational studies involving the residues Lys74, Asn85, Tyr87, Asp100, Asp127, Pro128, Ala129, Ala131, Glu132, Tyr134 and Thr135 is shown to reduce the sumoylation (Bernier-Villamor *et al.*, 2002; Yunus and Lima, 2006). Residues Glu132 and Tyr134 of Ubc9 serve as an additional binding surface for RanGAP1 which is in accordance with the fact that mutations of these two residues do not lower the sumoylation rate. (Bernier-Villamor *et al.*, 2002; Yunus and Lima, 2006).

1.1.8.3. Interactions of RanBP2. The C-terminal of IR1 domain of RanBP2 are in contact with the residues Arg8, Pro69, Pro105, Ala106 and the N-terminal of Ubc9. The Glu2671 and Asp2676 of RanBP2 are in contact with Arg13, Arg17 and Lys30 of Ubc9. The region of Phe2677-Leu2684 from RanBP2 packs on the β-sheet of Ubc9 which falls between Lys30-Asn40. The Leu2688, Tyr2689 and Leu2690 of RanBP2

make contacts with Phe22 and Asn40-Leu60 region of Ubc9. Gly47 and Gly55 of Ubc9 are within a hydrogen bonding distance with Tyr2689 of RanBP2 (Reverter and Lima, 2005).

It is shown that Arg17, Phe22-Asn31, Met36, Asn37, Asn40-Cys75, Glu78-Phe82, Gly115, Ile116, Phe155-Ala156 of Ubc9 are interacting with RanBP2 (Tatham *et al.*, 2005). Mutations of residues Phe22, Val25, Val27, Leu57 and Lys59 sevrely reduce RanBP2 binding efficiency, on the other hand mutations of residues Glu42, Lys48, Glu54 and Arg61 have comparatively moderate effect (Tatham *et al.*, 2005).

Considering the catalytic effects of RanBP2, Ile4, Val25 and Leu57 of Ubc9 are pointed out as RanBP2 interacting residues (Pichler *et al.*, 2004). The mutational studies involving the RanBP2 residues Leu2651, Leu2653 and Phe 2657 showed decreased RanBP2 function. Mutations of Pro2654, Cys2659, Asp2676 and Phe2677 are also known to effect RanBP2 function but less effectively (Pichler *et al.*, 2004).

1.2. Objective of the Study and Contribution to the Literature

The objective of this study is to verify the role of RanBP2 as an E3 for the sumoylation machinery and more importantly shed light on the controversial mechanism of sumoylation. Molecular Dynamics (MD) simulations Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 complexes are processed at a tempeature of 300 K. Model structure include the isopeptide bond between Gly97 of SUMO and 524Lys of RanGAP1 in order to shed light on the mechanism after covalent attachment of SUMO to target RanGAP1. To achieve consistency in results, parallel simulations are desiged for both structures.

Moreover, the presence of RanBP2 is pointed out to provide a packing motion at the Loop 2 region (Karaca *et al.*, 2010). Loop 2 region is Lys30-Met36 region of Ubc9 where some of the RanBP2 binding sites are located (Reverter and Lima, 2005). On that loop, Asp33 residue is observed to maintain a unique conformational state in which Asp33 makes a bending motion towards RanGAP1. In order to investigate

the role of Asp33 of Ubc9, Asp33Ala mutant is modeled. With the mutation, the significance of Asp33 along with the significance of Loop 2 region is underlined.

2. MATERIALS AND METHODS

2.1. Simulated Structures

Computations are based on the crystal structure of the hetero-tetramer complex with the pdb id 1Z5S (Lima and Reverter, 2005) for all models. Several structures produced from this complex structure are being simulated long enough to reveal the underlying mechanism of sumoylation. Those structures and simulation lengths are summarized on in Table 2.1 below.

Table 2.1. Simulated Models and Simulation Lengths

Simulated Models	Number of runs	Simulation Length (ns)
Ubc9-SUMO-E3-Target	2	59 & 49
Ubc9-SUMO-Target	2	48.5 & 51.5
ASP33ALA mutant of Ubc9-SUMO-E3-Target	1	43.5

2.1.1. Modeling of cases with the Peptide Linkage

The peptide bond is present between Gly97 of SUMO and Lys524 of Target, as in the crystal structure. The complex structure with the peptide bond represents the conformation before SUMO-Target binary complex is formed.

2.1.2. The Mutant Case

The mutation is designed in reference to the previous simulations on the SUMOylation cascade (Karaca *et al.*, 2010; Tozluoglu *et al.*, 2010) where Asp33 of Ubc9 display significant correlated fluctuations. The correlations increased considerably in the simulation of thioester bonded hetero-tetramer complex compared to the Ubc9-SUMO-Target case which lacks of the E3 ligase and has the thioester bonding between SUMO and Ubc9 (Karaca *et al.*, 2010). This correlation is coupled with the motion of the

flexible loop of Ubc9 between Thr29 and Asn37. Since this loop region of Ubc9 has correlations with E3 for all cases (Ubc9-SUMO-E3 and Ubc9-SUMO-Target-E3 with the thioester between SUMO-Ubc9 or with the peptide bond between SUMO-Target), this loop region is thought be functionally important and to prove that a mutant case is suggested.

To decide which amino acid to insert instead of the Aspartic acid at 33 of Ubc9, a sequence alignment among homologous structures is carried out (Larkin *et al.*, 2007). After sequence alignment, it has been observed that instead of Aspartic acid, Aspargine can be replaced. Since both Aspargine and Aspartic acid are polar mutant residue should be nonpolar. The most basic nonpolar amino acid is Glycine but since a Glycine mutation would provide flexibility, it is not preferred in this study. Instead of a Glycine mutation, an Alanine, which has a methyl as the side chain, mutation is thought to be more suitable for this case.

2.2. Molecular Dynamics Simulations

2.2.1. Theoretical Background

After proteins are transcripted, they undergo some conformational changes to be biologically functional. There are several ways to search for plausible conformations by both experimental and computational methods. The NMR (Nuclear Magnetic Resonance) relaxation techniques and X-ray crystallography, and recently Cryo-EM (Cryogenic Electron Microscopy) provide information about different conformational states of a given structure, but yet far from the description of a full spectrum of the available conformational states and the transitions between them. To this end, the Molecular Dynamics (MD) Simulation provides a means to explore the conformational space and dynamics involved with yet still some limitations.

In MD simulations, one uses numerical integration techniques to obtain approxi-

mate solutions to Newton's equations of motion (See Equation 2.1).

$$F_i = \mathrm{m_i} a_i \tag{2.1}$$

The output is a trajectory that specifies the positions of the atoms with time. (See Equation 2.2)

$$\frac{d^2 \mathrm{x_i}}{d \mathrm{t}^2} = m_i \frac{F_{x_i}}{m_i} \tag{2.2}$$

This equation explains the motion of the ith particle with mass mi along the coordinate x_i with F_{x_i} being the force on the i^{th} particle in x_i^{th} direction. (Leach, 2001)

The basis of MD depends on the potential function, V, of the system which is a function of the atomic positions. By using the potential function, the force exerted on each atom can be found. There are already empirically found parameters for both bonded and non-bonded interactions that are used in building of the force field. With those parameters, the potential function can be written as follows;

$$\begin{aligned}(\mathrm{V}(\mathrm{R}_1, ..., \mathrm{R}_N)) = & \sum_{bonds} \frac{k_i}{2}(l_i - l_{i,0})^2 + \\ & \sum_{angles} \frac{k_i}{2}(\theta_i - \theta_{i,0})^2 + \\ & \sum_{torsions} \frac{V_n}{2}(1 + cos(nw - \gamma)) + \\ & \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right)\end{aligned} \tag{2.3}$$

First term stands for primary interactions of the bonded atoms with l_i being the bond length, second term stands for the potential generated by the angle, θ_i of three successive atoms, the third term is to include the torsional potential. Those three terms

makes up the potential for bonded interactions. The last term is to incorporate the potential of nonbonded interactions, the first part of the 4^{th} term stands for Lennard-Jones 12-6 potential and the second part is counts for the electrostatic potential shown by Column potential. (Leach 2001)

2.2.2. Simulation Parameters for The System

The molecular dynamics simulation package AMBER is used for the simulations (Case *et al.*, 2004; Case *et al.*, 2005). The force field ff03 is chosen for the simulations. The protein is solvated explicitly in a truncated octahedron box using the TIP3P water box model (Jorgensen *et al.*, 1983). For start up, the initial velocities of atoms are initiated at 10 K by a Maxwellian distribution and then the temperature is gradually raised to 300 K. Once the temperature of the system has reached 300 K, 300 K is maintained. The pressure is kept at 1 bar by the Berendsen weak-coupling approach (Berendsen *et al.*, 1984). 2 fs was chosen as the time step along with the Leapfrog algorithm for the integration of Newton's equation of motion. At every 1 ps, the cartesian coordinates of atoms and energies are recorded.

2.2.3. Energy Minimization of The System

Stable state of molecular systems corresponds to global and local minimum on their potential energy surface. Starting from a non-equilibrium molecular geometry, energy minimization employs the mathematical procedure of optimization to move atoms so to reduce the net forces on the atoms until they become negligible. In order to obtain an energetically minimized structure as a starting conformation for the MD simulations, minimization algorithms are used. Energy minimization before launching MD simulations is useful as it removes the steric overlaps and relaxes the system.

2.3. Calculation of Root Mean Square Deviation, RMSD

The Root Mean Square Deviation (RMSD) is the measure of the average distance between the atoms of the superimposed conformations. In our case, the min-

imized structure is considered as the reference structure and all other snapshots of conformations created by MD simulation are superimposed onto that structure and the RMSD here reflect the deviation from the minimized X-ray crystal structure. The ptraj module of AMBER 8.0 (Case *et al.*, 2004; Case *et al.*, 2005) is used to calculate the Cα-RMSD values.

2.4. Calculation of Mean Square Fluctuations (MSF) and Correlation Between The Fluctutations

Mean Square Flucations (MSF) is a measure of the deviation Cα's of each residue from the mean structure, which is taken as the average structure calculated over the MD trajectory for the period where the system is observed to be under dynamic equilibrium. For all systems studied the equilibration periods are not included in the calculations of fluctuations. The equilibration point is determined via the RMSD plots. (The equilibration times are given in table)

Correlations between fluctuations, often referred as cross correlations, reveal the correlated or uncorrelated parts of the structure. Correlations of the ith residue with the jth residue are calculated as

$$C_{i,j} = \frac{\langle \Delta R_i \Delta R_j \rangle}{\langle \Delta R_i^2 \rangle^{1/2} \langle \Delta R_j^2 \rangle^{1/2}} \tag{2.4}$$

In Equation 2.4, ΔR_i is the fluctuation of the position vector R_i ; ΔR_i is the fluctuation of the position vector R_j. The position vectors are attained from the molecular dynamic trajectory after aligning to the minimized structure. In this formula, i and j stands for the residue indexes of interest.

For Equation 2.4, the brackets imply that the values are calculated by averaging over the ensemble of conformations starting from the equilibrated structure to the end of the trajectory. Cross correlations are calculated over a specified range of the trajectory with reference to the average structure of the specified range. The cross

correlation values vary over the range of -1:1 where -1 implies full anti-correlation, 1 implies full correlation and 0 implies no correlation of the fluctuations.

2.5. Calculation of The Time Delayed Autocorrelations of The Virtual Bond Vectors

The time delayed autocorrelations of the virtual bond vectors are calculated from the trajectory over the simulation time. The following formula is used for calculating correlations

$$\frac{1}{T(N-1)} \sum_{t=0}^{T} \sum_{i=0}^{N-1} \{r_{i,t+\tau} \times r_{i,t}\}^2 \tag{2.5}$$

where N is the residue number, T is the length of simulation, τ is the time delay and r is the unit orientational vector between two α-carbons of consecutive residues. This reveals the degree of maintaining a correlated motion.

2.6. Clustering

Large number of conformations is created via molecular dynamics simulations. To reduce the conformational space and identify the major conformational states, the conformations generated along the simulation time after the equilibration period are clustered by k-means clustering using the *kclust* module of Multiscale Modeling Tools for Structural Biology (MMTSB) Tool Set (Feig *et al.*, 2004) having RMSD of the alpha carbons as the similarity measure.

In k-means clustering, at first a collection of frames are selected randomly and this frames are assigned to centroids. Within the specified RMSD threshold, the snapshots are assigned to clusters based on their RMSD value with the centroids of the clusters. This procedure is repeated until every single frame is assigned to a cluster. The centroids and the clusters are updated after each iteration step.

In this study, the clustering is performed over the conformations aligned to the minimized structure and than average structure which is calculated over the conformations that reflects the equilibrium dynamics aligned to the minimized structure. The RMSD threshold is chosen such that distinct clusters are formed.

3. RESULTS AND DISCUSSIONS

Results and Discussions will be presented in 2 sections. In Section 3.1., by comparing two runs of Ubc9-SUMO-RanGAP1-RanBP2 and two runs of Ubc9-SUMO-RanGAP1 complexes and the role of RanBP2 on the Ubc9-SUMO-RanGAP1 complex with the Isopeptide Bond will be summarized. In the Section 3.2., the effect of Asp33Ala mutation will be discussed by comparing the simulation of Asp33Ala mutant structure of Ubc9-SUMO-RanGAP1-RanBP2 complex with two wild type simulations of the same complex.

3.1. The Role of RanBP2 on Ubc9-SUMO-RanGAP1 Complex with the Isopeptide Bond at 300 K

In this section, the simulations of the two plausible complex structures of sumoylation machinery adapted from the X-ray crystal structure, 1Z5S, (Reverter and Lima, 2005) are analyzed comparatively. These structures are Ubc9-SUMO-RanGAP1 and Ubc9-SUMO-RanGAP1-RanBP2, which are simulated at 300 K with the isopeptide bond between SUMO and RanGAP1. Parallel runs are performed for both complex structures and the results of the analysis are presented in detail in Appendixes A and B.

3.1.1. Root Mean Square Deviations (RMSD)

The Root Mean Square Deviations (RMSD) of trajectories are calculated by aligning the trajectories to the C-α atoms of the initial structure, which is the structure obtained after the energy minimization of the crystal structure. The whole 59000 snapshots, covering a time window of 59 ns, are considered in the averaging.

The RMSD plots are not descriptive for specific motions but they reflect the extent of motion from the initial conformation and the conformational changes throughout the trajectories. These plots are also used to identify the initial equilibration

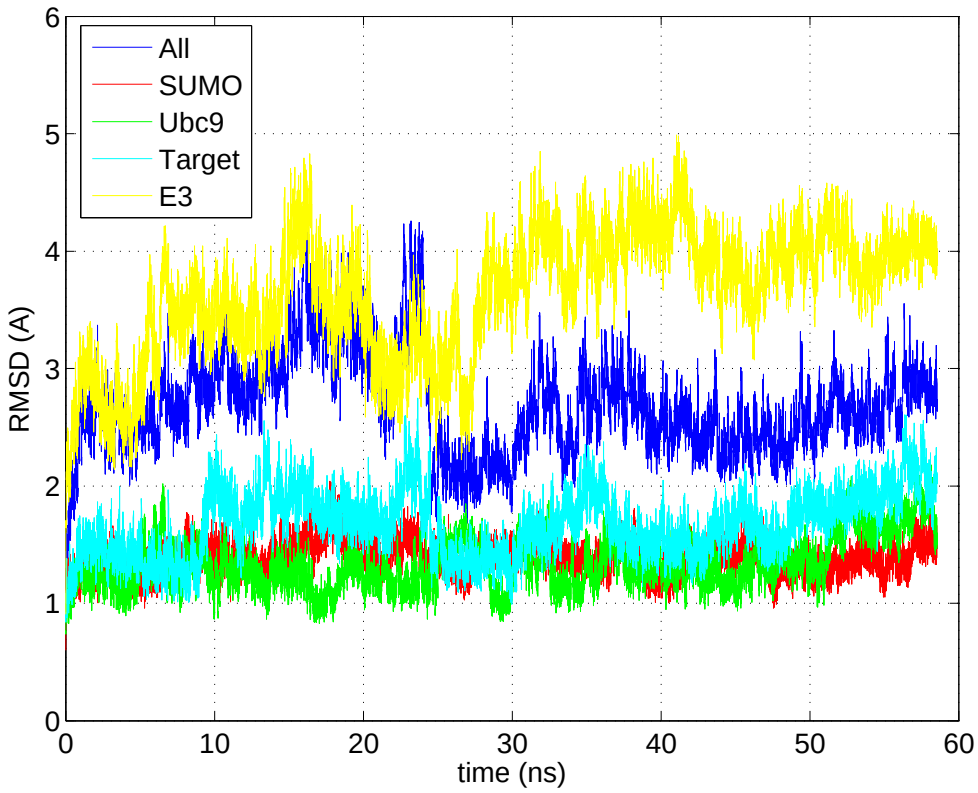

Figure 3.1. RMSD of Ubc9-SUMO-RanGAP1-RanBP2 tetramer complex from the initial minimized structure

periods prior to the dynamic equilibrium for the time window presented and analyzed here.

Figure 3.1 displays the RMSD values of Ubc9-SUMO-RanGAP1-RanBP2 complex structure and its individual chains; Ubc9, SUMO, RanGAP1, and RanBP2. The RMSD of Ubc9, SUMO and RanGAP1 fluctuate around the mean values 1.33, 1.38 and 1.67 Å, respectively, throughout the trajectory. Yet, the RMSD of RanBP2 displays high amplitude variations up to 5.0 Å. The latter is due to the large conformational changes of the RanBP2's loops, existing as a fragment in the crystal structure. The results are similar in the simulation carried out in parallel (Figure A.1). The RMSD values for Ubc9, SUMO and RanGAP1 fluctuate around the mean values 1.21, 1.32 and 1.42 Å, respectively, throughout the trajectory. On the other hand, the RMSD of RanBP2 displays high amplitude variations around 4.00 Å and that is reflected on the overall RMSD values. Merely looking at the profiles, it can be concluded that overall RMSD is mostly affected by the RMSD of RanBP2, besides the mean values are higher

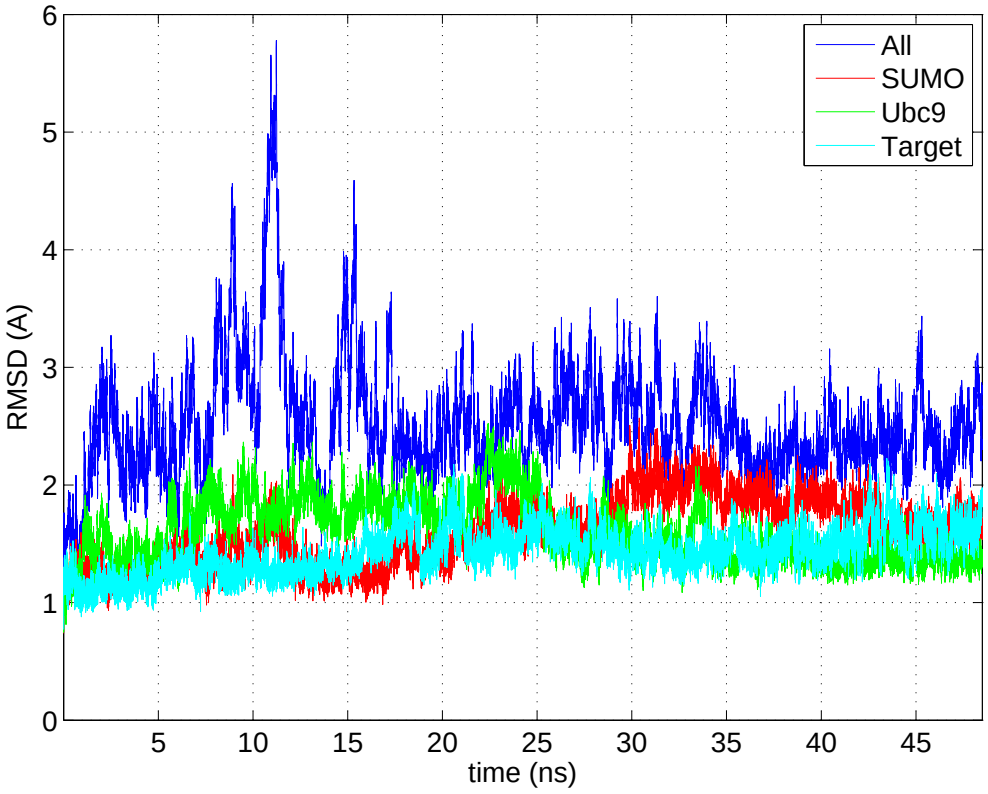

Figure 3.2. RMSD of Ubc9-SUMO-RanGAP1 ternary complex from the initial minimized structure

than the RMSD values of individual chains of Ubc9, SUMO and RanGAP1. The mean values for the overall RMSD are 2.69 Å and 2.60 Å (See Figure A.1) for the two runs which are obviously amplified by RanBP2 behavior.

Figure 3.2 displays the RMSD values of the Ubc9-SUMO-RanGAP1 complex structure and its individual chains, Ubc9, SUMO, and RanGAP1. The RMSD values for Ubc9, SUMO and RanGAP1 fluctuates around 2.00 Å. Nevertheless, the overall RMSD values for the Ubc9-SUMO-RanGAP1 complex structure are observed to display higher amplitude fluctuations that reaches up to 5.64 Å. There exists a peak around 11 ns, which should reflect the positional changes of the chains as a rigid body in the complex structure, where the individual chains display relatively low RMSD values.

Figure 3.3 displays the conformation at 11 ns and the initial minimized X-ray crystal structure aligned to the C-α atoms of overall structure. Overall alignment is

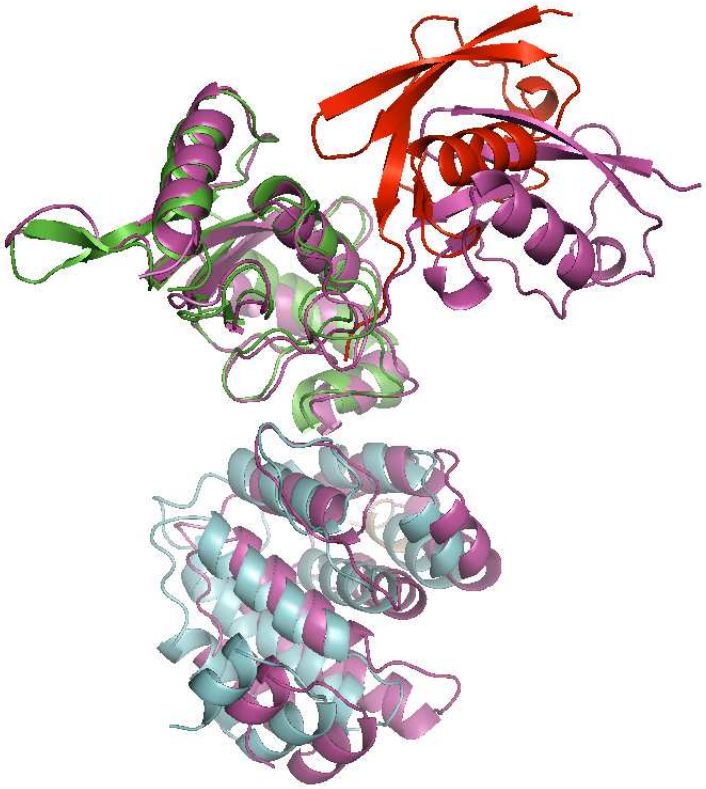

Figure 3.3. Alignment of Ubc9-SUMO-RanGAP1 mimized structure (Green: Ubc9, Red: SUMO, Blue: RanGAP1) onto the structure at 11 ns (Magenta) of the Ubc9-SUMO-RanGAP1 simulation

selected here to better observe the movement of chains with respect to each other. As one can see, the chains move apart from each other, being more dramatic for SUMO. SUMO moves away from its original position and approaches to RanGAP1.

Along, there are some conformational changes at Ubc9 and RanGAP1 interface and other parts of these two chains. The RMSD value of the alignment shown in Figure 3.3 is 5.64 Å, for which the rigid body motion of SUMO is mainly responsible. This might be an implication of SUMO-RanGAP1 dissociation from the Ubc9. However, the sampling of this conformation is very low throughout the simulation. SUMO goes back to orientation in its minimized structure in about 2 ns. Throughout the trajectory, this kind of rigid body motion of SUMO is not encountered again. In the parallel run (Figure B.1), the behavior of individual chains are similar to the results presented here, yet, with the lower overall RMDS values and less pronounced RMSD jumps. The peaks reach 4.0 Å the highest. The peaks around 12 ns and 50.5 ns are examples of

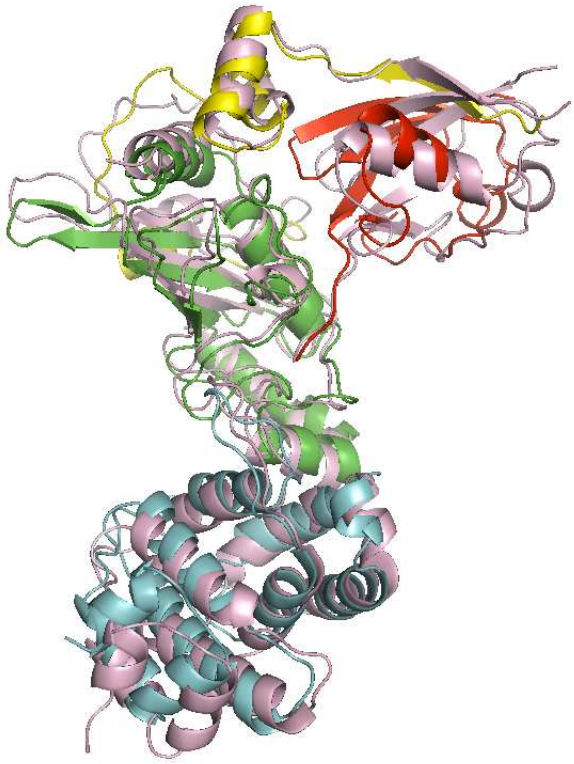

Figure 3.4. Alignment of Ubc9-SUMO-RanGAP1-RanBP2 mimized structure (Green: Ubc9, Red: SUMO, Blue: RanGAP1, Yellow: RanBP2) with the structure at 23.64 ns (Light Pink) of Ubc9-SUMO-RanGAP1-RanBP2 simulation which gives 2.75 Å for the individual RMSD of RanGAP1

the RMSD jumps.

The RMSD values of SUMO and Ubc9 chains do not differ significantly with and without RanBP2, yet fluctuate around 1.5 Å, respectively (Figure 3.1 versus Figure 3.2 and Figure A.1 vs Figure B.1). The mean values remain close whether RanBP2 is present or not, however, the amplitudes of fluctuations of Ubc9 and SUMO are lowered in the presence of RanBP2, that is to say RanBP2 stabilizes Ubc9 and SUMO. RanGAP1, on the other hand, has increased fluctuations in the presence of RanBP2. Without RanBP2 the RMSD of the RanGAP1 fluctuates about 1.40 Å but in the presence of RanBP2, the RMSD of RanGAP1 reaches up to 2.75 Å. In Figure 3.4, a snapshot having high RMSD of RanGAP1 is taken and aligned on the minimized Ubc9-SUMO-RanGAP1-RanBP2 structure. In Figure 3.4, the conformational change of RanGAP1 from Ubc9-SUMO-RanGAP1-RanBP2 simulation is obvious, the C-terminal

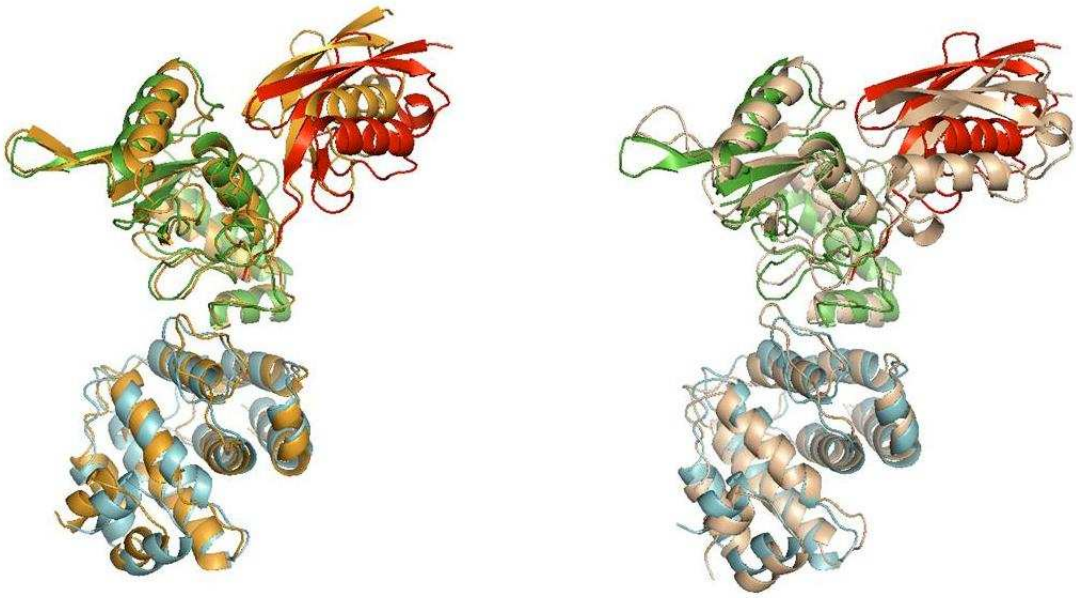

Figure 3.5. Alignment of Ubc9-SUMO-RanGAP1 mimized structure (Green: Ubc9, Red: SUMO, Blue: RanGAP1) with (a) the structure at 12 ns (Orange, on the left) and (b) the structure at 52.5 ns (Light Pink, on the right) taken from the parallel run of Ubc9-SUMO-RanGAP1 simulation

region half of the structure bends itself towards the SUMO causing the high RMSD values. However, the dramatic changes, such as the RMSD jump at 11 ns, of RMSD observed in Ubc9-SUMO-RanGAP1 are not observed in RanBP2 presence. RanBP2 seems to affect RanGAP1 such that RanGAP1, especially the C-terminal region, bends itself towards SUMO.

Figure 3.5 represents alignments of the snapshots at 12 ns and 52.5 ns with the minimized structure from the parallel run of Ubc9-SUMO-RanGAP1, respectively. Alignment is carried out with respect to the C-α atoms of the overall Ubc9-SUMO-RanGAP1 structure. In the (a) module, SUMO is observed to move backwards; however, in module (b) the SUMO bends itself towards RanGAP1. In both of the snapshots, Ubc9 and RanGAP1 make conformational changes, specifically the Loop 2 region of Ubc9 fluctuates. The movement of the loop is clearer in (b) module. Although, there are conformational changes at Ubc9 and RanGAP1, the RMSD jump is mainly caused by the rigid body motion of SUMO. It can be concluded that without RanBP2, the

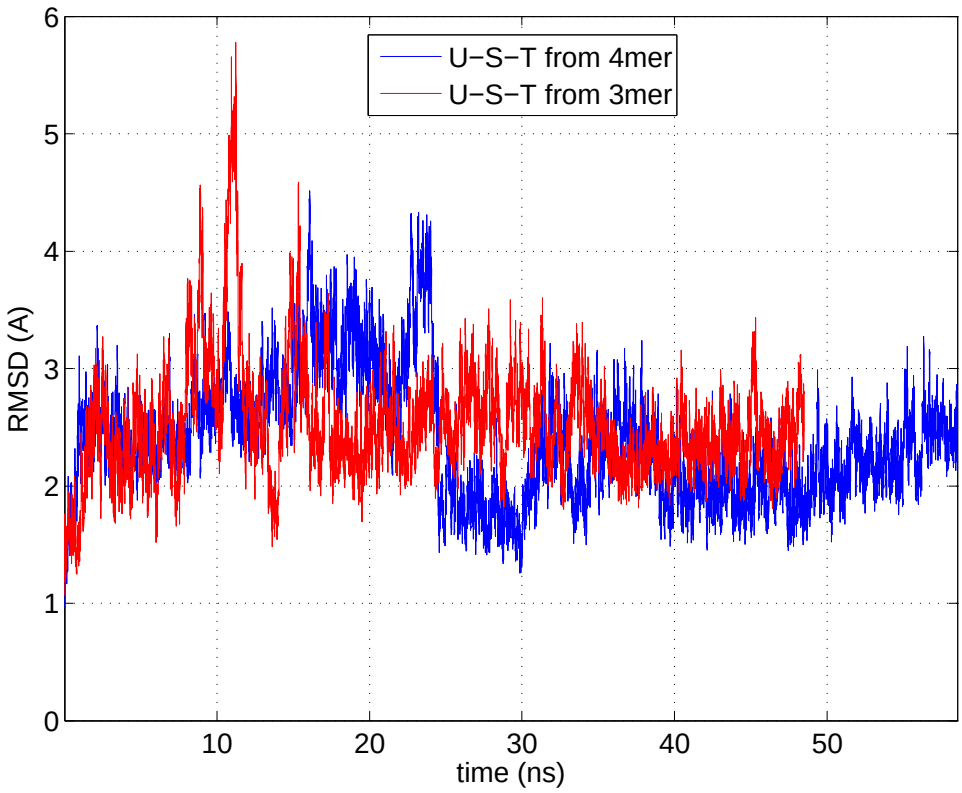

Figure 3.6. RMSD of Ubc9-SUMO-RanGAP1 from the simulations of Ubc9-SUMO-RanGAP1-RanBP2 (blue) structure and Ubc9-SUMO-RanGAP1 (red) structure from the initial minimized structure

position of SUMO is not stable.

The comparison of the overall RMSD values with and without RanBP2 might be misleading due to the large conformational changes of the RanBP2 fragment. For this, the RMSD values of the Ubc9-SUMO-RanGAP1-RanBP2 complex structure are calculated excluding the RanBP2 chain and compared with the RMSD values of the Ubc9-SUMO-RanGAP1 complex structure from the Ubc9-SUMO-RanGAP1 simulations (Figure 3.6). With RanBP2, high jumps in RMSD (like the jump at 11 ns observed in RMSD of Ubc9-SUMO-RanGAP1-RanBP2) are not observed and thus the amplitude of the average fluctuations of the RMSD values of the Ubc9-SUMO-RanGAP1 simulations is lower than those from the Ubc9-SUMO-RanGAP1-RanBP2 simulation. Apparently, RanBP2 stabilizes the conformational transitions by reducing the conformational space of the structure as an E3 ligase for the sumoylation system. The average of overall complex and individual chain RMSD values for each different

simulation are summarized by Table 3.1.

Table 3.1. Average RMSD values (in Å) of overall structures and individual chains

Simulation Type	All	Ubc9	SUMO	RanGAP1	RanBP2
Ubc9-SUMO-RanGAP1-RanBP2	2.69	1.33	1.38	1.67	3.64
Ubc9-SUMO-RanGAP1-RanBP2, second run	2.60	1.21	1.32	1.42	3.97
Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2	2.36	1.13	1.34	1.50	3.30
Ubc9-SUMO-RanGAP1	2.52	1.60	1.60	1.43	
Ubc9-SUMO-RanGAP1, second run	2.17	1.37	1.60	1.37	

Figures 3.1 and 3.2 suggest the equilibration periods as 10 ns and 3 ns, respectively, for the Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 complex structures. Furthermore, under the light of Figure A.1 and Figure B.1, the equilibration periods for the parallel runs are determined as 5 ns for each simulation.

3.1.2. Mean Square Fluctuations (MSF)

The mean square fluctuations (MSF) are calculated from the simulations after the equilibration period. The calculations are carried out over the conformations aligned to the C-α atoms of the minimized structures. The profile of the MSF suggests flexible and rigid parts of the structures.

Figure 3.7 displays the experimental temperature factors (B-factors) of the Ubc9-SUMO-RanGAP1-RanBP2 complex structure (Reverter and Lima, 2005) and the MSF of the Ubc9-SUMO-RanGAP1 and Ubc9-SUMO-RanGAP1-RanBP2 complex structures from the simulations. In the MSF presented here, the individual C-α alignments are carried out for each chain and results are merged MSF for the whole structure. As the mean-square fluctuations are evaluated on the individual chain basis, the behavior of fluctuations discussed here should mainly reflect the intramolecular interactions and/or implied intermolecular interactions. The calculated MSF results do not match exactly to the experimentally determined B-factors in magnitudes. B-factor is the $8\pi/3$ factor of MSF, yet the profile of the fluctuations is similar.

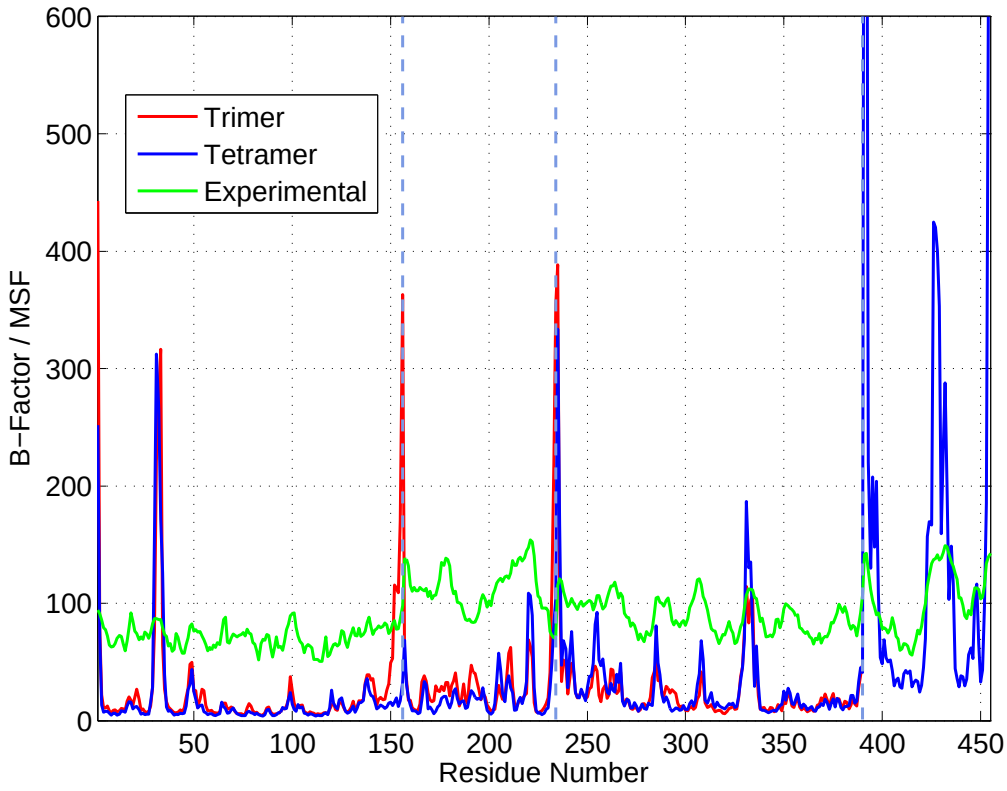

Figure 3.7. Comparison of the experimental B-factors and the MSF calculated from the trajectories through individual alignment for each chain (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1; 391-455: RanBP2)

Among the three chains Ubc9, SUMO and RanGAP1; SUMO and the RanBP2 region shows higher fluctuations and Asp33 of Ubc9 (residue 32) which resides on the Loop 2 region is said to be functionally important (Karaca *et al.*, 2010) and shows the largest fluctuations in Figure 3.7, except the chains ends which can move freely. The highest peak following Asp33 is Gly97 of the SUMO which makes a peptide bond Lys524 of RanGAP1.

Asp33 of Ubc9 displays the highest mobility in Ubc9 in both Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 complex simulations. This residue is in the middle of a mobile and flexible loop, Loop 2, which has recently been suggested as functional (Karaca *et al.*, 2010; Tozluoglu *et al.*, 2010). This region is in fact the RanBP2 binding site of Ubc9 (Table C.1). Another peak is at Gly50 of Ubc9 located in the middle of a mobile loop whose interaction with Gly68 of SUMO is already known, which is found to be anchoring towards Ubc9 in the absence of RanBP2 when

Ubc9 and SUMO are linked via a thioester linkage. (Karaca *et al.*, 2010). Asp100 and Lys101 (residue 99 and 100) take roles in RanGAP1 recognition (Tatham *et al.*, 2003; Yunus and Lima, 2006) and these residues are also observed to display relatively high mobility. The catalytic regions Lys74-Phe77, Cys93 and Ala131-Gln139, which acts as the built-in E3 of Ubc9, regions are not mobile according to the MSF results. These regions are already stabilized with SUMO binding (Karaca *et al.*, 2010) and the RanBP2 binding do not lead to a more enhanced stability.

In SUMO (region: 157:234) presence of RanBP2 only makes slight changes in MSF. The most significant fluctuation is observed for Gly68 (residue 205), which resides in the middle of a loop that interacts with Gly50 of Ubc9.

Pro451, Asp482, Ser505 and Lys528 (residues 254, 285, 308 and 331) are the residues of the RanGAP1 that display the highest mobility. Those residues are not functional in sumoylation machinery; however, their mobility decreases with RanBP2. Lys528 is important as it affects the orientation of the LKSE SUMO binding motif between residues 523-526 of RanGAP1. This implies that RanBP2 not only puts the SUMO and Ubc9 in right orientation, but also stabilizes the fluctuations of those regions of RanGAP1.

The MSF results for the parallel runs which are presented in Appendix A and B support the same outcomes. In Figure A2 and B2 experimental B-factors along with MSF for different runs are presented. Similar patterns are observed for both Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 simulations respectively.

Figure 3.8 displays the MSF of the Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 complex structures calculated with the overall complex structure alignment. Here, the configurational as well as conformational behavior is possibly assessed. As seen, SUMO (residues 157:234) displays the highest mobility with and without RanBP2, being more enhanced for the latter. This also confirms the RMSD profiles that SUMO moves itself towards RanGAP1 in Ubc9-SUMO-RanGAP1 simulations. The presence of RanBP2 restricts the SUMO's position in space with respect

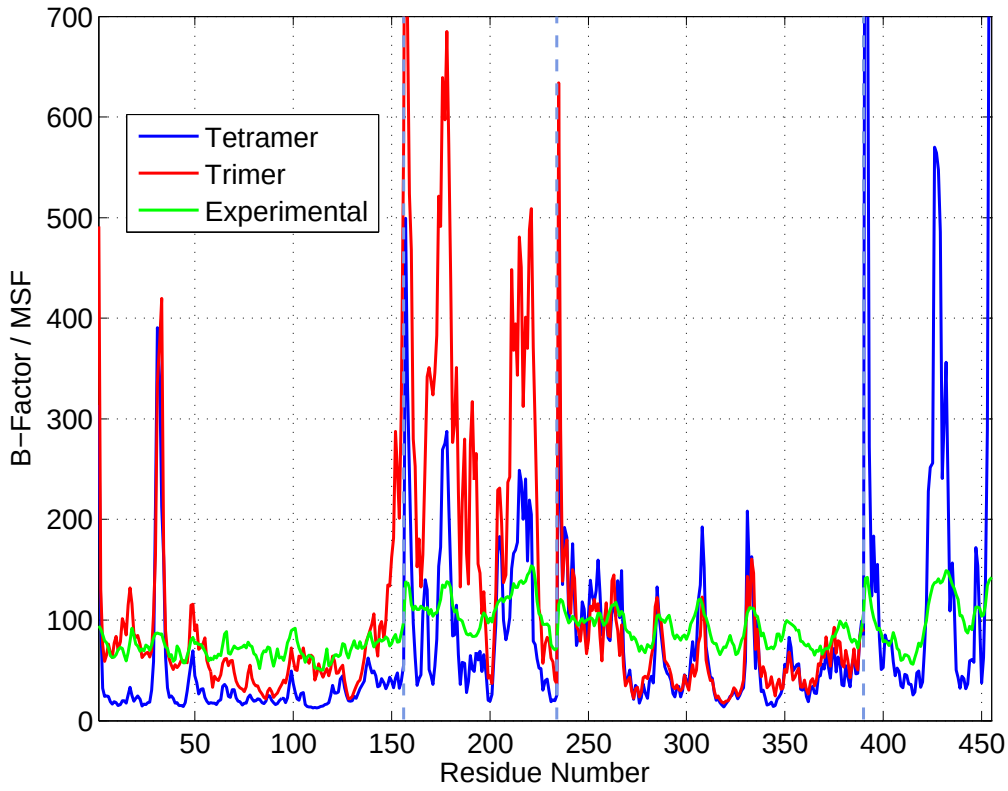

Figure 3.8. Comparison of the experimental B-factors and calculated from trajectories through overall alignment on the whole structure (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1; 391-455: RanBP2)

to the other chains.

Among the three chains Ubc9, SUMO and RanGAP1, SUMO and the RanBP2 region shows higher fluctuations and Asp33 of Ubc9 (residue 32) shows the largest fluctuations, except the chains ends which can move freely. The highest peak following Asp33 is Gly97 of the SUMO which protrudes its side chain towards Lys524 of RanGAP1.

RanBP2 holds the Ubc9 and SUMO in the right positions and eases the RanGAP1 binding (Reverter and Lima, 2005). Ubc9 displays quite restricted fluctuations, except Asp33. With RanBP2, Ubc9 gets even more restricted in its residues' fluctuations. On the other hand, contrary to the Ubc9 and SUMO, RanGAP1's fluctuations display similar behavior; however, the MSF of RanGAP1 in the presence of RanBP2 is about 10% higher.

In Figure 3.7 and Figure 3.8, MSF are calculated by the same principles, except that individual chain alignment and overall (complex) structure alignment are done respectively. Individual alignments mainly characterize the internal mobility profile within the chain with yet some implicit contributions from the other chains. Overall alignments, nevertheless, provides clues about the possible rigid body motion of the chains within the quaternary structure. Comparing the two figures, it is demonstrated that SUMO make a configurational change by moving towards RanGAP1 without RanBP2, which is as well visualized in Figure 3.5.

The control runs for both Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 complex structures yield a very similar behavior for the residues and chains in these complex structures (Appendix A and B.).

3.1.3. Correlation between Fluctuations

The correlations between the fluctuations of residues, cross correlations, are calculated by the analysis of the conformations aligned to the energy minimized initial structure. The correlations are normalized and the color code is from red to negative describing, respectively, the most positively to the most negatively correlated fluctuations.

Figure 3.9 displays the cross correlation map for the Ubc9-SUMO-RanGAP1-RanBP2 complex structure. Correlations between fluctuations, often referred as cross correlations, reveal the correlated or uncorrelated parts of the structure. It reflects the correlations of the deviations from C-α's of each residue from the mean structure. The positively correlated residues make the dynamic blocks in the complex structure, where some of the residues mediate the association between the chains in a complex structure. Here, it is possible to identify the chains Ubc9, SUMO, RanGAP1 and RanBP2 along the diagonal, where more residues with positively correlated fluctuations are populated. The coupling between the fluctuations of these regions are more negatively imposed, yet there are regions between that are observed to display positively correlated fluctuations to maintain the collectivity of the chains within the complex.

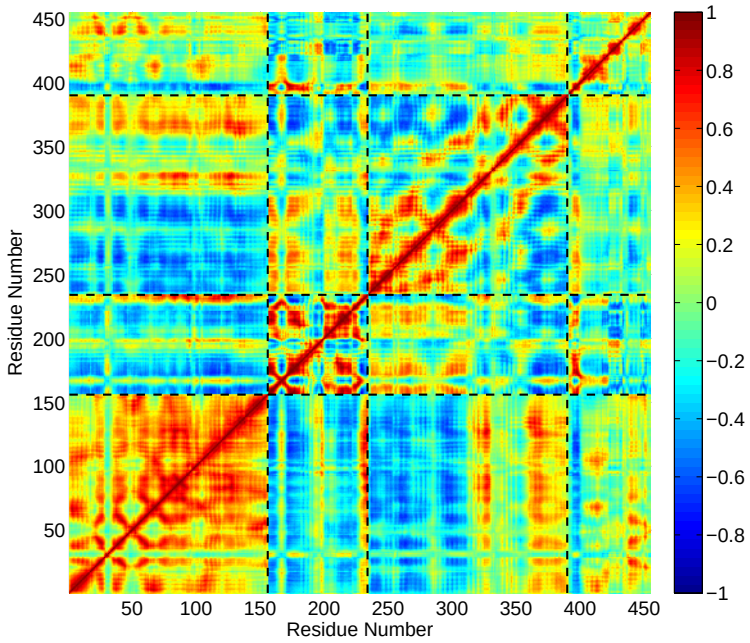

Figure 3.9. Cross-Correlations of the simulation of Ubc9-SUMO-RanGAP1-RanBP2 tetramer complex calculated by alignment onto the initial minimized structure (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1; 391-455: RanBP2)

In the cooperative fluctuations of the complex structure, it is observed that Asp33 (residue 32), which in the middle of Loop 2 of Ubc9 makes positive correlations with Thr41-Leu44 and His75-Glu85 regions of SUMO (residues 178-181 and 212-222), which are close to the RanBP2 interface yet far from the Ubc9 interfaces. On the other hand, the two sides of Loop 2 display positively correlated fluctuations with Thr511-Leu522, the LKSE motif between Leu523-Glu526 and Leu555-Pro566 (residues 314-329 and 358-369) regions of RanGAP1. Thr511-Leu522 and Leu555-Pro566 are the helices that face the built-in E3 of Ubc9 which nothing but the Ala131-Gln139 region of Ubc9 (residues 130-138). The previous MD simulations (Karaca *et al.*, 2010) has already underlined the functional role of Asp33 in the SUMO binding cascade. Further, Glu50 (residues 49) of Ubc9 also displays a weak correlations with the C-terminal of RanBP2. Asp 100 and Ala 101 (99 and 100th residues) that are known to be important for target recognition (Tatham *et al.*, 2003; Yunus and Lima, 2006), are correlated with RanGAP1 region, particularly with the LKSE motif of RanGAP1 (residues 326-329). As seen, Ubc9 is correlated with most of the RanGAP1, thus, the Ubc9 dissociation is not yet feasible at

thist point with high coupling of Ubc9-Target in their fluctuations. Asp100 and Ala101 (residues 99 and 100) are also correlated with Ser61 of SUMO (residue 198) which is close to the C-terminal tail of SUMO and might be important in positioning the Gly97 (residue 234) of SUMO in the right orientation. Gly97 of SUMO of is important as sumoylation occurs through binding of this residue with the Target binding sites.

In addition to the weak correlations with Ubc9, SUMO shows strong correlations with RanGAP1, implying that SUMO-RanGAP1 coupling is achieved. Particularly, Tyr21-Leu24 and Lys37-His43, which is the N-terminal region of SUMO and resides at the RanBP2 interface, along with the Gly68-Val90, which is mostly consisted of loops and Gly50 (residue 49) of Ubc9 anchors to this large loop, of SUMO (residues 158-161, 174-180, 205-227) shows high correlations with the loop that contains Asp462 and the loop that starts with Lys500 of RanGAP1 (residues 265 and 300). Those loops are the flexible regions of RanGAP1 that move towards SUMO in the presence of RanBP2 and causes considerable increases RMSD (Figure 3.5)

The results of the parallel run can be found in Appendix A (Figure A.4), which highly agree those presented in the main text.

Figure 3.10 presents the cross correlations of the Ubc9-SUMO-RanGAP1 complex structure. To start with Ubc9 region (1:156), the coupling of Ubc9's residues in their fluctuations is stronger than those with RanBP2. Ubc9 display correlated fluctuations with Ser61 and Arg70 (residues 198 and 207) of SUMO located at the Ubc9 interface. Interestingly, the residues of SUMO that maintain the association with Ubc9 are not the same with and without RanBP2. This suggests that Ubc9 and SUMO interface is still coupled in their fluctuations, yet in the presence of RanBP2 those correlations are not as strong and different residues in Ubc9 and SUMO show association. This implies the emergence of implying a new, but weak, interface formation.

Compared to the Ubc9-SUMO-RanGAP1-RanBP2 complex structures, the correlation of Loop 2 where Asp33 is located in the middle with Thr511-Leu522, the LKSE motif between Leu523-Glu526 and Leu555-Pro566 (residues 314-329 and 358-369) re-

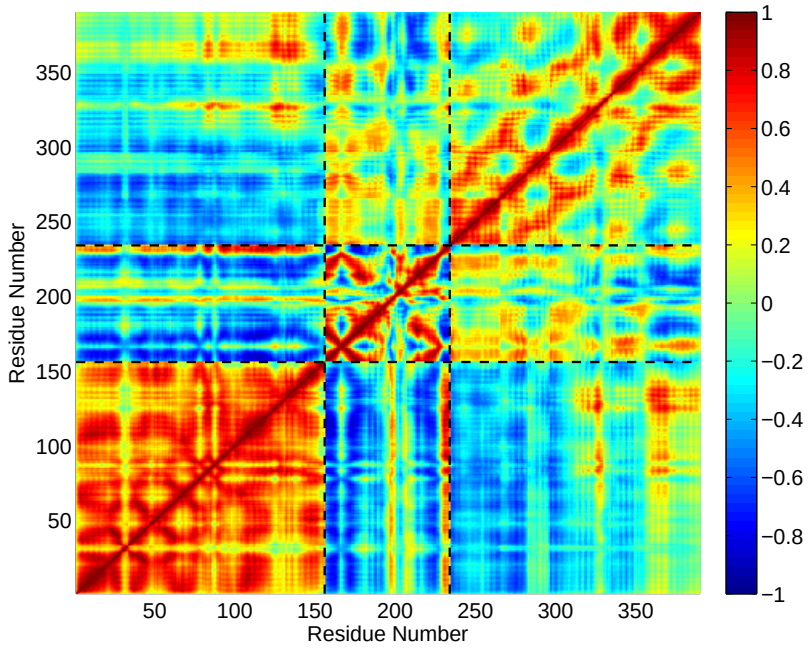

Figure 3.10. Cross Correlations of the simulation of Ubc9-SUMO-RanGAP1 tetramer complex calculated by alignment onto the initial minimized structure (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1)

gions of RanGAP1 has faded without RanBP2. The catalytic Cys97 and the surrounding loop between Lys74-Phe77 of Ubc9 (residues 96, 73-76) are less correlated with Thr511-Leu522, the LKSE motif between Leu523-Glu526 and Leu555-Pro566 (residues 314-329 and 358-369) regions of RanGAP1. The built-in E3 of Ubc9 which is Ala131-Gln139 (residues 130-138) region of Ubc9 is highly correlated with Thr511-Leu522, the LKSE motif between Leu523-Glu526 and Leu555-Pro566 (residues 314-325, 326-329 and 358-369) of RanGAP1. LKSE is the highly conserved SUMO binding motif whereas Thr511-Leu522 and Leu555-Pro566 regions are important as they face the Ala131-Gln139, the built-in E3 of Ubc9.

SUMO, on the other hand, maintains the same correlations with the RanGAP1 as in the Ubc9-SUMO-RanGAP1-E3 complex structure. Tyr21-Leu24 and Lys37-His43, which is the N-terminal region of SUMO and resides at the RanBP2 interface, along with the Gly68-Val90, which is mostly consisted of loops and anchored by Gly50 (residue 49) of Ubc9, of SUMO (residues 158-161, 174-180, 205-227) still has high

correlations with the loops that contains Asp462 and Lys500 (residues 265 and 300), which move towards SUMO in the presence of RanBP2 and increases in RMSD (Figure 3.5).

In the presence of RanBP2, Loop 2 which contains the Asp33 is coupled with Thr511-Leu522, the LKSE motif between Leu523-Glu526 and Leu555-Pro566 regions of RanGAP1. Increasing correlations of Loop 2 with the mentioned regions on RanGAP1 leads to increased correlations of SUMO with the RanGAP1. Without RanBP2, Ubc9 rearranges its cooperative fluctuations where new couplings are formed with Ser61 and Arg70 residues of SUMO and the correlations of Ubc9 with the RanGAP1 are weakened. Considering these, it can be suggested that the presence of RanBP2 exerts its effect by weakening the coupling Ubc9-Target and reorganizing the Ubc9-SUMO associations.

The results for the parallel run of Ubc9-SUMO-RanGAP1 complex can be found in Figure B.4 in Appendix B section.

3.1.4. Decay of Orientational Vectors by Time

Orientational correlations for the virtual bond vectors between two successive C-α atoms are calculated. The behavior of the virtual bond vectors can conveniently be analyzed throughout the simulations by the autocorrelation values as a function of time. The autocorrelation values at various time lags are presented for Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 structures in Figures 3.10 and 3.11, respectively. It is of interest here that which residues or which regions display the maximum loss or conservation in their autocorrelations with RanBP2.

Figure 3.11 displays the autocorrelation values for the bond vectors of the Ucb9-SUMO-RanGAP1-RanBP2 complex structure at various time delays. As the time delay increases the correlation values decrease considerably as one can expect, which means that the knowledge of the orientation value within the time window indicated is lost. Analysis of the minima revealed that Asp33, the bond vector between residues

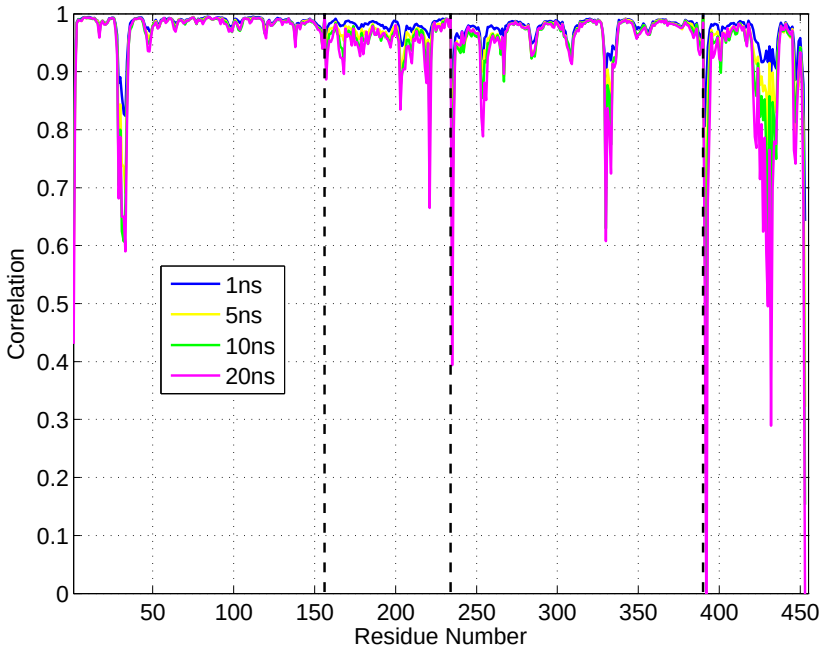

Figure 3.11. Decay of Orientational Vectors from simulation of Ubc9-SUMO-RanGAP1-RanBP2 complex calculated by alignment onto the initial minimized structure (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1; 391-455: RanBP2)

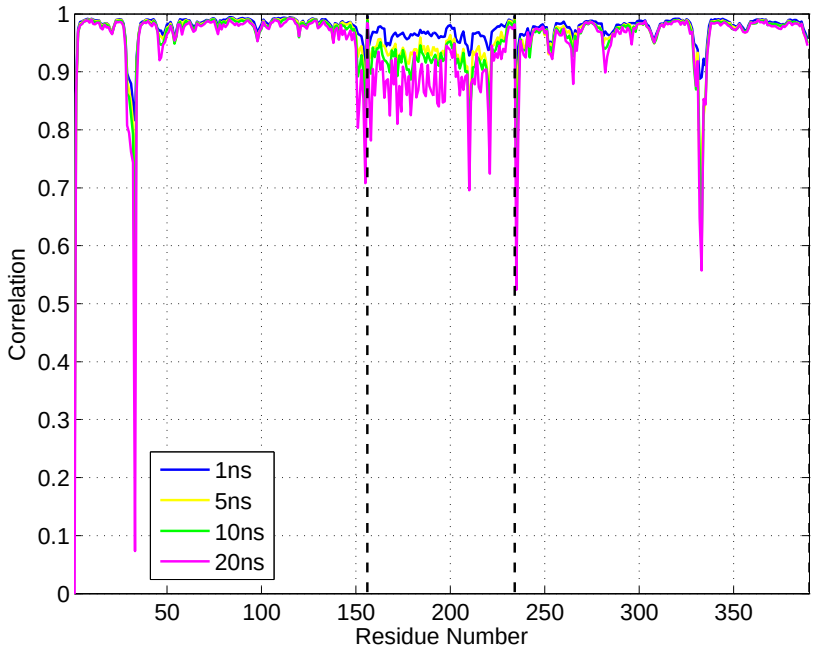

Figure 3.12. Decay of Orientational Vectors from simulation of Ubc9-SUMO-RanGAP1 complex calculated by alignment onto the initial minimized structure (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1)

32 and 33 could comparatively sample the orientational space more freely than the rest of residues in Ubc9 implying high degree of flexibility. The minima on SUMO, Ser31, Phe66 and Glu84 (residues 168, 203 and 221) corresponds to loop regions at the RanBP2 interface suggesting loose interactions at SUMO-RanBP2 interface. As for RanGAP1, the most pronounced minimum is observed for Lys528 (residue 330), which comes right after the LKSE (Leu523-Glu526) motif. This residue comes right after the conserved SUMO binding motif and has structural importance as the orientation of Lys528 affects the orientation of the LKSE motif. However, this residue can have functional role on positioning of the attacking of SUMO's tail to the Lys524. This will be further inspected.

Figure 3.12 displays the autocorrelation values for the bond vectors of the Ucb9-SUMO-RanGAP1 complex structure at various time delays. As the time delay increases the correlations values decrease faster than the Ubc9-SUMO-RanGAP1-RanBP2 complex structure. The minima, especially Asp33 (residue 32) of Ubc9 and SUMO region are similar with Ubc9-SUMO-RanGAP1-RanBP2 case only with larger decays. However, Glu85 (residue 221) of SUMO and Pro451 of RanGAP1 (residue 254) shows shorter decays implying stronger correlations for these two residues. Glu85 is in the C-terminal region of SUMO and decrease in correlation in the presence of RanBP2 can be interpreted as increased mobility caused by RanBP2. Pro451 of RanGAP1, on the other hand, resides on the same mobile loop that causes the RMSD of RanGAP1 to increase in Ubc9-SUMO-RanGAP1-RanBP2 case (Figure 3.4).

Despite a few residues, the autocorrelations of the orientational vectors decreases without RanBP2. This implies that the orientational space is less restricted without RanBP2. Considering the RanGAP1, in Ala531 (333th residue) is observed to lose correlations that were maintained in the Ubc9-SUMO-RanGAP1-RanBP2 complex structure. This residue is also located around the LKSE motif, which can have functional importance in positioning the Lys524 in right position for SUMO conjugation.

Results for the parallel runs (Figures A.5 and Figure B.5) for both structures confirm the observations presented here in Figure 3.11 and Figure 3.12.

3.1.5. Clustering

To obtain the main conformational states and the changes, the conformations from the trajectories are clustered via k-means clustering as RMSD being the similarity measure. The x-axis shows the time of the trajectory whereas the y-axis shows the RMSD of each conformation from the centroids of the respective clusters. See Methods section for the details.

3.1.5.1. Clustering of Ubc9-SUMO-RanGAP1-RanBP2 Trajectory. The trajectory of Ubc9-SUMO-RanGAP1-RanBP2 simulation is 59 ns long, the equilibration period is determined as 10 ns from the RMSD plots.

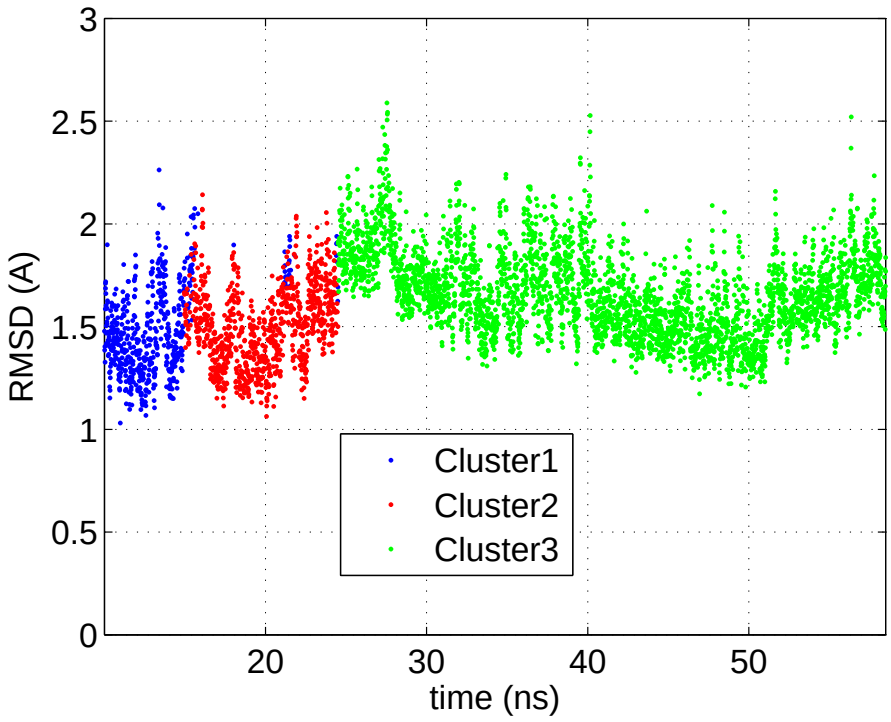

Figure 3.13. Clustering of the trajectory of Ubc9-SUMO-RanGAP1-RanBP2 tetramer complex at 3.00 Å as the threshold for the overall RMSD

Excluding the first 10 ns, the clustering analysis revealed three clusters when 3.0 Å threshold is selected for the RMSD of the heavy backbone atoms for the Ubc9-SUMO-RanGAP1-RanBP2 complex structure (Figure 3.13). The clusters are distinctly separated from each other. The third cluster (Green) starting from 25 ns takes the

largest portion of the conformational ensemble. The case is similar in the parallel simulation (See Figure A.6); three clusters among which the last cluster comprise the largest percentage of the sampled conformational space. The representative member of each cluster is presented in Figure 3.14, where the conformational changes between clusters can be inspected.

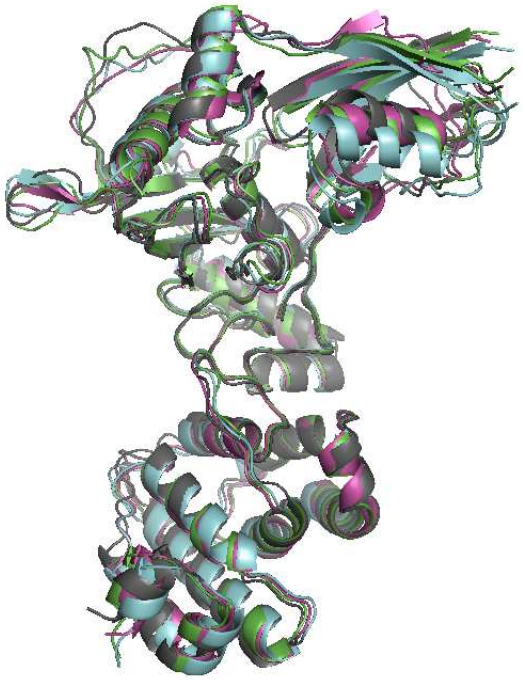

Figure 3.14. The representative members of each cluster aligned onto the minimized structure. (Grey: minimized structure; Green: Cluster1; Blue: Cluster2; Pink: Cluster3)

Slight changes are observed between the best members of the clusters. For instance, the Asp33 (residue 32) loop is puckering slightly in accordance with previous studies. (Karaca *et al.*, 2010) Along, SUMO bends and rolls itself towards RanGAP1 and moves back to original position in the minimized structure. With the highly mobile nature, the RanBP2 fragment shows very different conformations between clusters. Since the RanBP2 fragment is highly mobile, the clustering can be misleading due to increased RMSD caused by RanBP2. For that reason, clustering of the trajectory excluding the RMSD of RanBP2 fragment will be carried out.

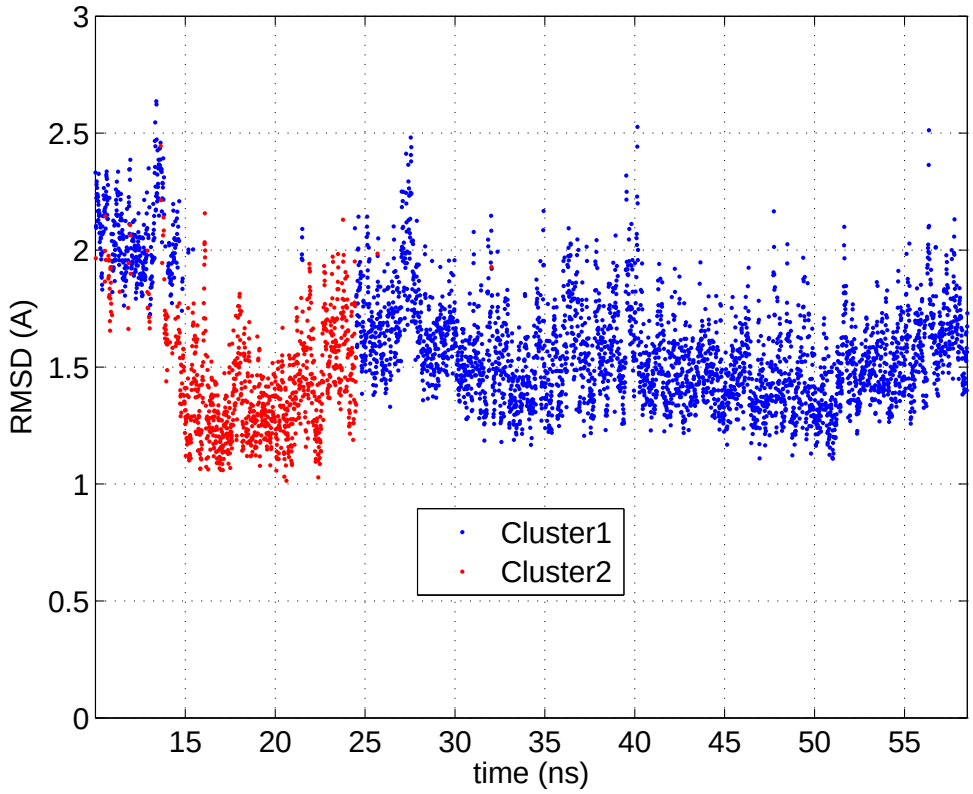

Figure 3.15. Clustering of the trajectory of Ubc9-SUMO-RanGAP1-RanBP2 tetramer complex at 3.00 Å as the threshold for the RMSD of Ubc9-SUMO-RanGAP1

Figure 3.15 represents the distribution of conformations sampled throughout the simulation of Ubc9-SUMO-RanGAP1-RanBP2 structure at 300 K. RMSD of Ubc9-SUMO-RanGAP1 structure is selected as the similarity measure. RanBP2 is extracted from the calculations as it fluctuates a lot and may defect the clustering result. Excluding the RMSD of RanBP2, clustering resulted in two main groups of clusters. However, it is pretty obvious that the first cluster (Blue) is the combined version of the first and third cluster shown in Figure 3.13. Since the RMSD of RanBP2 is excluded, 3.00 Å threshold is too much to distinguish different clusters, thus the RMSD threshold is lowered to 2.75 Å in order to obtain distinctive clusters.

Figure 3.16 displays the distribution of conformations created by the simulation of Ubc9-SUMO-RanGAP1-RanBP2 structure at 300 K. The similarity measure is again RMSD of Ubc9-SUMO-RanGAP1 structure, however, the threshold is lowered to 2.75 Å to be able to distinguish between clusters. As one can see, the pattern in Figure

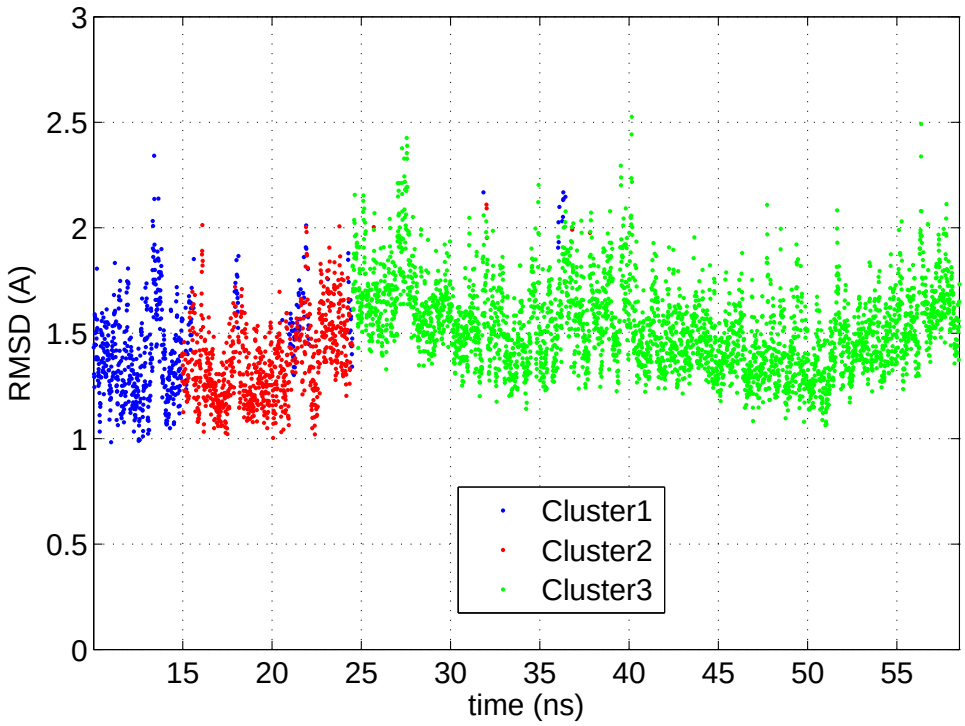

Figure 3.16. Clustering of the trajectory of Ubc9-SUMO-RanGAP1-RanBP2 tetramer complex at 2.75 Å as the threshold for the RMSD of Ubc9-SUMO-RanGAP1

3.13 is obtained again. Three different clusters with clear transitions are obtained. The conformations distributed to each cluster are approximately same as the case that Ubc9-SUMO-RanGAP1-RanBP2 simulation is clustered according to overall RMSD (Figure 3.13), hence, Figure 3.14 is still valid for pointing out the conformational differences between clusters. It is important to note that, Cluster1 and Cluster3 are close to each other as they merge and form Cluster1 in Figure 3.15 where Ubc9-SUMO-RanGAP1-RanBP2 simulation is clustered via RMSD of Ubc9-SUMO-RanGAP1 at 2.75 Å threshold. Cluster1 and Cluster3 are closer ensembles compared to Cluster2. This can be interpreted as although the conformational ensemble shifts to different space (Cluster2), it comes back to a more similar (compared to Cluster2) space of conformations which is Cluster3.

Figures 3.17-3.19 represent the cross correlations averaged over the conformations sampled in each cluster (See Figure A.7-A.9 for parallel runs). Close inspection of Figure 3.17, 3.18 and 3.19 reveal that correlations of LKSE motif of RanGAP1

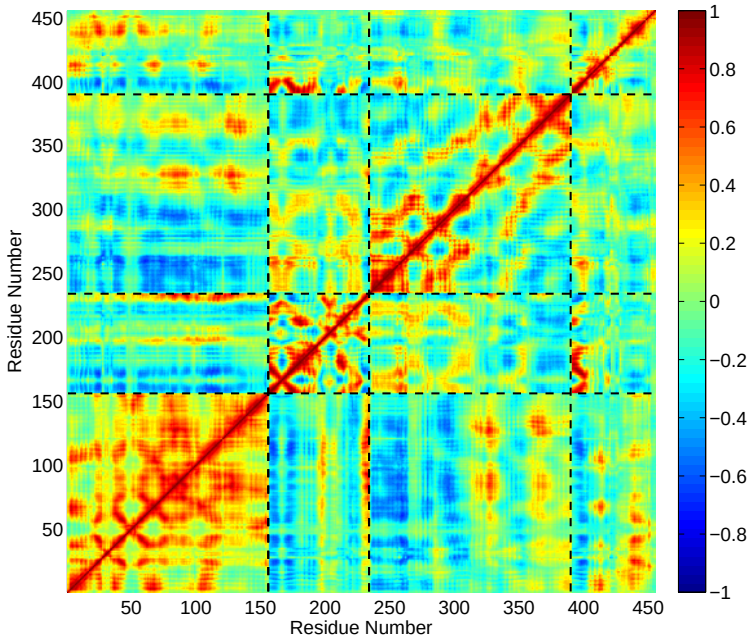

Figure 3.17. Cross Correlations of the simulation of Ubc9-SUMO-RanGAP1-RanBP2 complex (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1) calculated over the first cluster (10 ns - 15 ns)

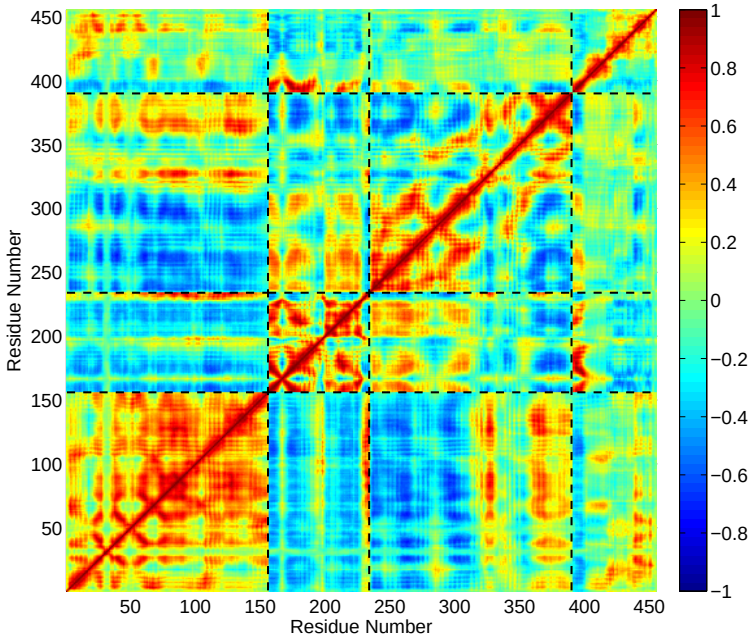

Figure 3.18. Cross Correlations of the simulation of Ubc9-SUMO-RanGAP1-RanBP2 complex (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1) calculated over the second cluster (15 ns - 25 ns)

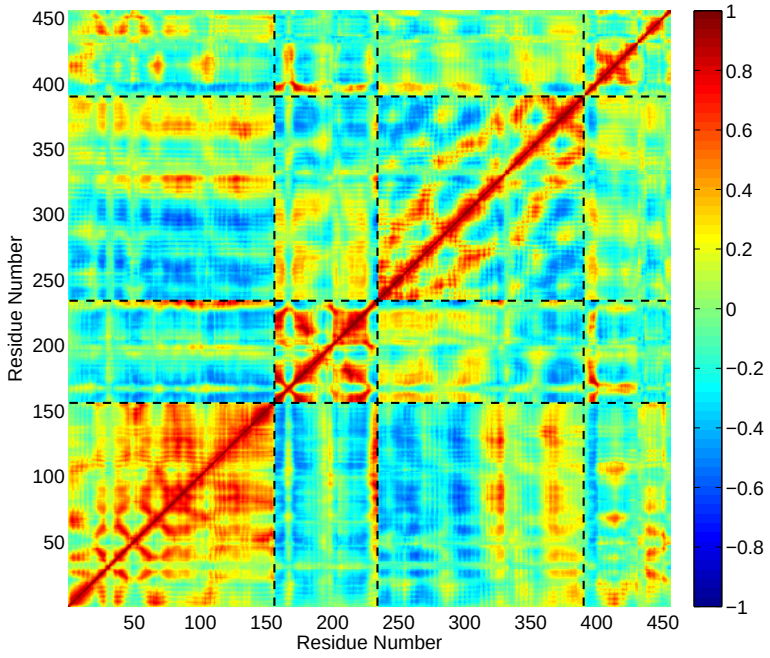

Figure 3.19. Cross Correlations of the simulation of Ubc9-SUMO-RanGAP1-RanBP2 complex (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1) calculated over the third cluster (25 ns - 59 ns)

(residues 326-329) with Asp100-Lys101 of Ubc9 do not remain strong throughout the simulation. Whenever the loop that contains Asp33 of Ubc9, Loop 2, shows higher correlations with the C-terminal region of RanBP2, the correlations of Ubc9 with the RanGAP1 increases as in cross correlations of second cluster (Figure 3.18). Besides, the increased correlations of Loop 2 region and Ala131-Gln139 (residues 130-138) of Ubc9 with Leu555-Pro566 (residues 358-369) leads to stronger coupling of SUMO-Target. Moreover, the loop of Cys93 (residue 92) of Ubc9 and the loop formed by Lys74-Phe77 (residues 73-76) which faces the catalytic Cys93 makes strong correlations with the SUMO binding motif Leu523-Glu526 of RanGAP1 (residues 326-329). These correlations are also observed to increase with increased Ubc9-RanBP2 correlations.

Furthermore, Ubc9 shows positive correlations with SUMO residues: Gln29, Asn60, Gly68 and Gly97 (residues 166, 197, 205 and 234). These correlations gets weaker as simulation proceeds, the weakening of these correlations from cluster to clusters can be observed in Figure 3.17, 3.18 and 3.19.

On the second cluster, the correlations of Tyr21-Leu24, Lys37-His43 , Gly68-Val90 of SUMO (residues 158-161, 174-180, 205-227) with the N-terminal region of RanBP2 and the two loops of RanGAP1 which contain Asp462 and Lys500 (residues 265 and 300) is stronger than the other two clusters. Giving benefit of a doubt, one can suggest that an increase in the correlations of the loop that contains Asp33 of Ubc9, Loop 2 with the C-terminal region of RanBP2 leads to an increase in the coupling between SUMO and RanGAP1. Ubc9-RanBP2 interactions can be suggested as a reason for the increased correlations of SUMO and RanGAP1 but not vice versa because it is already known that RanBP2 exerts its allosteric effect as an E3 via Ubc9. Yet, it is not well established how Ubc9 transfers this effect to SUMO.

The results for the parallel simulation are presented on Appendix A (See Figure A.7-A.9 for parallel runs), they are observed to agree the results and discussion presented here.

3.1.5.2. Clustering of Ubc9-SUMO-RanGAP1 Trajectory. The trajectory of Ubc9 - SUMO-RanGAP1 simulation is 48.5 ns long and the equilibration period is determined as 3 ns from the RMSD analysis. Leaving out the first 3 ns, the clustering analysis revealed five and six clusters when 3.00 Å and 2.75 Å threshold is selected respectively for the RMSD of the heavy backbone atoms (Figure 3.20 and Figure 3.21). The distribution of the conformations into the clusters does not follow the same profiles in the case of the Ubc9-SUMO-RanGAP1-RanBP2 complex structure; i.e. the members of the clusters may appear at different times of the trajectory. On the other hand, similar to the Ubc9-SUMO-RanGAP1-RanBP2 complex simulation, the parallel simulation (see Figure B.6) yields only three clusters with 3.00 Å selected as threshold for RMSD. Yet, the distribution of the conformations of the clusters in time reflects the effect of the absence of RanBP2: The conformational transitions between different conformational states are more frequent without RanBP2. Both results imply that RanBP2 affects the kinetics of Ubc9-SUMO-RanGAP1 system.

As the threshold for RMSD, 2.75 Å can be selected for Ubc9-SUMO-RanGAP1

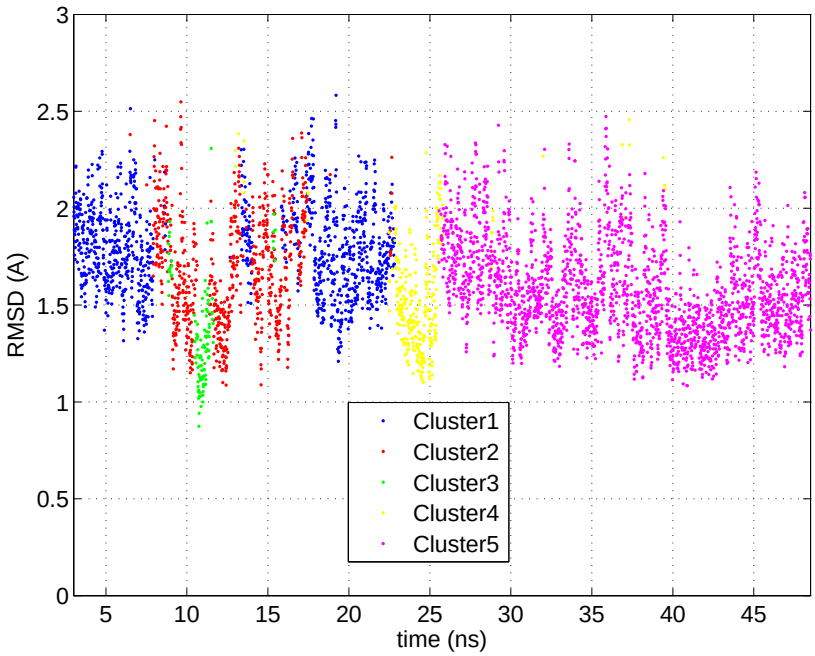

Figure 3.20. Clustering of the simulation of Ubc9-SUMO-RanGAP1 ternary complex at 3.00 Å threshold for RMSD

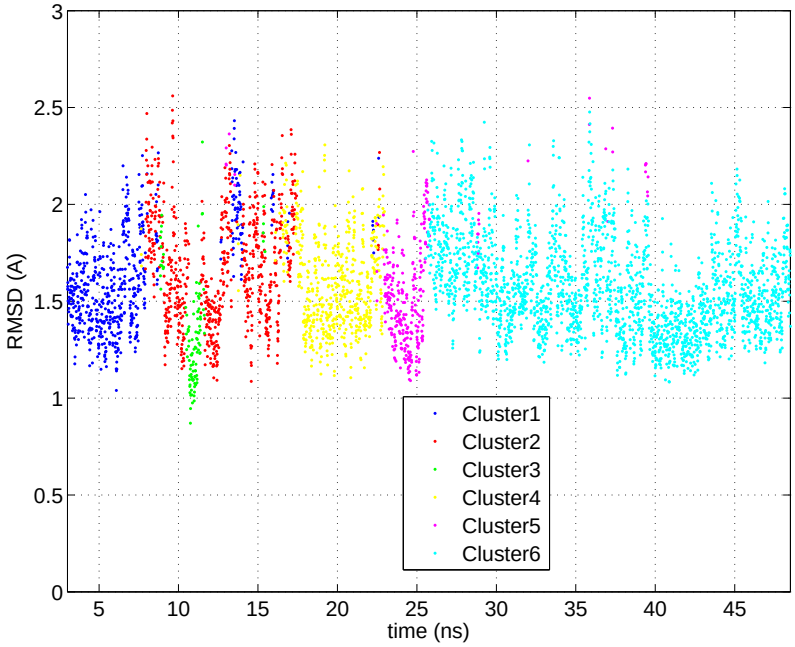

Figure 3.21. Clustering of the simulation of Ubc9-SUMO-RanGAP1 ternary complex at 2.75 Å threshold for RMSD

system. The superimposed cluster centroids formed by clustering at 2.75 Å can be seen In Figure 3.22. The four clusters that form the first 25 ns of the trajectory are not well separated from each other. After 25 ns, a single cluster, which spans the conformations generated until the end of the trajectory, is formed. This suggests that at around 25 ns the system reaches a new conformational state that is more stable than the former fours clusters. In Figure B.6, the results for the parallel simulation are displayed. Three clusters are formed in parallel simulation; however, the transitions between conformations are more frequent than the Ubc9-SUMO-RanGAP1-RanBP2 simulation. Furthermore, Cluster1 and Cluster2 are revisited many times after 25 ns, but after 47 ns a third cluster is formed similar to the shift a third cluster after 25 ns in the original Ubc9-SUMO-RanGAP1 simulation presented in this section.

Figure 3.22. The representative members of the clusters are aligned onto the minimized structure. (Green: Cluster1; Blue: Cluster2; Purple: Cluster3, Yellow: Cluster4, Pink: Cluster5, Grey: Cluster6)

The main differences between the clusters can be observed in Figure 3.22. Without RanBP2, SUMO first moves towards RanGAP1, then comes back to the original

position and stays there for the rest of the simulation time. This is the same motion which was observed from time to time in Ubc9-SUMO-RanGAP1-RanBP2 simulation in Figure 3.4. Thus, it can be concluded that RanBP2 helps SUMO to maintain the right orientation for sumoylation.

On the two clustering results, Figure 3.16 and Figure 3.20, comparison of two clustering analysis, carried out at 3.00 Å as the RMSD threshold, reveals that presence of RanBP2 reorganizes the conformational ensemble into a more stable level. Without RanBP2, till 25 ns (47 ns for the parallel run) the conformations revisit the first cluster of conformations, moreover the transitions between clusters are not distinct. However, in the presence of RanBP2, there are two clusters until 25 ns, which are separated neatly, and after 25 ns system reaches a new conformational ensemble and chose to sample similar conformations.

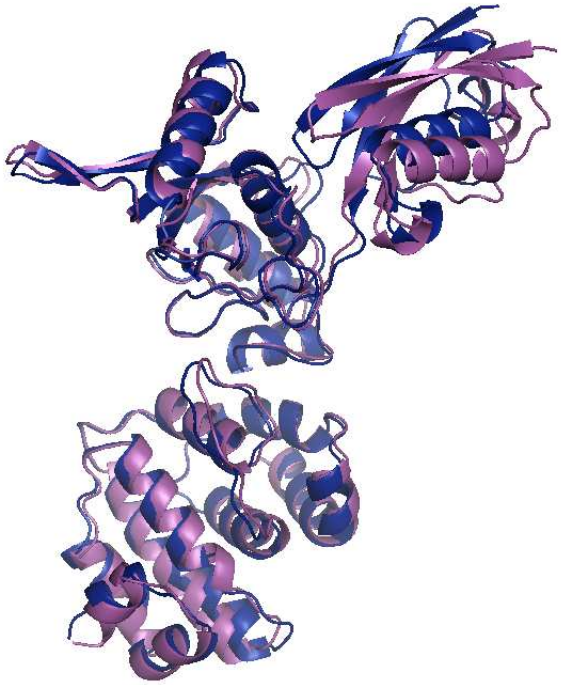

Figure 3.23. The representative members of the last clusters generated from the simulations of Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 structures (Blue: sixth cluster of Ubc9-SUMO-RanGAP1 simulation, Pink: third cluster of Ubc9-SUMO-RanGAP1-RanBP2 simulation)

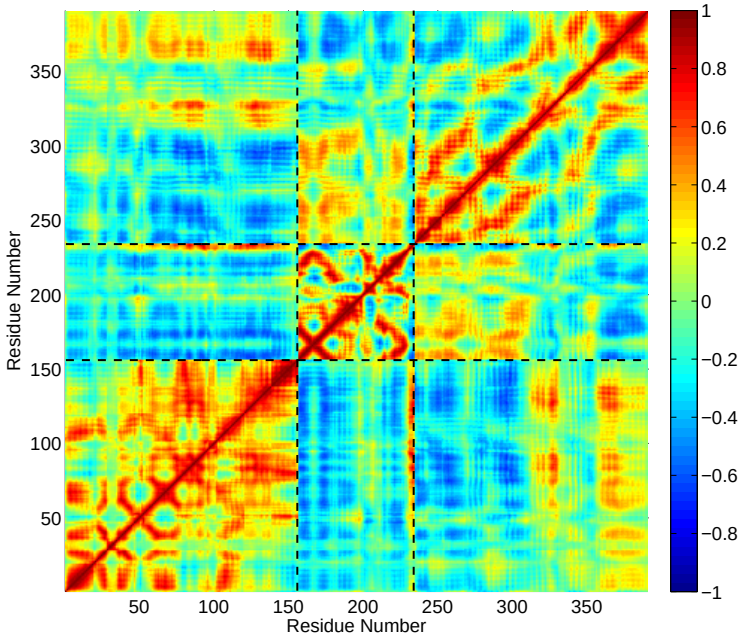

Figure 3.24. Cross Correlations from the simulation of Ubc9-SUMO-RanGAP1 complex (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1) calculated over the first cluster

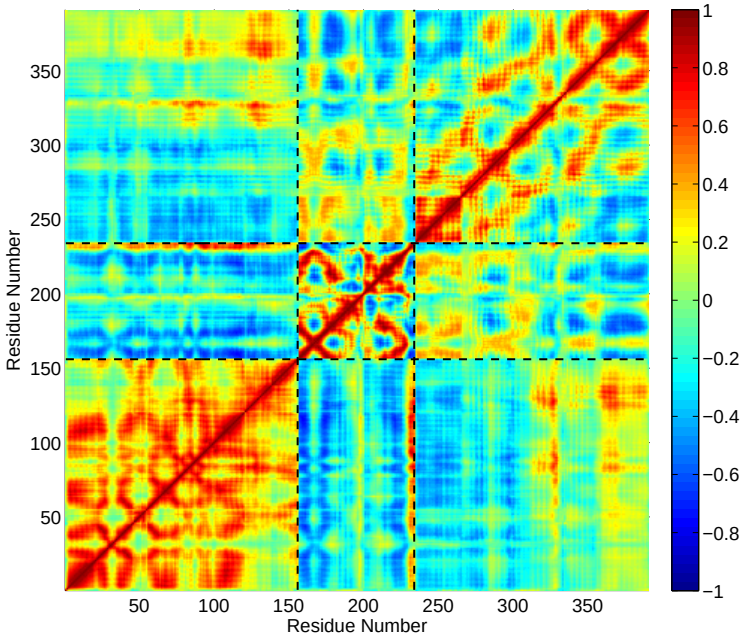

Figure 3.25. Cross Correlations from the simulation of Ubc9-SUMO-RanGAP1 complex (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1) calculated over the second cluster

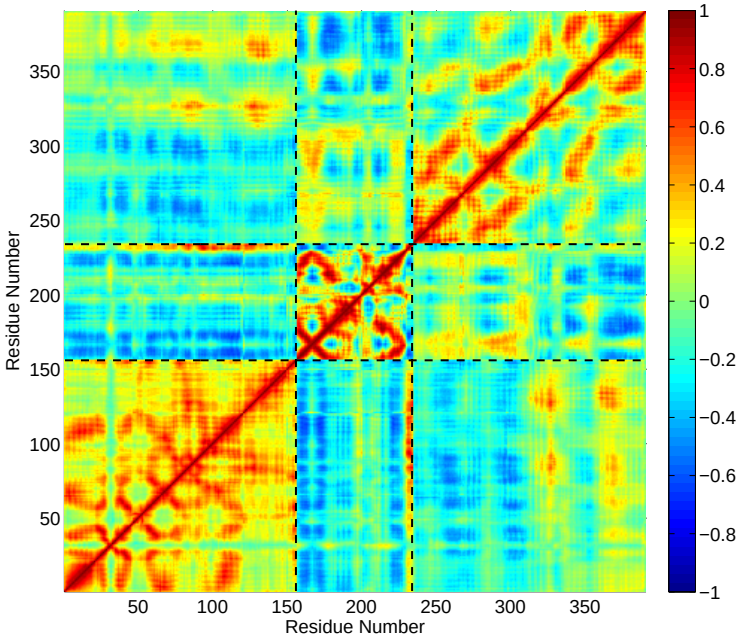

Figure 3.26. Cross Correlations from the simulation of Ubc9-SUMO-RanGAP1 complex (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1) calculated over the third cluster

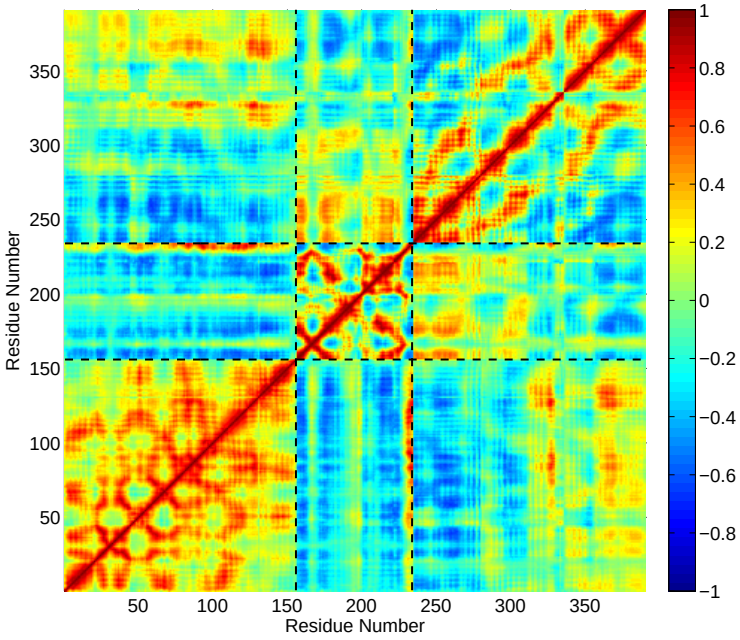

Figure 3.27. Cross Correlations from the simulation of Ubc9-SUMO-RanGAP1 complex (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1) calculated over the fourth cluster

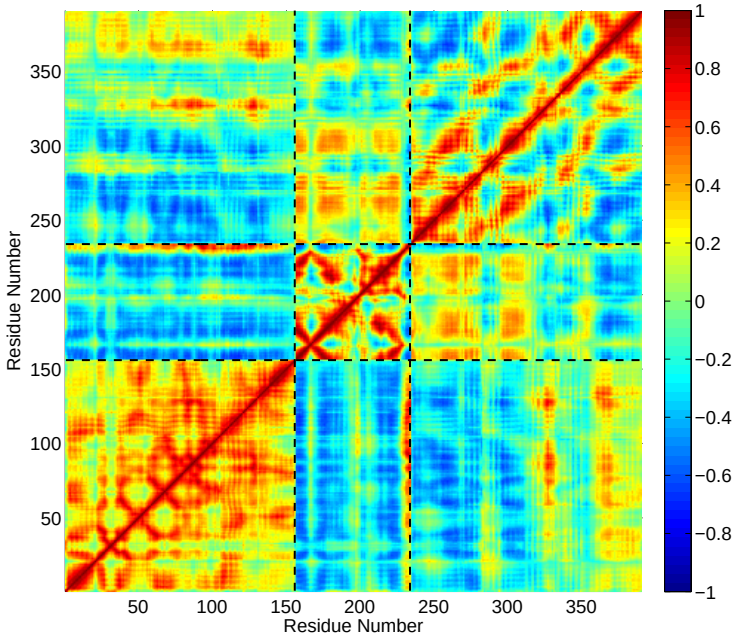

Figure 3.28. Cross Correlations from the simulation of Ubc9-SUMO-RanGAP1 complex (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1) calculated over the fifth cluster

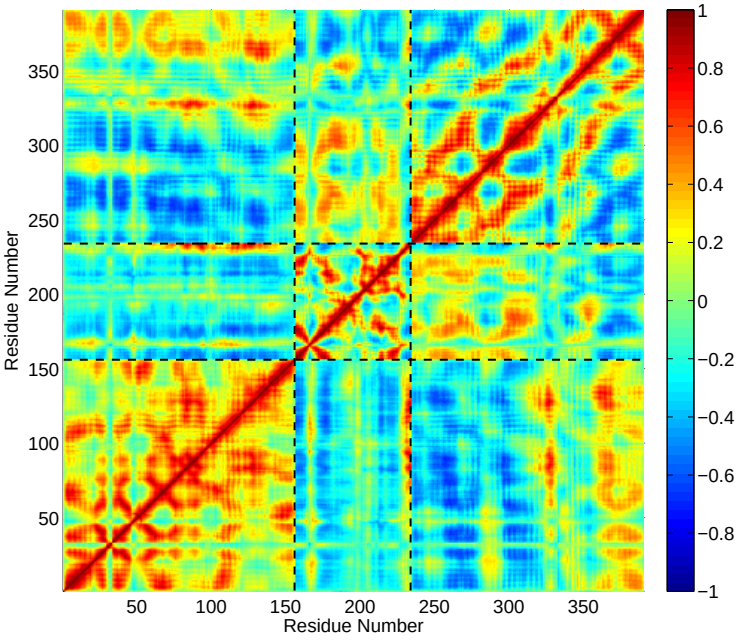

Figure 3.29. Cross Correlations from the simulation of Ubc9-SUMO-RanGAP1 complex (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1) calculated over the sixth cluster

In Figure 3.23, Ubc9-SUMO-RanGAP1 structures taken from the centroid of the third cluster generated by the clustering of Ubc9-SUMO-RanGAP1-RanBP2 simulation and centroid of the sixth cluster generated by the clustering of Ubc9-SUMO-RanGAP1 simulation are aligned. As one can see, the conformations of Ubc9 and RanGAP1 resembles however in Ubc9-SUMO-RanGAP1-RanBP2 case, SUMO bends itself towards RanGAP1. RanBP2 helps SUMO to maintain its original position in the crystal structure.

Loop 2 region (Lys30-Met36, residues 39-35), which contains the RanBP2 binding site of Ubc9 (Val27-Glu42, residues 26-41) shows strong correlations with Leu555-Pro566 of RanGAP1 (residues 358-369). Leu555-Pro566 of RanGAP1 is the Ubc9 binding site which faces the built-in E3 motif of Ubc9 (Ala131-Arg141 of Ubc9, residues 130-140). This correlation is observed in Cluster 1, 4, 5, 6 being more pronounced in Clusters 1, 5 and 6.

Furthermore, the catalytic loop, which holds catalytic Cys93 (residue 92), Lys74-Phe77 (residue 73-76), His83-Ser89 (residues 82-88), which helps orientation of Gly97 of SUMO, the catalytic site Cys93, Asp100-Lys101 (residues 99-100), which are responsible for target recongnition, Asn124-Pro128 (residues 123-127), which is in contact with the conserved SUMO binding motif of target (LKSE), the built-in E3 and specific target binding site Ala131-Arg141 (residues 130-140) of Ubc9 makes positive correlations with Thr511-Leu522 (residues 314-325) of RanGAP1, which faces the built-in E3 region of Ubc9 and Leu523-Gly526 (residues 326-329) of RanGAP1, which is the conserved SUMO binding motif. These correlations are observed in all clusters however amplified in Clusters 1, 5 and 6.

Besides, Tyr21-Leu24, Lys37-His43, Gly68-Val90 of SUMO (residues 158-161, 174-180, 205-227) with the two loops of RanGAP1 that contain Asp462 and Lys500 (residues 265 and 300). These correlations are observed in all clusters but gets stronger in Clusters 1, 5 and 6.

It is interesting to note that whenever Loop 2 region is correlated with Thr511-

Leu522, Leu523-Gly526 (residues 314-329) and Leu555-Pro566 (residues 358-369) of RanGAP1, SUMO shows strong correlations with RanGAP1 too. The results for the parallel simulation also support this argument (See Figure B.7 to Figure B.9). The same behavior is also observed with the Ubc9-SUMO-RanGAP1-RanBP2 complex structure. Thus, the functional importance of this residue suggested by Karaca *et al.*, 2010 is futher confirmed in the present analysis. It is apparent that Loop 2 region is of great importance in SUMO-RanGAP1 coupling, which was already shown in Karaca *et al.*, 2010; Tozluoglu *et al.*, 2010. It can be suggested that Loop 2 may have role in SUMO and RanGAP1 conjugation.

It can be concluded that the coupling of Loop 2 region with the RanGAP1 is provided by the built-in E3 region of Ubc9. This can be claimed due to the fact that increasing correlations of Ala131-Arg141 (built-in E3) of Ubc9 with Thr511-Leu522, Leu523-Glu526 and Leu555-Pro566, increases the coupling of Tyr21-Leu24, Lys37-His43, Gly68-Val90 of SUMO (residues 158-161, 174-180, 205-227) with the two loops of RanGAP1 that contain Asp462 and Lys500 (residues 265 and 300). The case is similar in Ubc9-SUMO-RanGAP1-RanBP2, however in that sequence of events start with coupling of RanBP2 with Loop 2. In Ubc9-SUMO-RanGAP1 case, the coupling is triggered by the built-in E3 region of Ubc9, Ala131-Arg141.

3.1.5.3. Clustering of Different Trajectories. Having both the Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 trajectories in hand, along with the similar clustering profiles discussed in Section 3.1.5.1. and 3.1.5.2., the clustering analysis on the merged trajectory of the two simulations are carried out. To have a fair comparison, the RanBP2 fragment in the Ubc9-SUMO-RanGAP1-RanBP2 complex structure is not taken into consideration. Having two different trajectories, the Ubc9-SUMO-RanGAP1 structures which obtained from the Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 are merged and clustered at 4.00 Å as the RMSD threshold. Although, the clustering of single trajectories are carried out at 3.00 Å, a higher value is needed here to group two different trajectories under similar clusters. The results are given in Figure 3.30.

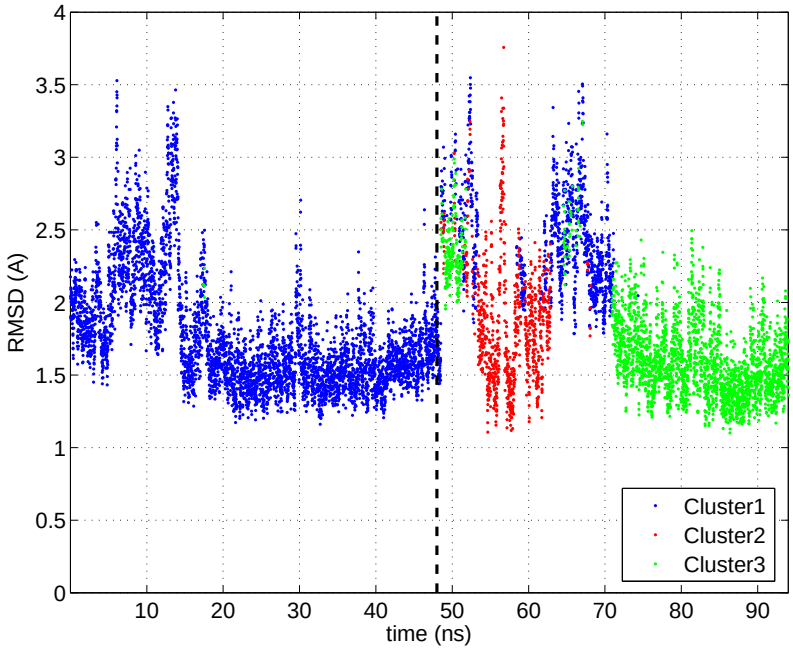

Figure 3.30. Clustering profile for Ubc9-SUMO-RanGAP1 structures obtained from Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 trajectories at 4.00 Å RMSD (Left portion of the dashed line: Ubc9-SUMO-RanGAP1-RanBP2 trajectory, Right portion of the dashed line: Ubc9-SUMO-RanGAP1 trajectory)

In Figure 3.30, the clustering of Ubc9-SUMO-RanGAP1 structures obtained from simulation of Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 complexes are presented. Left side of the dashed line is composed of the Ubc9-SUMO-RanGAP1 conformations that are obtained from the Ubc9-SUMO-RanGAP1-RanBP2 simulation and the right part is composed of structures that are obtained from the Ubc9-SUMO-RanGAP1 simulation. As a result, three main clusters are formed. With RanBP2 (left side of the graph), all of the conformation created by Ubc9-SUMO-RanGAP1-RanBP2 simulation fits into a single cluster (Blue). However, without RanBP2 conformations belonging to three different clusters are formed. Moreover, only about 15% of the conformation created belongs to the same cluster that Ubc9-SUMO-RanGAP1-RanBP2. Presence of RanBP2 shifts conformations in such a way that the occurance of blue conformations increase dramatically. It is obvious that the presence of RanBP2 stabilizes the Ubc9, SUMO and RanGAP1.

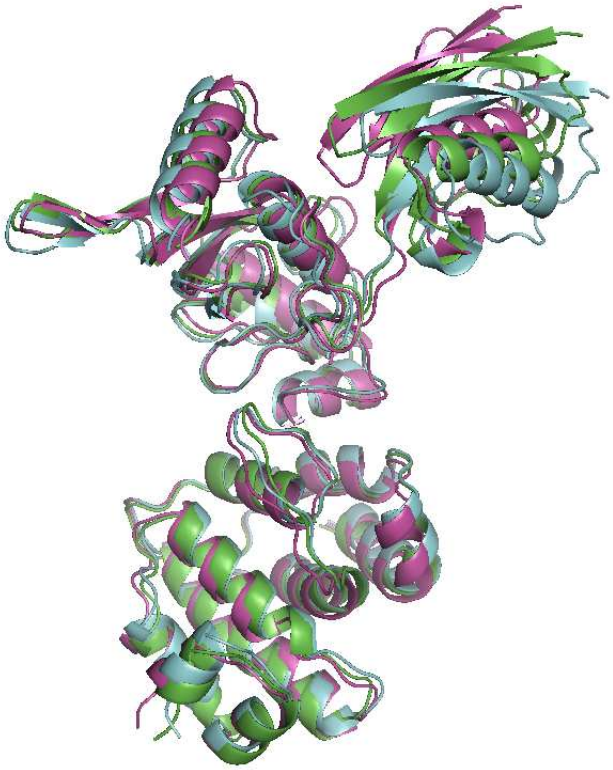

Figure 3.31. Representative members of each cluster formed by clustering of simulations of Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 complexes at 4.00 Å as the RMSD threshold. (Green: Cluster1; Blue: Cluster2; Pink: Cluster3)

In Figure 3.31 the centroids obtained from the clustering of of Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 simulations at 4.00 Å RMSD threshold for the RMSD of Ubc9-SUMO-RanGAP1 chains from both simulations. Green structure is the centroid of the first cluster and this is the closest to the crystal structure. The RMSD of the first centroid is 1.58 Å whereas second and third have 2.77 Å and 2.00 Å. We can conclude that RanBP2 provides Ubc9, SUMO and RanGAP1 to maintain the original position in the crystal structure. Without RanBP2, SUMO moves back and forth towards RanGAP1.

In addition to the Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 simulations, there are parallel runs for each of them. Having four trajectories, Ubc9-SUMO-RanGAP1 regions are merged trajectories and clustered. In Figure 3.32, clus-

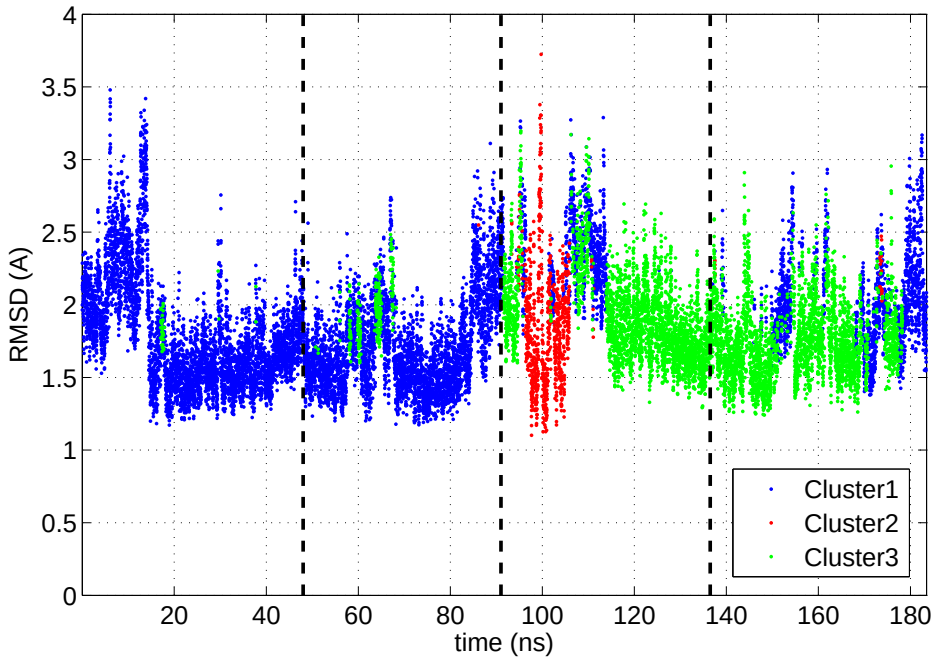

Figure 3.32. Clustering profile for two simulations of Ubc9-SUMO-RanGAP1-RanBP2 complex and two simulations of Ubc9-SUMO-RanGAP1 complex at 4.00 Å RMSD threshold

tering results for the two parallel simulations of Ubc9-SUMO-RanGAP1-RanBP2 and two parallel runs of Ubc9-SUMO-RanGAP1 structures are presented. The RMSD threshold is selected as 4.00 Å for the overall RMSD of Ubc9, SUMO and RanGAP1. First portion belongs to the conformations of Ubc9-SUMO-RanGAP1-RanBP2 simulation, third portion is from the Ubc9-SUMO-RanGAP1 simulation, and second and fourth portions are for the parallel runs for each of them. The Ubc9-SUMO-RanGAP1-RanBP2 simulations sample the conformations mostly from the same blue cluster. About 99% of the conformations sampled in Ubc9-SUMO-RanGAP1-RanBP2 simulations belong to the first cluster, however, this percentage falls down to 15% in Ubc9-SUMO-RanGAP1 simulations. Not only the percentage of the conformations that belongs to first cluster (blue) decreases, the distribution of the conformations that belong to that cluster is distributed over 20 ns of time in the absence of RanBP2. There is a conformational shift in the ensemble of conformations of Ubc9-SUMO-RanGAP1 without RanBP2. Presence of RanBP2 restricts the conformational ensemble.

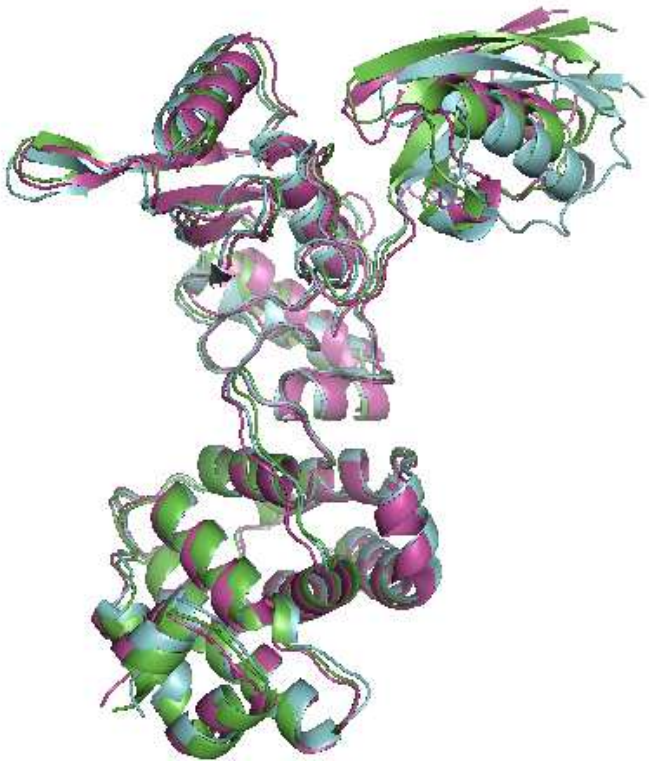

Figure 3.33. Representative members of the clusters formed by clustering of simulations of Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 complexes at 4.00 Å RMSD. (Green: Cluster1; Blue: Cluster2; Pink: Cluster3)

In Figure 3.33, the centroids of each cluster formed from Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 simulations along with their parallel runs at 4.00 Å being RMSD threshold for the RMSD of Ubc9-SUMO-RanGAP1 chains from each simulation. Green structure is the centroid of the first cluster and this is the closest to the crystal structure. The RMSD of the first centroid is 1.53 Å whereas second and third have 2.79 Å and 1.94 Å. It can well be concluded that RanBP2 helps to maintain the original positions of Ubc9, SUMO and RanGAP1 chains as in the crystal structure. In the absence RanBP2, SUMO is unstable and moves back (Cluster 3) and forth (Cluster 2) towards RanGAP1.

3.2. Effect of Asp33Ala mutation

In this section, the results for the simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 complex structure (at 300 K) are discussed. The mutant case is designed to validate the importance of Asp33 (residue 32) of Ubc9, which is already referred to as functionally important by Karaca *et al.*, 2010. The coupling of this residue with SUMO and RanGAP1 is also underlined in this work too. The comparative analysis of the mutant structure with the two wild type simulations will solidify our hypothesis that the increased coupling of Asp33 with C-terminal region of RanGAP1 increases SUMO-RanGAP1 coupling.

3.2.1. Root Mean Square Deviations (RMSD)

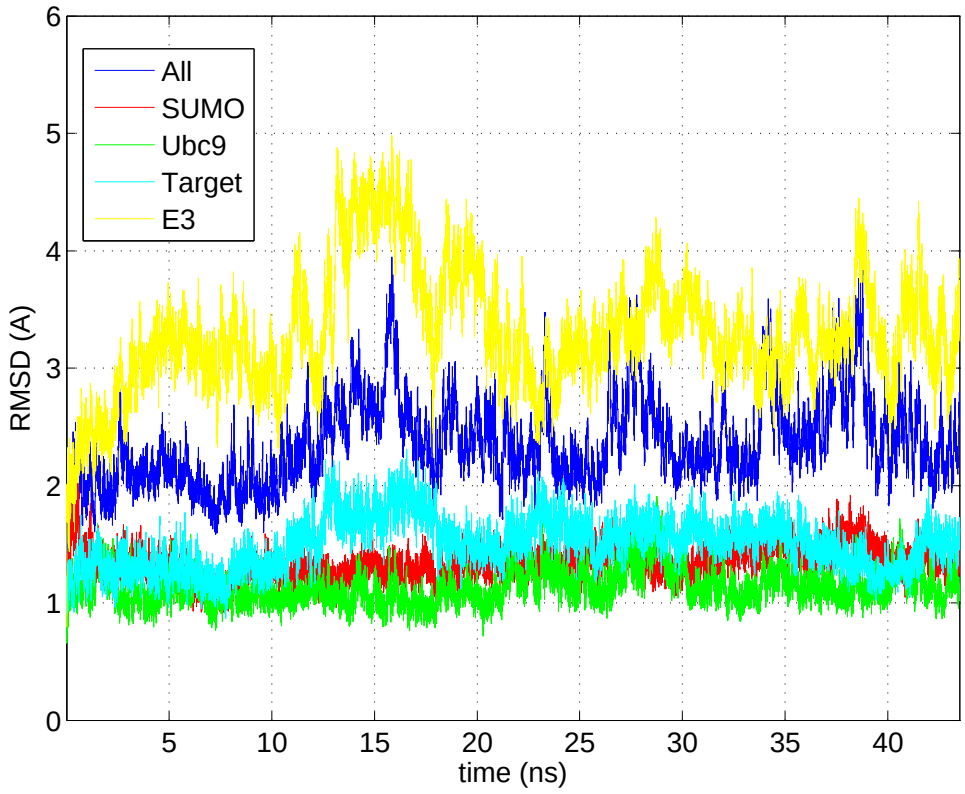

Figure 3.34. RMSD of simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 complex

In Figure 3.34, the RMSD values are presented for the Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 complex structure simulated at 300 K and its individual chains Ubc9, SUMO, RanGAP1 and RanBP2. Ubc9, SUMO and RanGAP1 are ob-

served to display rather stable RMSD profiles. The mean values of RMSD is 1.13 Å, 1.34 Å and 1.50 Å for Ubc9, SUMO and RanGAP1 respectively. RanBP2, on the other hand, displays a higher trend. The mean value of RanBP2 RMSD is 3.30 Å with a maximum of 5.00 Å. The trend of RanBP2 reflects the overall RMSD and causes it reach up 5.00 Å. The case was similar in the wild type simulations. As seen, the effect of mutation is not reflected on the RMSD values of individual; the values in the two cases are pretty close.

Here, it is nevertheless interesting to note the RanBP2 and RanGAP1 make a jump causing the overall RMSD to boost at around 16 ns. The reason of the RMSD jump can be observed in Figure 3.35.

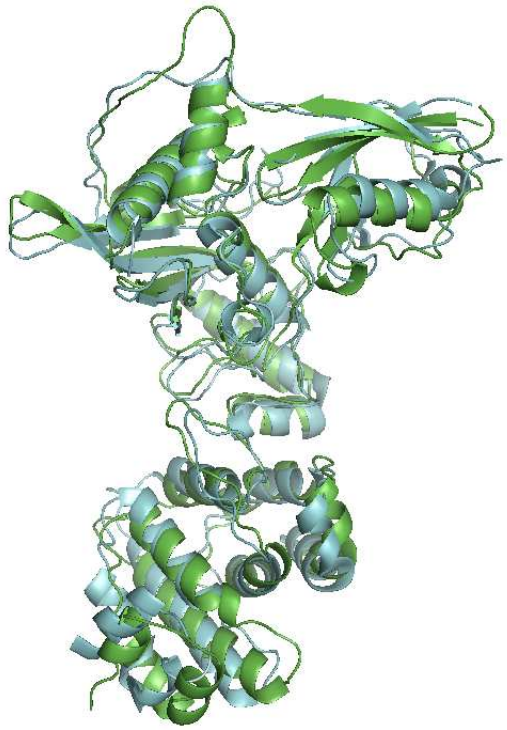

Figure 3.35. The alignment structure generated at 15.88 ns in simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 structure (Green) onto the crystal structure (Blue)

In Figure 3.35 structure generated at 15.88 ns in simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 structure (Green) is aligned onto the crystal structure (Blue), resuting in a RMSD value of 4.00 Å. The RMSD jump is mainly caused

by the rigid body motion of RanGAP1 towards SUMO. Especially, the N-terminal half bends towards SUMO. However, although RanGAP1 moves towards SUMO, SUMO and Ubc9 approaches itself which might make the sumoylation process harder. Besides, large conformational change is observed in the RanBP2 region.

3.2.2. Mean Square Fluctuations (MSF)

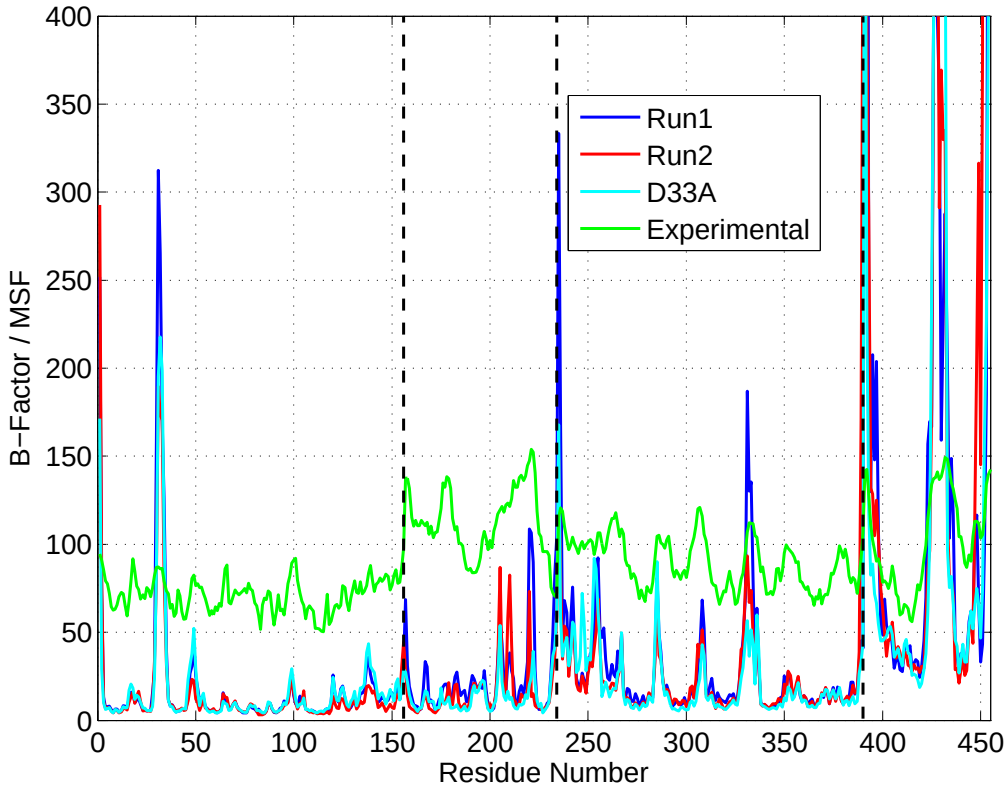

Figure 3.36. Experimental B-factors and calculated MSF results for the simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 via individual alignment

Figure 3.36 represents the MSF profile of the Asp33Ala mutant of the Ubc9-SUMO-RanGAP1-RanBP2 complex structure along with the MSF profiles for the wild type simulations of the same Ubc9-SUMO-RanGAP1-RanBP2 complex structure and an additional experimentally determined B-factor profile. The MSF values in Figure 3.36 calculated by the individual chain alignment. As seen, the mutation does not change the overall MSF profile. Even, the mobility of Asp33Ala (residue 32) is not altered significantly. Among the four chains, SUMO shows the maximum fluctuations. Residues Gly68, Asp73 and Glu83 (residues 205, 210 and 220) of SUMO, which reside on flexible loops, display decreased mobility. Particularly, Gly68 and Asp73 are located

on the loop that approaches the loop of Gly50 of Ubc9.

RanGAP1's, on the other hand, Glu444 (residue 247) shows increased mobility in the mutant structure. This residue is located on the flexible loop between Phe440-Lys445 (residues 243-248). This loop region of RanGAP1 belongs to the N-terminal region mentioned in Section 3.4.1 as moving towards SUMO in the mutant simulation at 15.88 ns. Ser505 and Lys528 of RanGAP1 (residues 308 and 331) display decreased mobility in the mutant case. Ser505 is located on a loop between Phe502-Asn508 (residues 305-311). These residues and the Phe502-Asn508 loop are found to be less mobile without RanBP2 in Section 3.1.2. The mutation might be blocking the transfer of the allosteric signal from RanBP2.

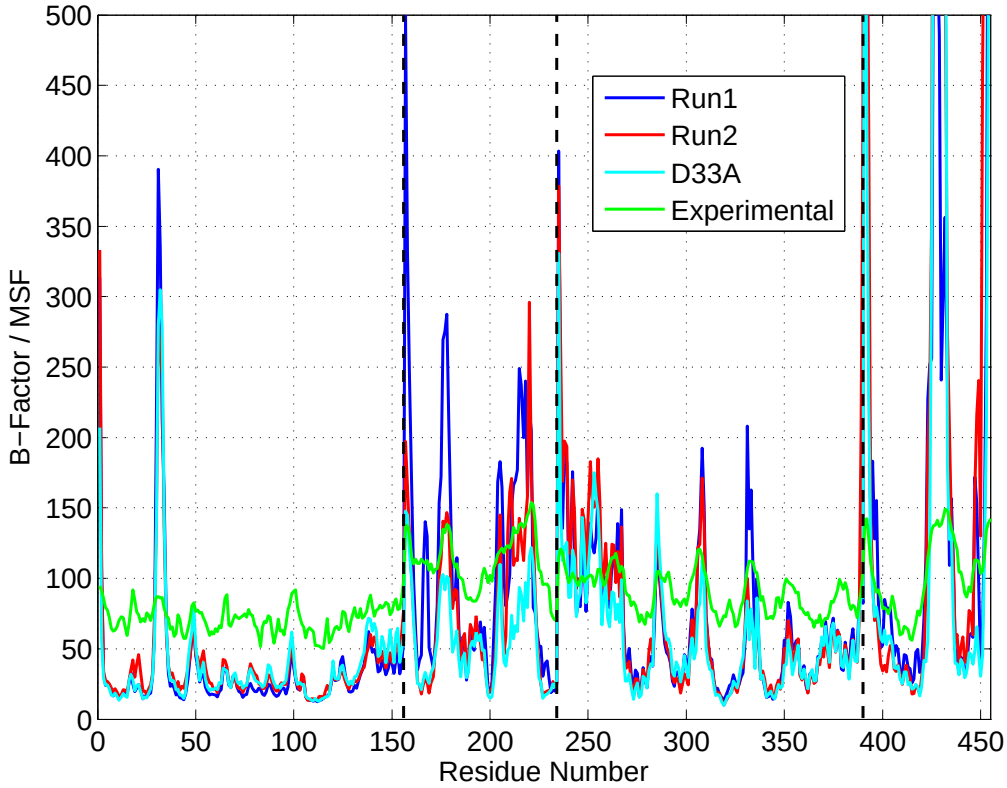

Figure 3.37. Experimental B-factors and calculated MSF results for the simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 via overall alignment (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1; 391-455: RanBP2)

Figure 3.37 represents the MSF profile of the Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 complex along with the MSF profiles for the wild type simulations of the same Ubc9-SUMO-RanGAP1-RanBP2 complex and an additional experimen-

tally determined B-factor profile. The MSF values in Figure 3.37 calculated by the overall alignment.

As seen, the mutation does not change the MSF profile for Ubc9. The mobility of Asp33Ala (residue 32) is not even changed. Among the four chains SUMO shows the maximum change. Asp30, Lys37-His43 and Gly68-Glu84 (residues 167, 174-180 and 205-221) of SUMO show decreased mobility in the mutant structure. Asp30 resides at the Ubc9 interface and Lys37-His43 region is at the RanBP2 interface. Gly68-Glu84 of SUMO is the loop region that approaches Gly50 of Ubc9 in wild type simulations. These regions lose mobility and rigidifies in the presence of Asp33Ala mutation.

RanGAP1, on the other hand, shows decreased mobility for Ala439, Ile456-Asp464 and Lys528-Tyr550 (residues 242, 259-267 and 331-353) regions. Ala439 and Ile456-Asp464 are on the N-terminal region and they are comprised of small loops. Several changes are listed about the change in mobility in the N-terminal region of RanGAP1. Those regions have no reported significance, however, they might be on the transmission pathway of the signal from Asp33.

3.2.3. Correlation between Fluctuations

Figure 3.38 shows the cross correlation map for the Ubc9-SUMO-RanGAP1-RanBP2 complex structure simulated with Asp33Ala mutation. Compared to the wild type simulation, Loop 2 with Asp33 makes weaker correlations with Thr511-Leu522, the SUMO binding motif: Leu523-Glu526 and Leu555-Pro566 (residues 314-325, 326-329 and 358-369) regions of RanGAP1 compared to wild type simulations of Ubc9-SUMO-RanGAP1-RanBP2. Thr511-Leu522 and Leu555-Pro566 regions face the Ala131-Arg141 region of Ubc9 which is known as built-in E3 of Ubc9. The built in E3 region of Ubc9 (Ala131-Arg141, residues 130-140) show weaker correlations with the Thr511-Leu522, Leu523-Glu526 and Leu555-Pro566 (residues 314-325, 326-329 and 358-369) of RanGAP1. More importantly, correlations of Asp100-Ala101 of Ubc9 (residues 99-100) show only weak positive correlation with the SUMO binding motif (Leu523-Glu526 of RanGAP1, residues 314-325), implying that Target recognition

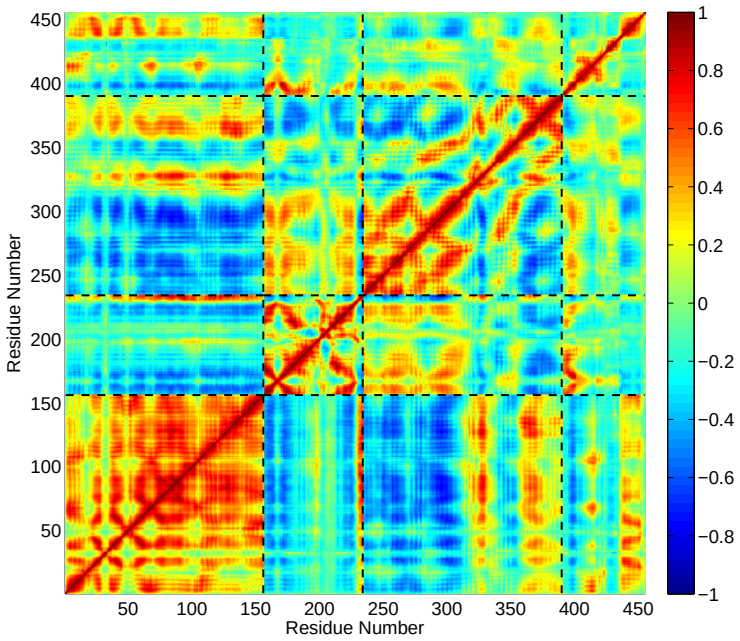

Figure 3.38. Cross Correlations of simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 tetramer complex calculated by alignment onto the initial minimized structure (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1; 391-455: RanBP2)

might be impaired.

Moreover, the loop Phe440-Lys445 (residue 243-248), which has increased mobility in the mutant case, of RanGAP1 shows positive correlations with Val27-Glu42 region, which is the RanBP2 binding site and includes the flexible Loop2 region, and Gly50 (residue 26-41 and 49) of Ubc9. Phe440-Lys445 (residue 243-248) of RanGAP1 showed increased mobility in the mutant case, and the presence of positive correlations in the latter indicates that the motion of those regions are still coupled yet with weaker correlations.

Gly50, Asp100 and Ala101 of Ubc9 are less correlated with Ser61 (residue 198) of SUMO (residue 198) compared to wild type simulations. Besides, correlations of Ser61 (residue 198) of SUMO with N-terminal region of RanGAP1 region has been faded. In addition, the correlations of Gly68 of SUMO with Thr511-Leu522, the SUMO

binding motif, Leu523-Glu526 and Leu555-Pro566 (residues 314-325, 326-329 and 358-369) regions of RanGAP1 has diminished with the presence of mutation. SUMO and RanGAP1 coupling gets weaker with the mutation.

The findings overall indicate that Asp33 of Ubc9 provides the coupling of Gly68 and Ser61 of SUMO with Ubc9 and RanGAP1. Presence of Asp33Ala mutation blocks the coupling of these residues with Gly50, Asp100 and Ala101 of Ubc9, N-terminal region of RanBP2 and the specified Thr511-Leu522, Leu523-Glu526 and Leu555-Pro566 (residues 314-325, 326-329 and 358-369) of RanGAP1. The increased correlations of the loop between Phe440-Lys445 (residue 243-248) of RanGAP1 with Val27-Glu42 region and Gly50 (residue 26-41 and 49) of Ubc9 indicate that Ubc9-RanGAP1 coupling is strengthened whereas coupling of SUMO-Ubc9 and SUMO-RanGAP1 weakened. In a nutshell, the Loop2 region especially the Asp33 residue of Ubc9 is of great importance in sumoylation dynamics.

3.2.4. Decay of Orientational Vectors by Time

The profile of Ubc9 (1-156) is similar to the wild type cases, only with a larger decay for Ala33 (Asp33Ala, residue 32 in Figure 3.39). About SUMO region (residues 157-234), large decays down to 0.67 in the wild type case is not encountered here, implying that SUMO motion is restricted in the mutant structure. RanGAP1 region, on the other hand, the loop that contains the Lys530 and Ile532 (residue 333 and 335) shows larger decays. This loop has functional importance on the grounds of the fact that the SUMO binding motif (LKSE) Leu523-Glu526 is located on that loop. The residues except the SUMO binding motif do not have functional significance but obviously the orientation of the residues on the loop affects the orientation of the binding motif thus affects the sumoylation machinery. Observing larger decays for Lys530 and Ile532 implies that this residue loses its memory faster than the wild type cases which makes the loop that the LKSE motif stands on less rigid.

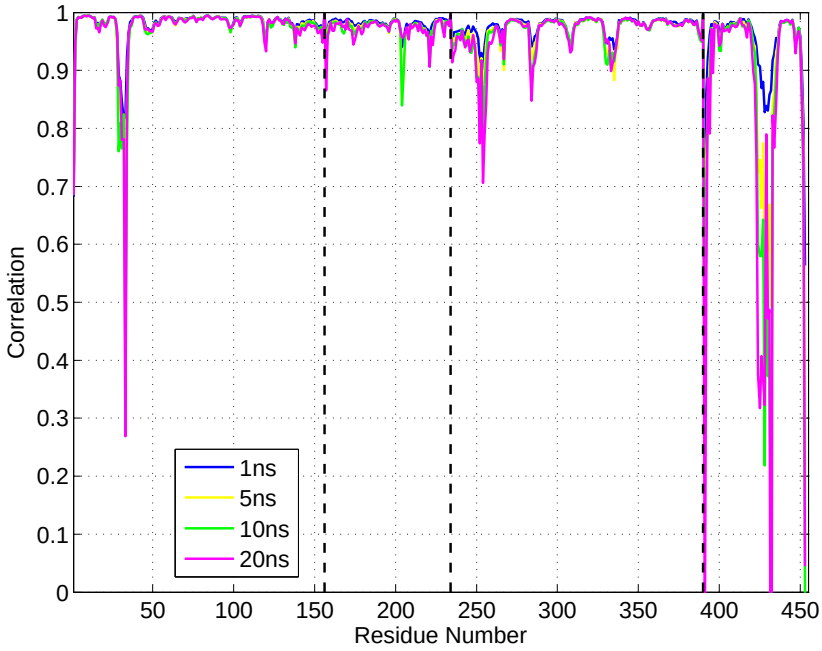

Figure 3.39. Decay of Orientational Vectors of simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 tetramer complex calculated by alignment onto the initial minimized structure (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1; 391-455: RanBP2)

3.2.5. Clustering

3.2.5.1. Clustering of the Mutant Ubc9-SUMO-RanGAP1-RanBP2 trajectory. Figure 3.40 shows the clustering of conformations sampled in the simulation of Asp33Ala mutant of the Ubc9-SUMO-RanGAP1-RanBP2 complex structure at 3.00 Å RMSD for the heavy backbone atoms. The RMSD of SUMO-Ubc9-RanGAP1 chains are taken into account, the RMSD of RanBP2 chain is exluded as it is highly flexible and might be misleading. The clustering profile given in Figure 3.40 for the mutant case is very different from the wild type simulations. The initial cluster (blue) is revisited after 30 ns. Such case is not also encountered in the parallel run for Ubc9-SUMO-RanGAP1-RanBP2 complex structure. Results imply that the ensemble of sampled conformations is altered. The representative members for the clusters are displayed in Figure 3.41.

In Figure 3.41, the centroids of the clusters obtained by clustering the simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 structure with 3.00 Å RMSD

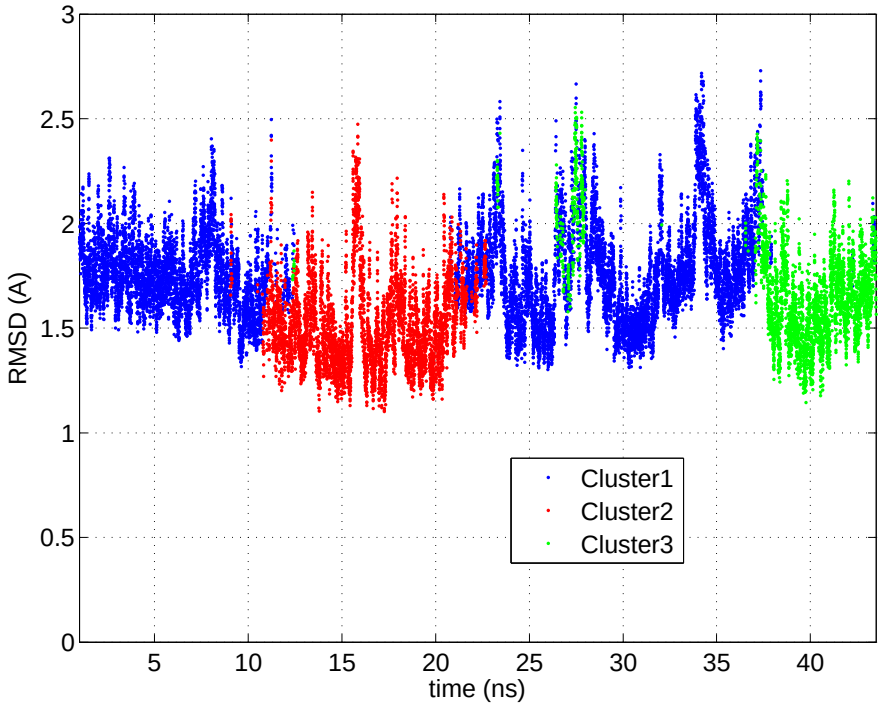

Figure 3.40. Clustering profile for simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2

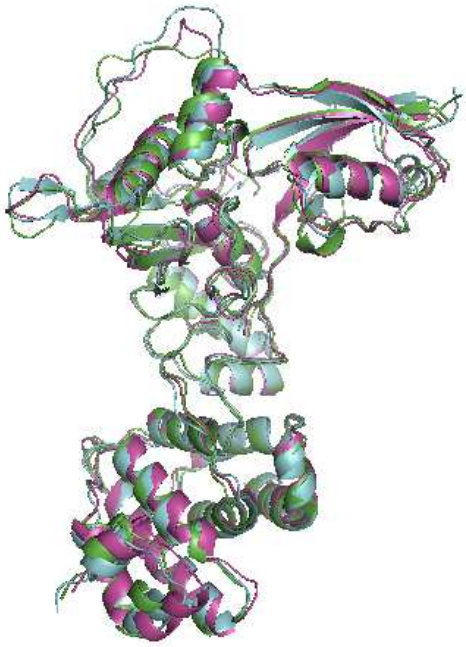

Figure 3.41. Representative members of the clusters formed bu clustering of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 simulation at 3.00 Å RMSD. (Green: Cluster1; Blue: Cluster2; Pink: Cluster3)

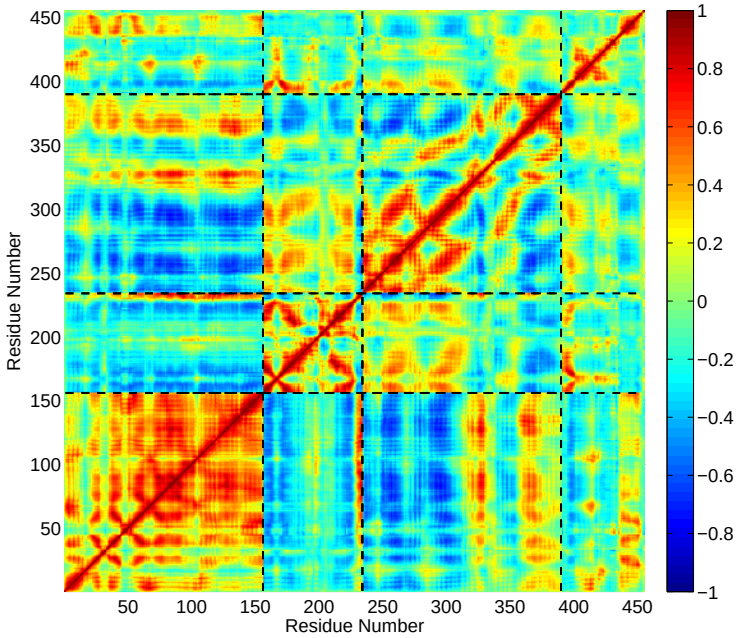

Figure 3.42. Cross Correlation of simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1; 391-455: RanBP2) calculated over the first cluster

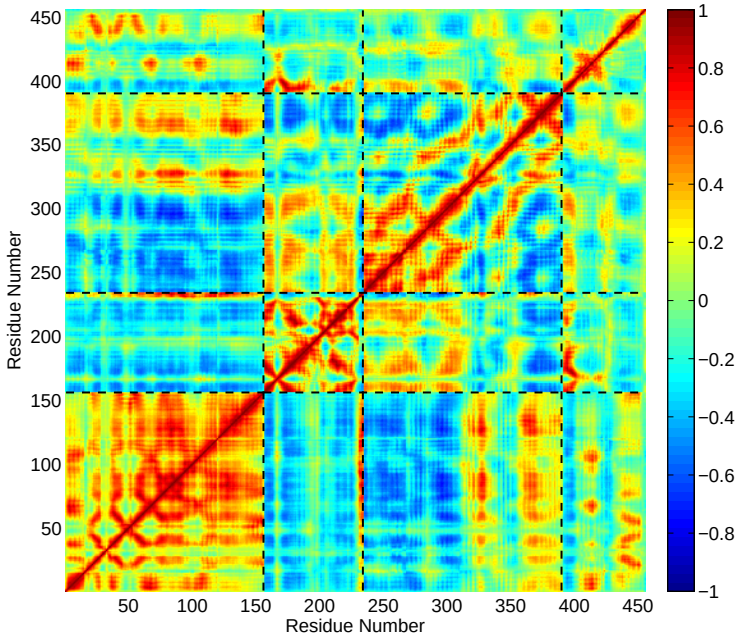

Figure 3.43. Cross Correlation of simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1; 391-455: RanBP2) calculated over the second cluster

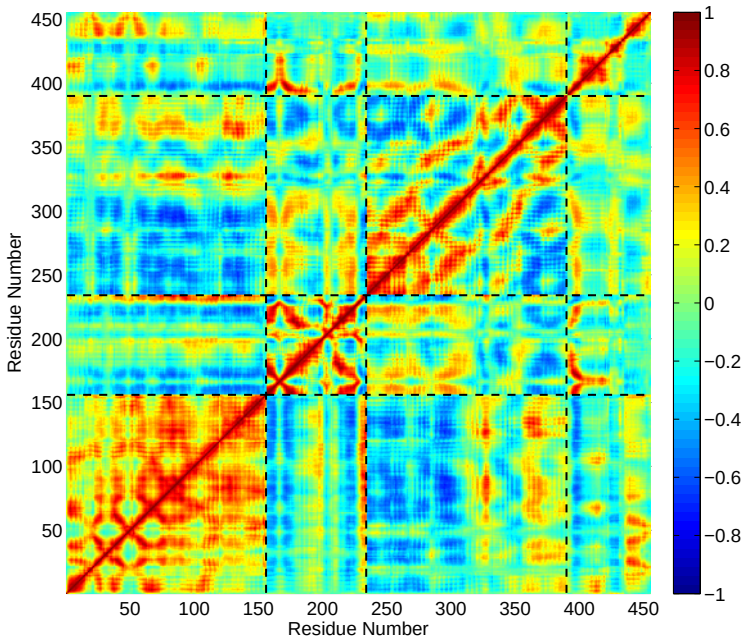

Figure 3.44. Cross Correlation of simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1; 391-455: RanBP2) calculated over the third cluster

threshold are displayed. The maximum change occurs at RanBP2 region as expected. Besides, N-terminal region of RanGAP1 show conformational change at Phe440-Lys445 region (residues 243-248). This region is found to be more mobile in the mutant in Section 3.2.2.

In Figure 3.42, 3.43 and 3.44 shows the cross correlations for the centroids of the clusters obtained through clustering of simulation of Asp33Ala mutant form of Ubc9-SUMO-RanGAP1-RanBP2 structure.

The Loop 2 region, which Asp33Ala, makes looser correlations with Thr511-Leu522, the SUMO binding motif: Leu523-Glu526 and Leu555-Pro566 (residues 314-325, 326-329 and 358-369) regions of RanGAP1 from cluster to cluster, having the third cluster the weakest correlations. As the correlations of Loop 2 decrease, the correlations of SUMO with Thr511-Leu522, Leu523-Glu526 and Leu555-Pro566 (residues 314-325, 326-329 and 358-369) of RanGAP1 decrease. Particularly, Tyr21-Leu24, Lys37-His43,

which resides at the RanBP2 interface, and the Gly68-Val90, to which Gly50 (residue 49) of Ubc9 anchors, of SUMO (residues 158-161, 174-180, 205-227) show high correlations with Phe440-Lys445, Ile456-Asp464 and Phe502-Asn508 regions of RanGAP1 (residues 243-248, 259-267 and 305-311).

Correlations of Ser61 (residue 198) with Phe440-Lys445, Ile456-Asp464 along with Phe502-Asn508 regions of RanGAP1 (residues 243-248, 259-267 and 305-311) regions have been lost through the trajectory. Additionally, Gly68 of SUMO with Thr511-Leu522, the SUMO binding motif, Leu523-Glu526 and Leu555-Pro566 (residues 314-325, 326-329 and 358-369) regions of RanGAP1 has faded throughout the trajectory.

3.2.5.2. Clustering of Ubc9-SUMO-RanGAP1-RanBP2 Simulations. In order to understand the extent of conformational shift caused by the mutation, all three simulations of (two parallel wild type simulations and an Asp33Ala mutant simulation) Ubc9-SUMO-RanGAP1-RanBP2 structure are merged and clustered at 3.50 Å RMSD threshold.

In Figure 3.45 the clustering profile for two wild type simulations and the Asp33Ala mutant simulation of Ubc9-SUMO-RanGAP1-RanBP2 is given. The clustering is based on the RMSD of Ubc9-SUMO-RanGAP1 regions only, excluding the RMSD of RanBP2 which is a flexible fragment. First two portions are comprised of conformations from the two wild type simulations, whereas the last portion of the graph (on the right) is composed of conformation created in the mutant simulation. As one can see, mutation not only affected coupling of chains with each other but the ensemble of conformations is shifted to different conformational ensemble on the selected rmsd measure.

In Figure 3.46, the representative members of the clusters obtained by clustering the simulation of two parallel wild type simulations and Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 structure with 3.00 Å RMSD threshold are given. The most distinctive conformational change is, as one can see, in the N-terminal region of RanGAP1 show conformational change at Phe440-Lys445, Ile456-Asp464 and Phe502-

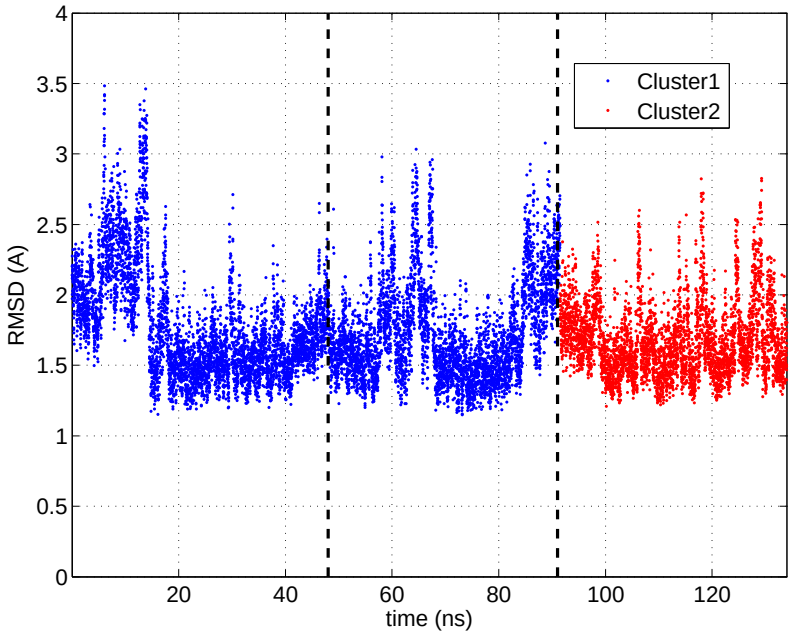

Figure 3.45. Clustering profile for two simulations of Ubc9-SUMO-RanGAP1-RanBP2 and simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 at 3.50 Å RMSD threshold (Left to Right: Wild type simulation 1, Wild type simulation 2, Mutant simulation)

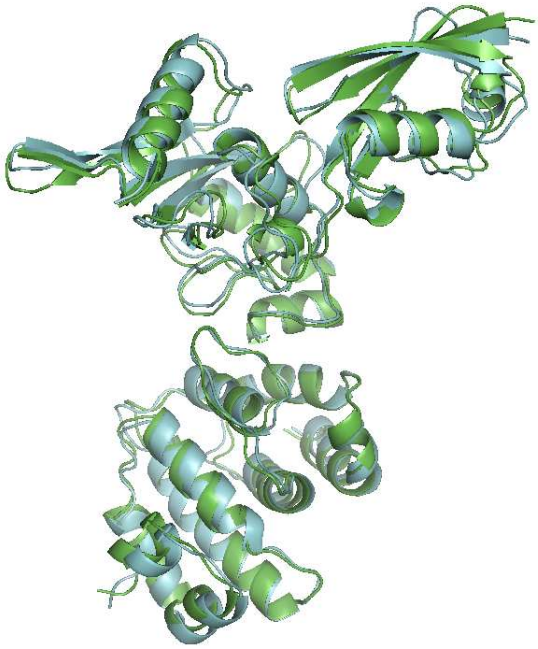

Figure 3.46. Representative members of the clusters formed by clustering of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 simulation and wild type simulations at 3.00 Å RMSD. (Green: Cluster1; Blue: Cluster2)

Asn508 regions (residues 243-248, 259-267 and 305-311).

3.2.5.3. Clustering of All Five Trajectories. To understand the extent of conformational change that the mutation caused, all three simulations of (two parallel wild type simulations and an Asp33Ala mutant simulation) Ubc9-SUMO-RanGAP1-RanBP2 structure and two parallel simulations of Ubc9-SUMO-RanGAP1 structure are merged and clustered at 3.50 Å RMSD threshold.

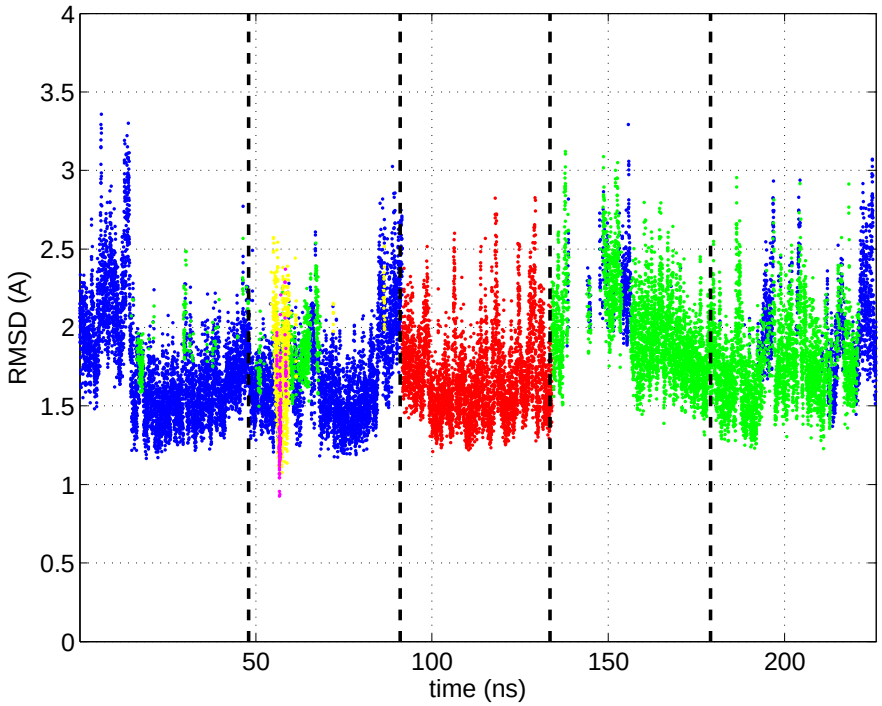

Figure 3.47. Clustering profile for two simulations of Ubc9-SUMO-RanGAP1-RanBP2, simulation of Asp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 and two simulations of Ubc9-SUMO-RanGAP1 trajectories at 3.50 Å RMSD threshold

The clustering results for two wild type simulations of Ubc9-SUMO-RanGAP1-RanBP2, the Asp33Ala mutant simulation of Ubc9-SUMO-RanGAP1-RanBP2 and two parallel simulations of the Ubc9-SUMO-RanGAP1 structure is given. First two portions are comprised of conformations from the two wild type simulations, the third region of the graph is composed of conformation created in the mutant simulation and the last two regions on the right is coming from the two parallel runs of Ubc9-SUMO-RanGAP1 structure. It is again obvious that the absence of RanBP2 distors

the stability of the conformation ensemble and the mutation results in a shift in the ensemble of conformations. Since the first and second clusters perfectly match those clusters shown in Figure 3.45, the representative members of the clusters also match to those presented in Figure 3.46.

4. CONCLUSIONS and FUTURE STUDIES

4.1. Conclusions

Sumoylation, the covalent attachment of SUMO (Small Ubiquitin Like Modifier) to target proteins is a posttranslational modification that can alter intracellular localization, interactions with other proteins or result in modifications by other posttranslational modifiers. The objective of this study is to understand the mechanism of sumoylation process just after the covalent attachment of SUMO. Ubc9-SUMO-RanGAP1-RanBP2 complex structure is chosen as the focus of this study. In order to understand the role of RanBP2, which is known as an E3 ligase for the sumoylation machinery, Ubc9-SUMO-RanGAP1-RanBP2 structure and Ubc9-SUMO-RanGAP1 complex, which contains the isopeptide bond between Gly97 of SUMO and 524Lys of RanGAP1, are simulated at 300 K with molecular dynamics method. In order to get coherent results, parallel simulations of both structures are designed to run. Moreover, inspired by the work by Karaca *et al.*, 2010 a mutation is design to further investigate the system. The presence of RanBP2 is found to be imposing a packing motion for the Loop 2 region, which is Lys30-Met36 region of Ubc9. On this loop Asp33 residue is found to maintain a unique conformational state, where Asp33 bends itself towards RanGAP1. (Karaca *et al.*, 2010) In order to clarify the role of Asp33 of Ubc9, Asp33Ala mutant is designed. In Section 4.1.1 the conclusions from comparisons of simulations of Ubc9-SUMO-RanGAP1-RanBP2 and Ubc9-SUMO-RanGAP1 structures are declared whereas in Section 4.1.2 the conclusions gained from the mutant simulations is given.

4.1.1. The Effect of RanBP2 Presence

4.1.1.1. Stabilization and Restriction of Conformational Space. Presence of RanBP2 does not reduce the mean RMSD values, however, stabilizes the fluctuations of Ubc9 and SUMO. Without RanBP2, SUMO is observed to reach conformations with high RMSD in which SUMO bends itself towards RanGAP1. In the presence of RanBP2, the rigid body motions of Ubc9 and SUMO avoided, but two loops of RanGAP1 (Phe440-

Lys445, Ile456-Asp464 and Phe502-Asn508 regions) bends itself towards SUMO. The MSF of RanGAP1 in the presence of RanBP2 is observed to be about 10% higher.

Both the clustering of trajectories at same RMSD thresholds and the decay of auto-correlations of the virtual bond vectors supported the idea that RanBP2 restricts the conformational space of Ubc9-SUMO-RanGAP1 structure via stabilizing the fluctuations of Ubc9 and SUMO. Without RanBP2, Ubc9 and SUMO shows larger decays in auto-correlations of the virtual bond vectors which imply that these two chains are more mobile and lose their memory faster compared to case that RanBP2 present. SUMO and Ubc9 shows decreased stability without RanBP2, however, two loops regions, Phe440-Lys445, Ile456-Asp464 and Phe502-Asn508 of RanGAP1 display decreased autocorrelations in the presence of RanBP2.

Clustering, on the other hand, gives information of the distribution of generated conformations on an RMSD-based algorithm. Even including the RMSD of the flexible RanBP2 fragment, in Ubc9-SUMO-RanGAP1-RanBP2 fewer clusters form compared to Ubc9-SUMO-RanGAP1 case. This alone stands as evidence for the fact that RanBP2 restricts the conformational space of Ubc9-SUMO-RanGAP1 structure. Furthermore, in the absence of RanBP2 revisiting of former conformational groups (clusters in which conformations are distinguished via clustering) are observed. Additionally, simultaneous clustering of cases with and without RanBP2 revealed that, presence of RanBP2 provides generation of conformation that are closer to (on RMSD bases) the crystal structure. About 99% of the conformations sampled in Ubc9-SUMO-RanGAP1-RanBP2 simulations belong to the first cluster which has 1.53 Å RMSD to the starting structure. However, this percentage is reduced to 15% in Ubc9-SUMO-RanGAP1 simulations. This implies not only that the system makes higher conformational changes, but it implies lack of conformational stability in the absence of RanBP2.

All in all, combinining these findings with previous work (Pichler *et al.*, 2002; Reverter and Lima, 2005, Karaca *et al.*, 2010, Tozluoglu *et al.*, 2010) RanBP2 to limit the available conformations/configurations of Ubc9-SUMO-RanGAP1 complex

by providing the optimum orientation and conformation of Ubc9 and SUMO proteins.

4.1.1.2. Change in Couplings and Importance of Loop2. Loop 2 (Lys30-Met36) is important due to that fact that the some of the RanBP2 binding sites (Val27-Glu42) are located on this flexible loop (Reverter and Lima, 2005). Moreover, it is highly mobile and suggested to have functional role in trasmission of the effect of RanBP2 (Karaca *et al.*, 2010). Since this loop is highly associated with the RanBP2, it has no doubt that changes in its coupling will be observed in the absence of RanBP2.

Ubc9 and SUMO interface is still coupled in their fluctuations in the absence of RanBP2, yet in the presence of RanBP2 those correlations are not as strong and different residues in Ubc9 and SUMO show association. In the absence of RanBP2, Ubc9 rearranges its cooperative fluctuations in which new correlations rise with Ser61 and Arg70 residues of SUMO whereas Ubc9:RanGAP1 coupling is weakened. It can be concluded that, the presence of RanBP2 weakens the coupling Ubc9:RanGAP1 and reorganizes the Ubc9-SUMO associations.

In the presence of RanBP2, Loop 2 is positively correlated with Thr511-Leu522, the LKSE motif between Leu523-Glu526 and Leu555-Pro566 regions of RanGAP1. Thr511-Leu522 and Leu555-Pro566 regions are functionally important on the grounds of the fact that they face the built-in E3 region of Ubc9 (Ala131-Arg141) (Bernier-Villamor *et al.*, 2002). Without RanBP2, the catalytic Cys97 and the surrounding loop between Lys74-Phe77 of Ubc9 are less correlated with Thr511-Leu522, the LKSE motif between Leu523-Glu526 and Leu555-Pro566 regions of RanGAP1. Moreover, Ala131-Gln139 region of Ubc9, which is known as the built-in E3 of Ubc9, is strongly correlated with Thr511-Leu522 and Leu555-Pro566 of RanGAP1. It can be deducted that, in the absence of an actual E3 ligase, the built-in E3 region of Ubc9 tries to exert the same allosteric effect by maintaining the correlations with the specified regions of RanGAP1.

It is interesting to note that, increasing correlations of Loop 2 with those regions

on RanGAP1 (Thr511-Leu522, the LKSE motif between Leu523-Glu526 and Leu555-Pro566) leads to increased correlations of SUMO with the RanGAP1. The coupling of Tyr21-Leu24, Lys37-His43, Gly68-Val90 of SUMO with the two loops of RanGAP1 (Phe440-Lys445, Ile456-Asp464 and Phe502-Asn508 regions) strengthens whenever the Loop 2 correlations with RanGAP1 increase.

Correlations of fluctuations suggest that, in the presence of RanBP2, Loop 2 correlations with RanGAP1 are provided by the interactions of RanBP2 and Loop 2 of Ubc9. In the absence of RanBP2, the built-in E3 region of Ubc9 makes couplings with Thr511-Leu522 and Leu555-Pro566 regions of RanGAP1. The correlations of Loop2 gradually increase along the course of simulation as built-in E3 and RanGAP1 correlations gets stronger. At that point, a sequence of events in the signal transmission can be suggested. If RanBP2 is present, Loop 2:RanBP2 and SUMO (N-terminal region):RanBP2 correlations are provided. If RanBP2 is not present, built-in E3 site of Ubc9 (Ala131-Arg141) takes the control, makes coupling with Thr511-Leu522 and Leu555-Pro566 regions of RanGAP1. In either way, coupling of Loop 2 with Thr511-Leu522, the LKSE motif between Leu523-Glu526 and Leu555-Pro566 regions of RanGAP1 are observed. Whenever, this coupling is achieved, correlations of Tyr21-Leu24, Lys37-His43, Gly68-Val90 of SUMO with the N-terminal domain (particularly, Phe440-Lys445, Ile456-Asp464, Phe502-Asn508 regions) of RanGAP1 increase. Strong correlations mean stronger interactions with SUMO and RanGAP1. These coupling are accompanied by coupling of Gly50 of Ubc9 with Gly68 of SUMO. At that point the role of those resides on the signaling pathway cannot be determined, requires further focus.

4.1.2. The Role of Asp33 of Ubc9

4.1.2.1. Effect of Mutation on Mobility. The mutation is not expressed on the RMSD values, however, RanGAP1 is observed to make abrupt changes in RMSD which is mainly caused by the rigid body motion towards SUMO. This motion was observed in the wild type simulation however with lower RMSD. Although, RanGAP1 approaches towards SUMO, it is still not feasible for sumoylation to occur as Ubc9 and SUMO

approaches to each other too.

The mutation did not affect the mobility of Ubc9 itself but MSF revealed that Asp30, Lys37-His43 and Gly68-Glu84 of SUMO show decreased mobility in the mutant structure. Asp30 is located at the Ubc9 interface and Lys37-His43 region resides at the RanBP2 interface. Gly68-Glu84 of SUMO is nothing but the loop that anchors Gly50 of Ubc9 in wild type simulations. Those regions of SUMO display decreased mobility and rigidifies with the Asp33Ala mutation. Some regions of RanGAP1 have no reported significance, however, considering the MSF and couplings of the cross correlation maps, they might be on the transmission pathway of the signal from Asp33. For instance, Phe440-Lys445, Ile456-Asp464, Phe502-Asn508 and Lys528-Tyr550 regions of RanGAP1 is found to be less mobile without RanBP2. Presence of mutation mimicked the RanBP2-absent case, implying that allosteric signaling caused by the RanBP2 is blocked.

4.1.2.2. Effect of Mutations on Couplings. In simulation of mutant structure, Loop 2 makes looser correlations with Thr511-Leu522, the SUMO binding motif: Leu523-Glu526 and Leu555-Pro566 regions of RanGAP1 compared to wild type simulations. The built in E3 region of Ubc9 (Ala131-Arg14) show weaker correlations with the Thr511-Leu522, Leu523-Glu526 and Leu555-Pro566 region of RanGAP1. This implies that, even if the actual RanBP2 is present along with the built-in E3 region of Ubc9, the desired couplings of Loop 2 cannot be achieved in the mutant case. It is also worth noting that, correlations of Asp100-Ala101 of Ubc9, which are responsible for target recognition, display only weak correlation with the SUMO binding motif (Leu523-Glu526 of RanGAP1) implying that recognition of RanGAP1 might be impaired thus blocking the sumoylation process.

4.1.2.3. Effect of Mutation on the Conformational Ensemble. Individual and simultaneous clustering of the simulation of the Aspp33Ala mutant of Ubc9-SUMO-RanGAP1-RanBP2 revealed that mutation not only affected coupling of chains with each other but the ensemble of conformations is shifted to a different conformational ensemble on

the selected RMSD measure.

Inspection of representative members of the clusters showed that N-terminal region of RanGAP1 show conformational change at Phe440-Lys445, Ile456-Asp464 and Phe502-Asn508 region. Those regions have no reported significance however worth further investigation due to the fact that they are mobile upon RaBP2 presence and loose mobility with the introduction of mutation mimicking the RanBP2 absent case.

4.1.3. Future Studies

There are over 300 publications on SUMO and sumoylation however the exact mechanism of sumoylation or the exact pathway that Ubc9, SUMO, RanGAP1 and RanBP2 communicate with each other is still not clear.

Under the light of this work, mutations that impair coupling of Gly68 of SUMO with Gly50 of Ubc9 can be carried out on the grounds of the fact that this coupling is observed to be associated with signaling pathway of the allosteric effect of RanBP2.

Moreover, Phe440-Lys445, Ile456-Asp464 and Phe502-Asn508 regions of RanGAP1 worth investigation as their mobility is highly dependent on RanBP2.

APPENDIX A: Results for Parallel Run of Ubc9-SUMO-RanGAP1-RanBP2 Complex Structure

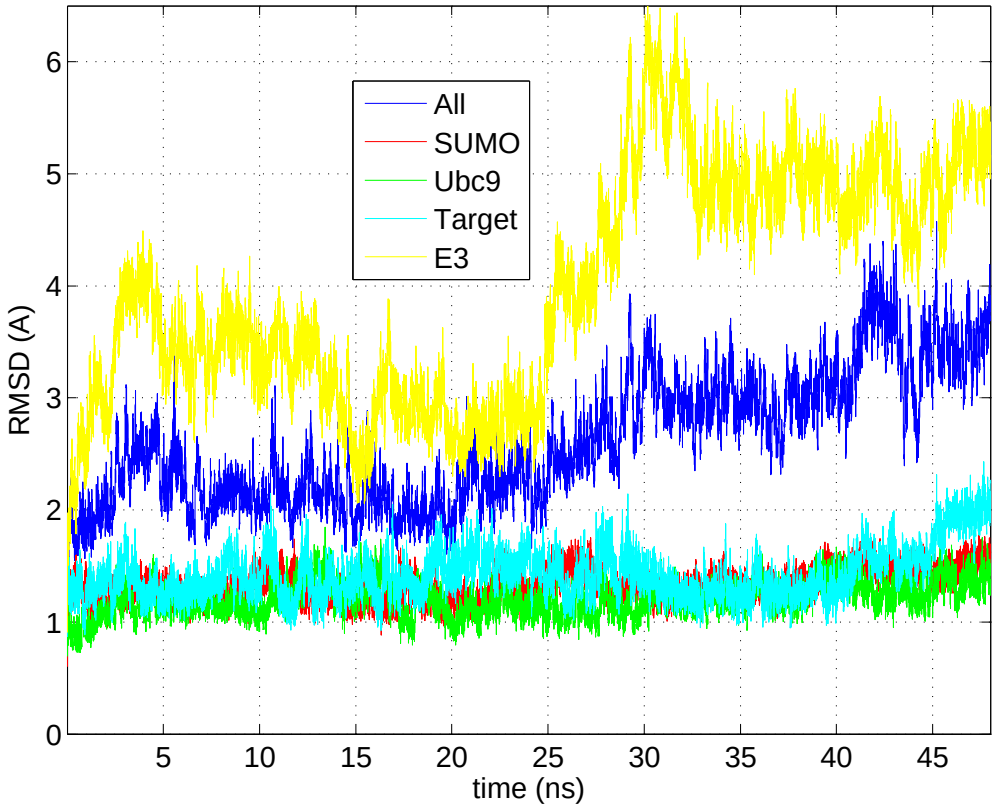

Figure A.1. RMSD of parallel run of Ubc9-SUMO-RanGAP1-RanBP2

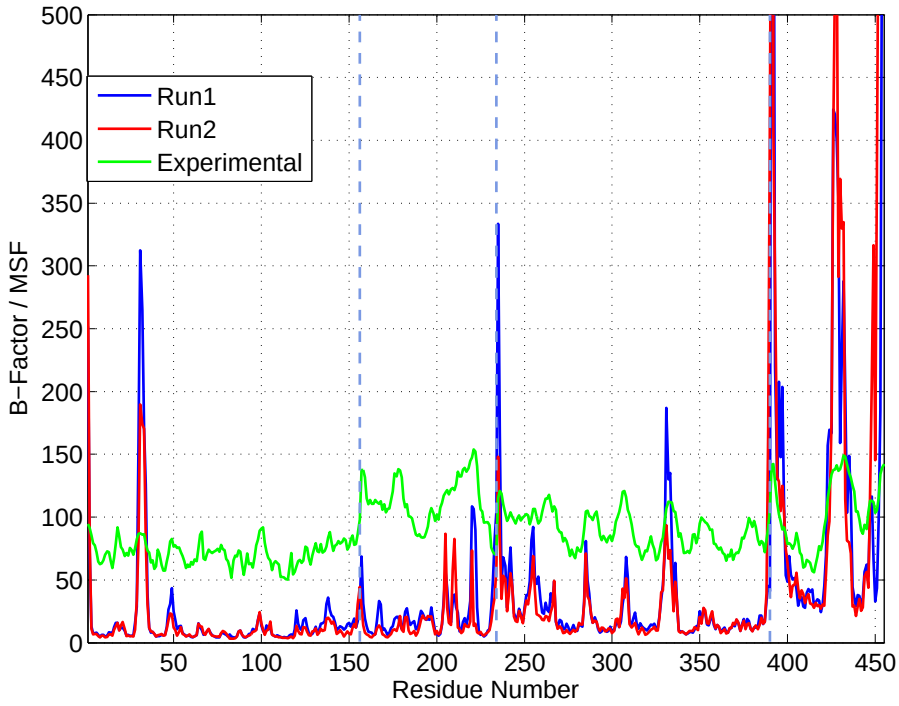

Figure A.2. B-factors of parallel run of Ubc9-SUMO-RanGAP1-RanBP2 via individual alignment (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1, 391-455: RanBP2)

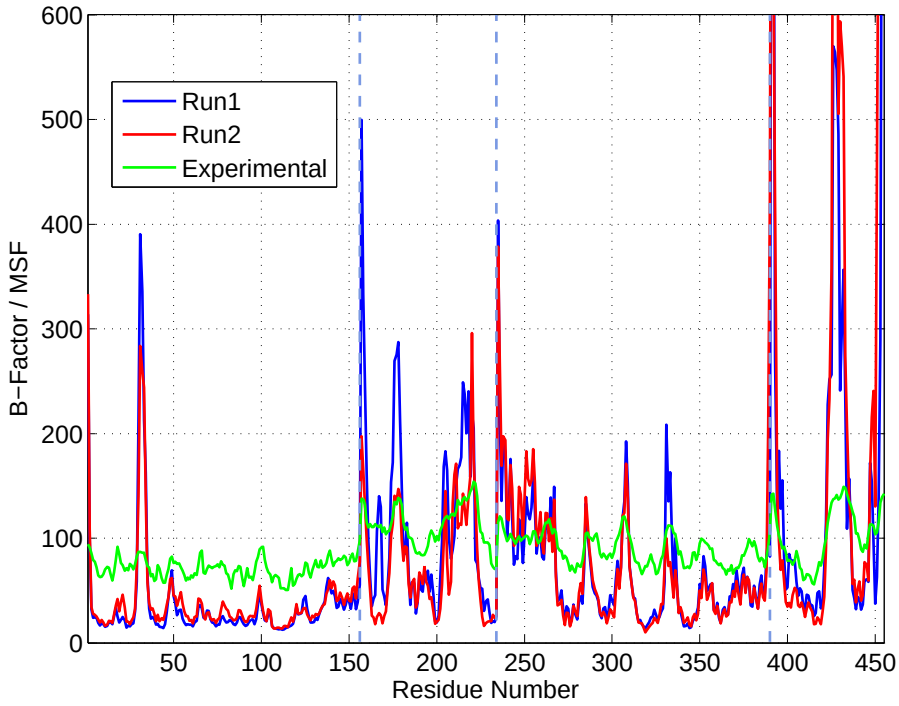

Figure A.3. B-factors of parallel run of Ubc9-SUMO-RanGAP1-RanBP2 via overall alignment (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1, 391-455: RanBP2)

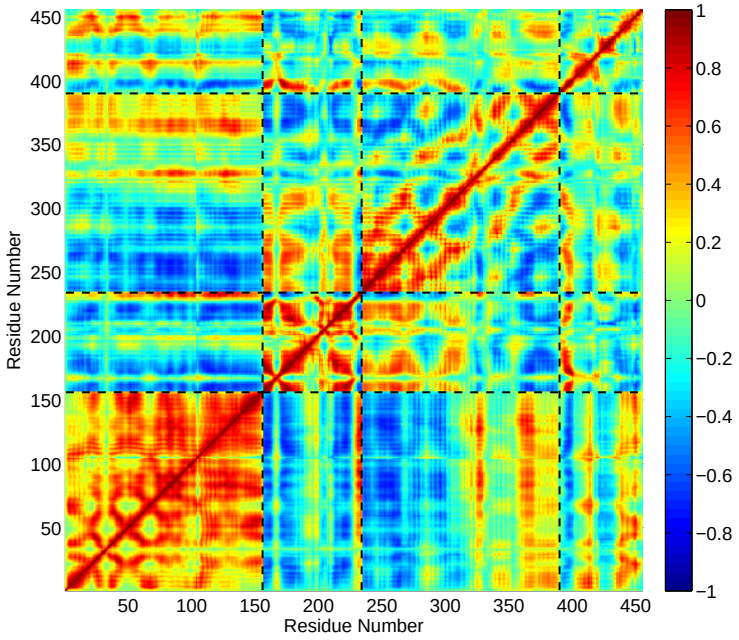

Figure A.4. Cross Correlations of parallel run of Ubc9-SUMO-RanGAP1-RanBP2 tetramer complex from the initial minimized structure (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1, 391-455: RanBP2)

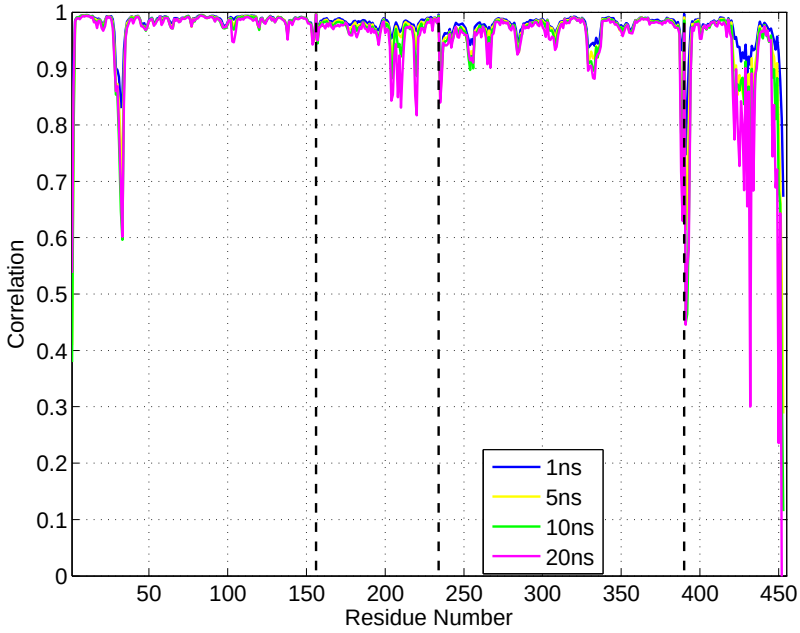

Figure A.5. Decay of Orientational Vectors from parallel run of Ubc9-SUMO-RanGAP1-RanBP2 tetramer complex from the initial minimized structure (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1; 391-455: RanBP2)

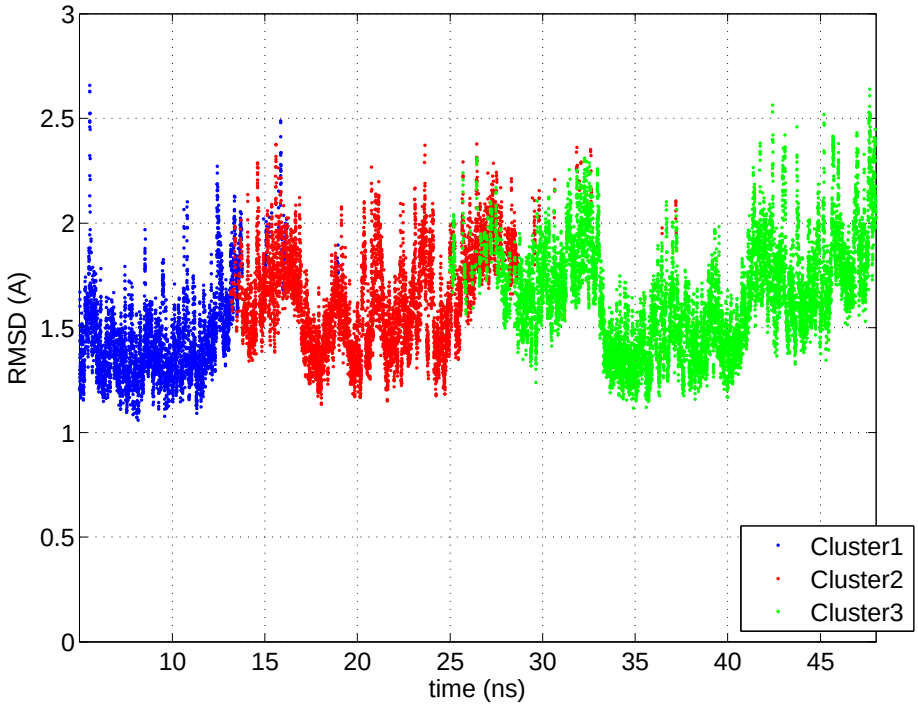

Figure A.6. Clustering profile for parallel run of Ubc9-SUMO-RanGAP1-RanBP2 structures at 3.00 Å RMSD as Threshold

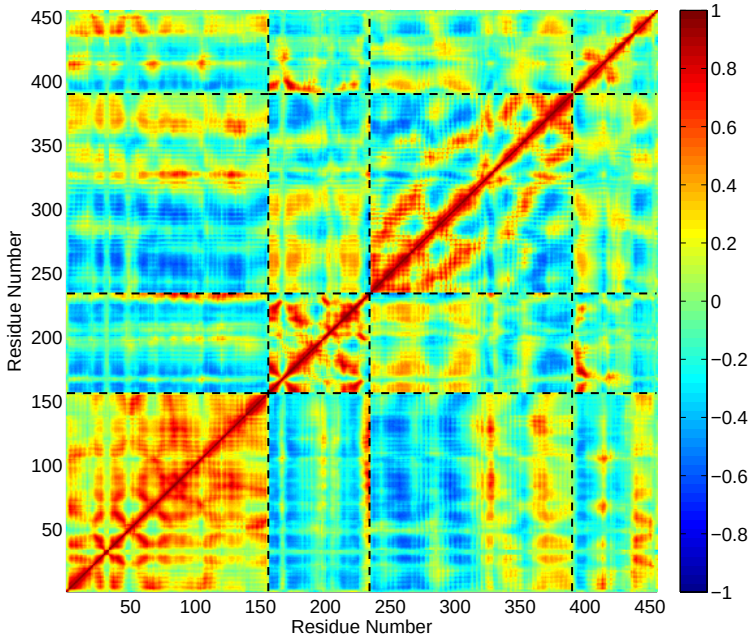

Figure A.7. Cross Correlations of parallel run of Ubc9-SUMO-RanGAP1-RanBP2 (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1, 391-455: RanBP2) from the first cluster

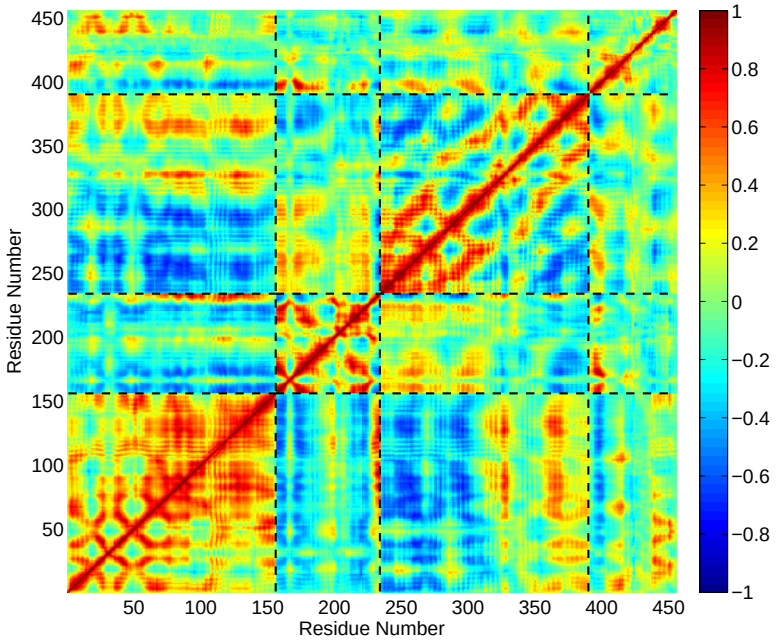

Figure A.8. Cross Correlations of parallel run of Ubc9-SUMO-RanGAP1-RanBP2 (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1, 391-455: RanBP2) from the second cluster

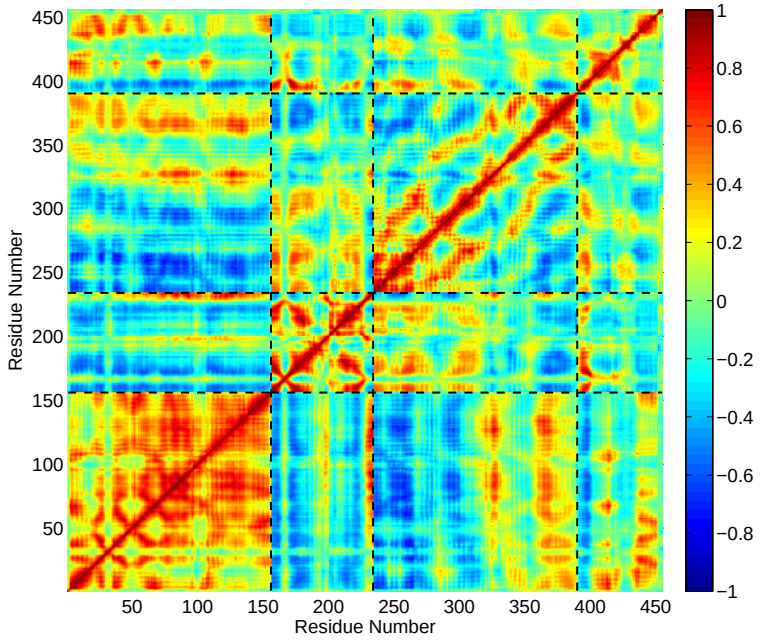

Figure A.9. Cross Correlations of parallel run of Ubc9-SUMO-RanGAP1-RanBP2 (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1, 391-455: RanBP2) from the third cluster

APPENDIX B: Results for Parallel Run of Ubc9-SUMO-RanGAP1 Complex Structure

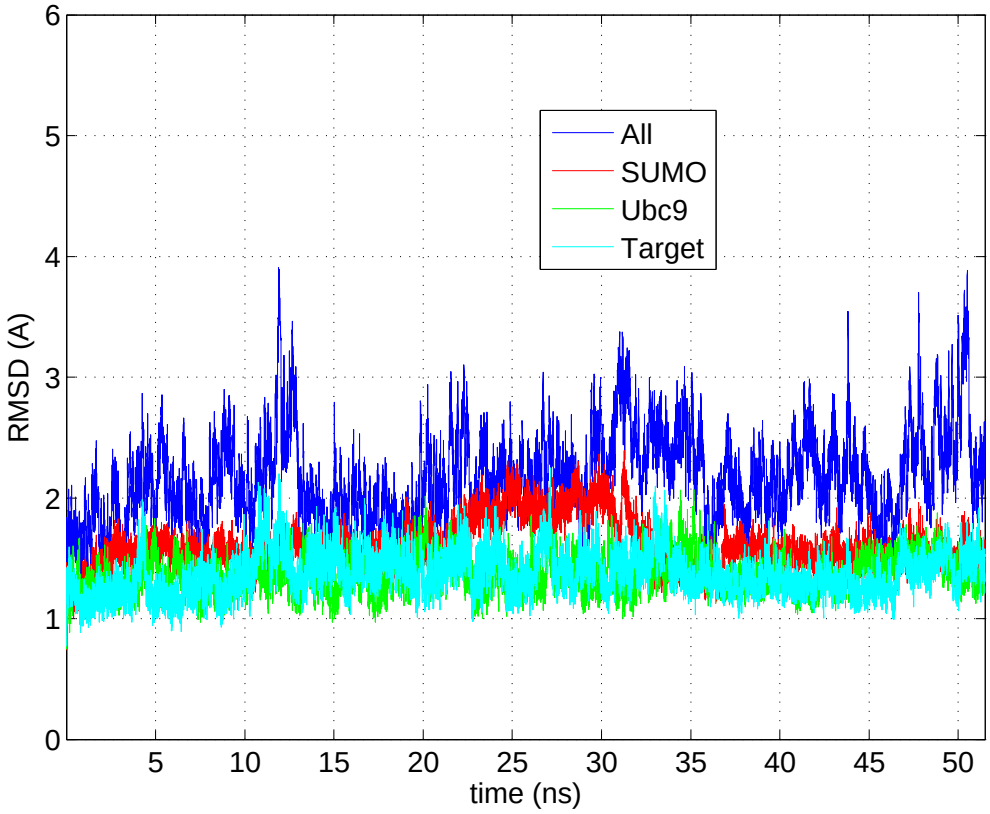

Figure B.1. RMSD of parallel run of Ubc9-SUMO-RanGAP1-RanBP2

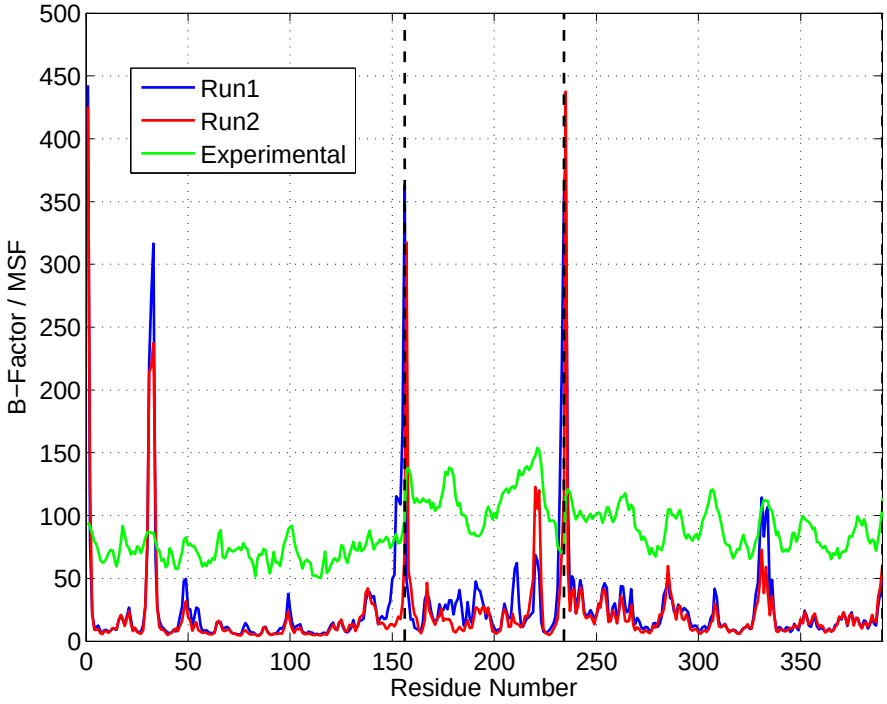

Figure B.2. B-factors of parallel run of Ubc9-SUMO-RanGAP1 via individual alignment (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1)

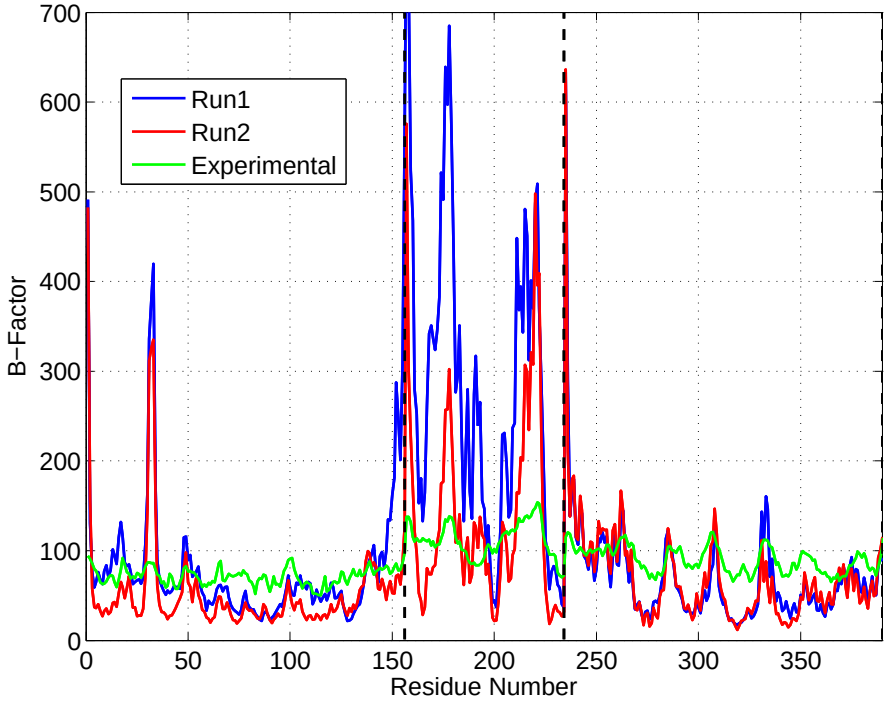

Figure B.3. B-factors of parallel run of Ubc9-SUMO-RanGAP1 via overall alignment (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1)

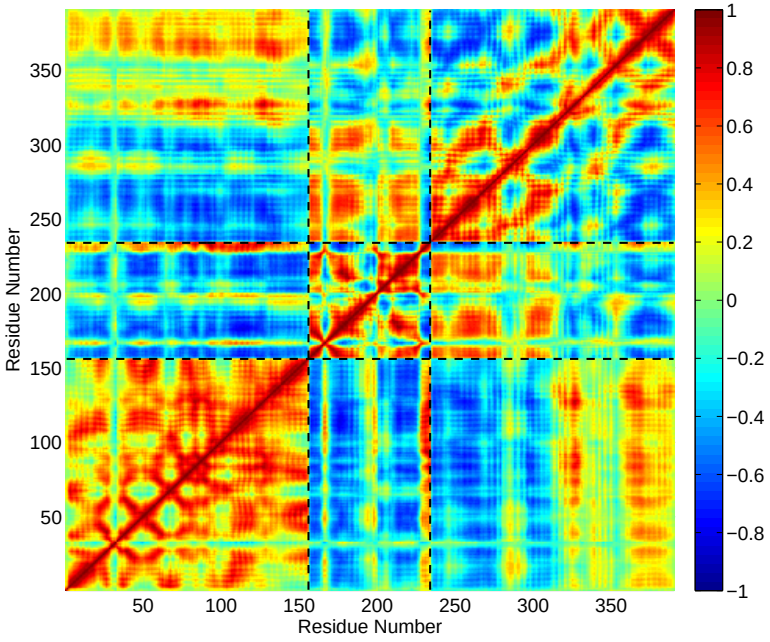

Figure B.4. Cross Correlations of parallel run of Ubc9-SUMO-RanGAP1 tetramer complex from the initial minimized structure (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1)

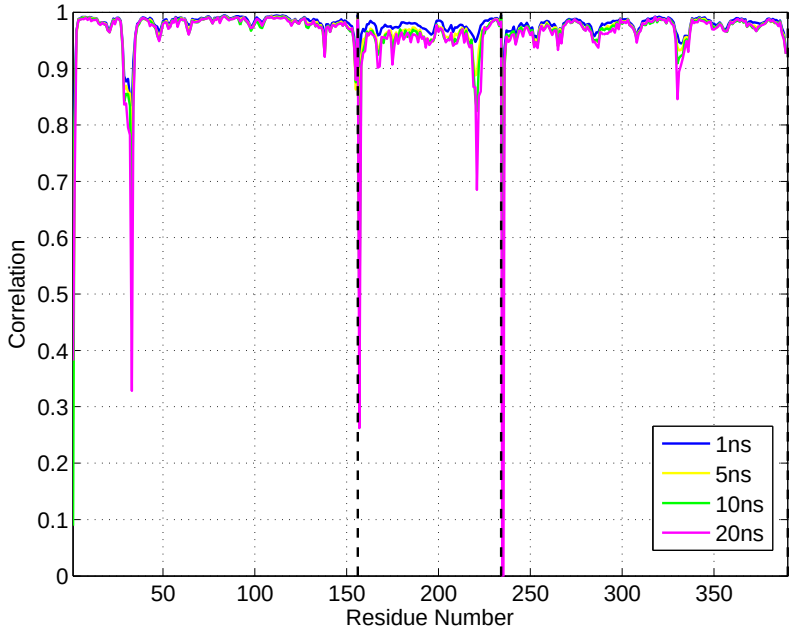

Figure B.5. Decay of Orientational Vectors from parallel run of Ubc9-SUMO-RanGAP1 tetramer complex from the initial minimized structure (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1)

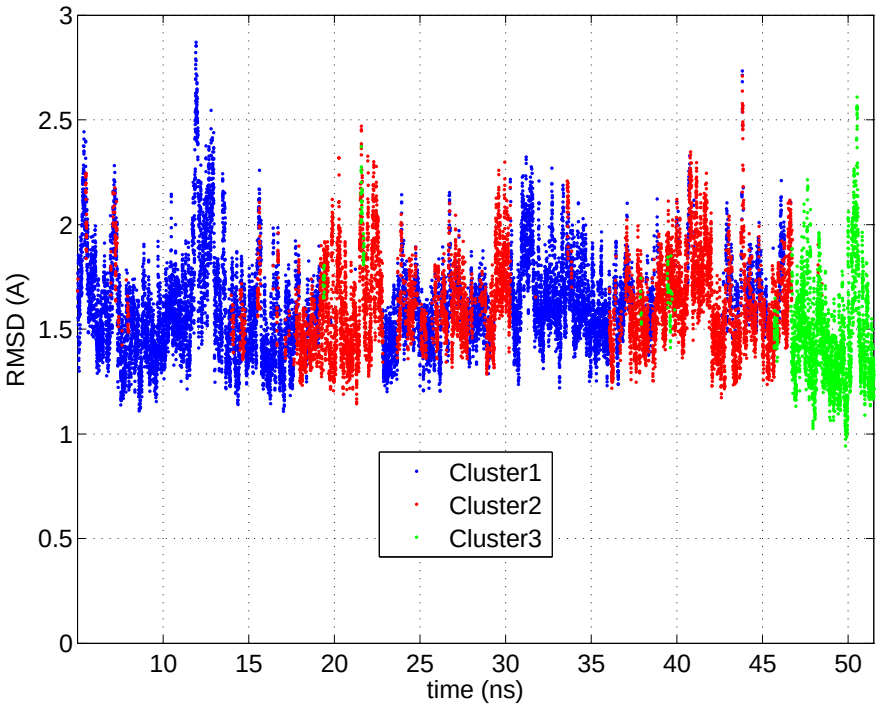

Figure B.6. Clustering profile for parallel run of Ubc9-SUMO-RanGAP1 structure at 3.00 Å RMSD as Threshold

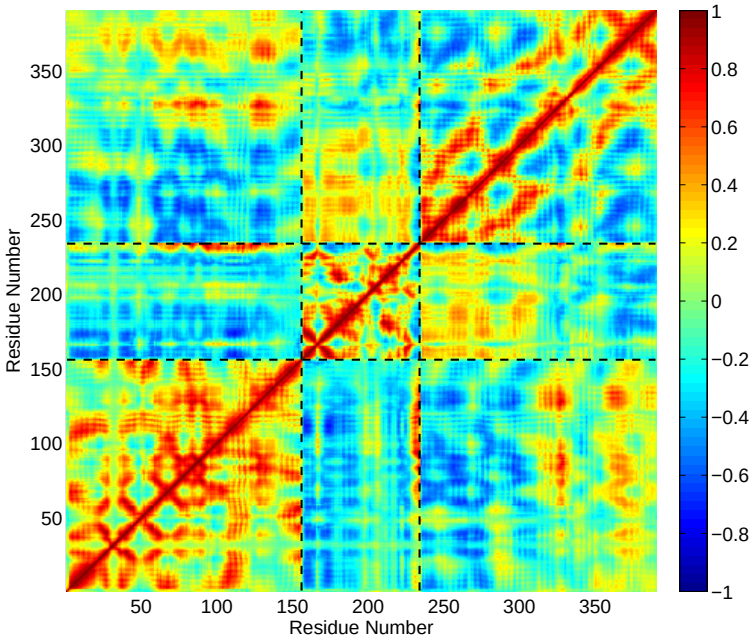

Figure B.7. Cross Correlations of parallel run of Ubc9-SUMO-RanGAP1 (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1) from the first cluster

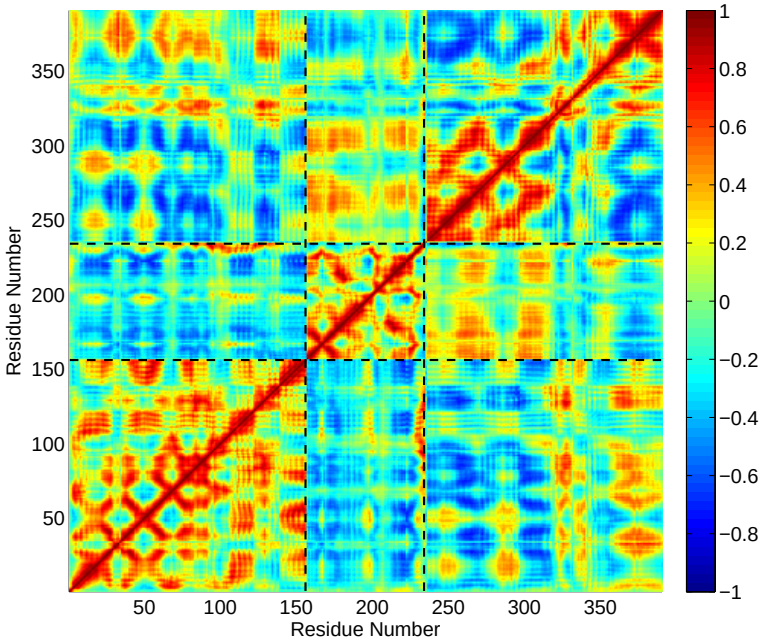

Figure B.8. Cross Correlations of parallel run of Ubc9-SUMO-RanGAP1 (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1) from the second cluster

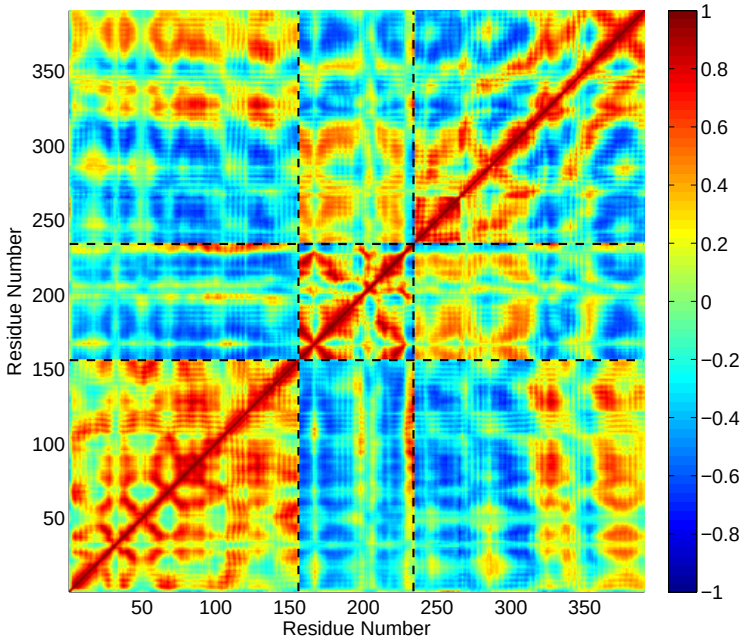

Figure B.9. Cross Correlations of parallel run of Ubc9-SUMO-RanGAP1 (1-156: Ubc9; 157-234: SUMO; 235-390: RanGAP1) from the third cluster

APPENDIX C: Summary of Significant Sites

Table C.1. Summary of functional residues in Ubc9

Site	Description
Val27-Glu42	The loop between beta-sheets, RanBP2 binding site (Reverter and Lima, 2005)
Loop 2, Lys30-Met36	Highly mobile loop, has functional role in trasmission of the effect of RanBP2 (Karaca *et al.*, 2010; Tozluoglu *et al.*, 2010)
Lys74-Phe77	Catalytic Loop that holds catalytic Cys93 (Bernier Villamor *et al.*, 2001)
His83-Asn85	HPN motif (Bernier Villamor *et al.*, 2001;Yunus and Lima, 2006; Wu *et al.*, 2003)
His83-Ser89	Help orientation of the Gly97 of SUMO (Bernier Villamor *et al.*, 2001; Yunus and Lima, 2006; Wu *et al.*, 2003)
Asn124-Pro128	Loop region which is contact with the SUMO binding motif in Target (Bernier Villamor *et al.*, 2001)
Cys93	Catalytic Cysteine (Melchior, 2000; Bernier Villamor *et al.*, 2001; Tang *et al.*, 2009)
Asp100-Lys101	Responsibles for Target recognition (Tatham *et al.*, 2003; Yunus and Lima, 2006)
Ala131-Arg141	Built-in E3 and RanGAP1 specific target binding site (Hochstreasser 2002; Martin *et al.*, 2007; Bernier Villamor *et al.*, 2002; Andrea *et al.*, 2004; Yunus and Lima, 2006; Wu *et al.*, 2003

Table C.2. Summary of functional residues in SUMO

Site	Description
Gly97	Catalytic site (Hochstrasser, 2002; Nalepa *et al.*, 2006; Kirkin and Dikic, 2007; Watson and Irwin, 2006; Hoeller *et al.*, 2006; Dye and Schulman, 2007; Capilli and Lima, 2007)

Table C.3. Summary of functional residues in RanGAP1

Site	Description
Thr511-Leu522	Faces Built-in E3 of Ubc9 (Bernier Villamor *et al.*, 2002)
Leu523-Glu526	Conserved LKSE motif for Sumoylation (Bernier Villamor *et al.*, 2002, reverter and Lima, 2005; Duda and Schulman, 2005; Heun, 2007; Geiss-Friedlander and Melchior, 2007)
Leu555-Pro566	Ubc9 binding site, Faces Built-in E3 of Ubc9 (Bernier Villamor *et al.*, 2002)

REFERENCES

Alarcon-Vargas, D., Z. Ronai, 2002, "SUMO in Cancer Wrestlers Wanted", *Cancer Biology & Therapy*, Vol. 1:3, pp. 237-242.

Berendsen, H. J. C., J. P. M. Postma, W. F. Van Gunsteren, A. DiNola and J. R. Haak, 1984, "Molecular dynamics with coupling to an external bath", *Journal of Chemical Physics*, Vol. 81, pp. 3684-3690.

Bernier-Villamor, V., D. A. Sampson, M. J. Matunis and C. D. Lima, 2002, "Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin conjugating enzyme Ubc9 and RanGAP1", *Cell*, Vol. 108, pp. 345-356.

Boddy, M. N., K. Howe, L. D. Etkin, E. Solomon and P. S. Freemont, 1996, "PIC 1, a novel ubiquitin-like protein which interacts with the PML component of a multiprotein complex that is disrupted in acute promyelocytic leukaemia", *Oncogene*, Vol. 13, pp. 971-982.

Capili, A. D. and C. D. Lima, 2007, "Structure and Analysis of a Complex between SUMO and Ubc9 Illustrates Features of a Conserved E2-Ubl Interaction", *Journal of Molecular Biology*, Vol. 369, pp. 608-618.

Capili, A. D. and C. D. Lima, 2007, "Taking it Step by Step: Mechanistic insights from structural studies of Ubiquitin/Ubiquitin-like protein modification pathways", *Current Opinion in Structural Biology*, Vol. 17, pp. 726-735.

Case, D. A., T. A. Darden, T. E. I. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, B. Wang, D. A. Pearlman et al., 2004, "AMBER 8", *Science*, University of California, San Francisco, 2004.

Case, D. A., T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Jr. Merz, A. Onufriev, C. Simmerling, B. Wang and R. Woods, 2005, "The Amber biomolecular

simulation programs", *Journal of Computational Chemistry*, Vol. 26, pp. 1668-1688.

Cheng, J., T. Bawa, P. Lee, L. Gong and T. H. Yeh, 2006, "Role of Desumoylation in the Development of Prostate Cancer", *Neoplasia*, Vol. 8, No. 8, pp. 667-676.

Desterro, J. M. P., J. Thomson and R. T. Hay, 1997, "Ubch9 conjugates SUMO but not ubiquitin", *FEBS Letters*, Vol. 417, pp. 297-300.

Desterro, J. M. P., M. S. Rodriguez, G. D. Kemp, and R. T. Hay, 1999, "Identification of the Enzyme Required for Activation of the Small Ubiquitin-like Protein SUMO-1", *The Journal of Biological Chemistry*, Vol. 274, No. 15, pp. 10618-10624.

Dohmen, R. J., 2004, "SUMO protein modification", *Biochimica et Biophysica Acta*, Vol. 1695, pp. 113-131.

Duda, D. M., B. A. Schulman, 2005, "Tag-Team SUMO Wrestling", *Molecular Cell*, Vol. 18, pp. 612-614.

Feig, M. J., J. Karanicolas and C. L. Brooks, 2004, "MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology", *Journal of Molecular Graphics and Modelling*, Vol.22, pp. 377-395.

Ferrier, V., 2002, "Getting hit by SUMO", *Nature Cell Biology*, Vol. 4., pp. E57.

Flory, P. J., 1969, *"Statistical mechanics of chain molecules"*, John Wiley & Sons Inc.

Geiss-Friedlander, R. and F. Melchior, 2007, "Concepts in sumoylation: a decade on", *Nature Reviews Molecular Cell Biology*, Vol. 8, pp. 947-956.

Gill, G., 2003, "Post-translational modification by the small ubiquitin-related modifier SUMO has big effects on transcription factor activity", *Current Opinion*, Vol. 13, pp. 108-113.

Gill, G., 2004, "SUMO and ubiquitin in the nucleus: different functions, similar mech-

anisms?", *Genes Dev.*, Vol. 18, pp. 2046-2059.

Giraud, M. F., J. M. Desterro, J. H. Naismith, 1998, "Structure of ubiquitin conjugating enzyme 9 displays significant differences with other ubiquitin conjugating enzymes which may reflect its specificity for sumo rather than ubiquitin", *Acta. Crystallogr. D. Biol. Crystallogr.*, Vol. 54, pp. 891-898

Girdwood, D. W. H., M. H. Tatham, R. T. Hay, 2004, "SUMO and transcriptional regulation", *Seminars in Cell & Developmental Biology*, Vol. 15, pp. 201-210.

Gong, L., T. Kamitani, K. Fujise, L. S. Caskey and E. T. H. Yeh, 1997, "Preferential Interaction of Sentrin with a Ubiquitin-conjugating Enzyme, Ubc9", *The Journal of Biological Chemistry*, Vol. 272, No. 45, pp. 28198-28201.

Gorlich, D., I. W. Mattaj, 1996, "Nucleocytoplasmic Transport ", *Science*, Vol. 271. no. 5255, pp. 1513-1519.

Hay, R. T., 2001, "Protein modification by SUMO", *TRENDS in Biochemical Sciences*, Vol. 26 No. 5, pp. 332-333.

Hecker, C. M., M. Rabiller, K. Haglund, P. Bayer and I. Dikic, 2006, "Specification of SUMO1- and SUMO2-interacting Motifs", *The Journal of Biological Chemistry*, Vol. 281, NO. 23, pp. 16117-16127.

Herrmann, J., L. O. Lerman and A. Lerman, 2007, "Ubiquitin and Ubiquitin-Like Proteins in Protein Regulation", *Journal of American Heart Association*, Vol. 100, pp. 1276-1291.

Hershko, A., H. Heller, S. Elias and A. Ciechanover, 1983, "Components of Ubiquitin-Protein Ligase System", *The Journal of Biological Chemistry*, Vol. 258, No. 13, pp. 8206-8214.

Heun, P., 2007, "SUMOrganization of the nucleus", *Curr Opin Cell Biol.*, Vol. 19, pp. 350-355.

Hilgarth, R. S., L. A. Murphy, H. S. Skaggs, D. C. Wilkerson, H. Xing and K. D. Sarge, 2004, "Regulation and Function of SUMO Modification", *The Journal of Biological Chemistry*, Vol. 279, No. 52, pp. 53899-53902.

Hochstrasser, M., 2002, "New structural clues to substrate specificity in the ubiquitin system", *Molecular Cell*, Vol. 9, pp. 453-454.

Iniguez-Lluhi, J. A., 2006, "For a Healthy Histone Code, a Little SUMO in the Tail Keeps the Acetyl Away", *Acs Chemical Biology*, Vol. 1, pp. 67-74.

Jacquiau, H. R., R. C. A. M. V. Waardenburg, R. J. D. Reid, M. H. Woo, H. Guo, E. S. Johnson and M. A. Bjornsti, 2005, "Defects in SUMO (Small Ubiquitin-related Modifier) Conjugation and Deconjugation Alter Cell Sensitivity to DNA Topoisomerase I-induced DNA Damage", *The Journal of Biological Chemistry*, Vol. 280, No. 25, pp. 23566-23575

Jentsch, S., G. Pyrowolakis, 2000, "Ubiquitin and its kin: how close are the family ties?", *Trends In Cell Biology*, Vol. 10, pp. 335-342.

Johnson, E. S., G. Blobel, 1997, "Ubc9p Is the Conjugating Enzyme for the Ubiquitin-like Protein Smt3p", *The Journal Of Biological Chemistry*, Vol. 272, pp. 26799-26802.

Johnson, E. S., 2004, "Protein modification by SUMO", *Annual Review of Biochemistry.*, Vol. 73, pp. 355-382.

Jorgensen, W. L., J. Chandrasekhar, J.D. Madura, R.W. Impey and M.L. Klein, 1983,"Comparison of simple potential functions for simulating liquid water", *Journal of Chemical Physics*, Vol. 79, pp. 926-935.

Kamitani, T., H. P. Nguyen and E. T. H. Yeh, 1997, "Preferential Modification of Nuclear Proteins by a Novel Ubiquitin-like Molecule", *The Journal Of Biological Chemistry*, Vol. 272, pp. 14001-14004.

Karamouzis, M. V., P. A. Konstantinopoulos, F. A. Badra, A. G. Papavassiliou, 2008, "SUMO and estrogen receptors in breast cancer", *Department of Biological Chemistry*, Vol. 107, pp. 195-210.

Kerscher, O., 2007, "SUMO junction-whats your function? New insights through SUMO-interacting motifs", *European Molecular Biology Organization*, Vol. 8, NO 6, pp. 550-555.

Kim, K. I., S. H. Baek and C. H. Chung, 2002, "Versatile Protein Tag, SUMO: Its Enzymology and Biological Function", *Journal of Cellular Physiology*, Vol. 191, pp. 257-268.

Kim, K. I. and S. H. Baek, 2006, "SUMOylation Code in Cancer Development and Metastasis", *Molecules and Cells*, Vol. 22, pp. 247-253.

Knipscheer, P., W. J.V. Dijk, J. V. Olsen, M. Mann and T. K. Sixma, 2007, "Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation", *The EMBO Journal*, Vol. 26, pp. 2797-2807.

Kretz-Remy, C. and R. M. Tanguay, 1999, "SUMO/sentrin: protein modifiers regulating important cellular functions", *Biochem. Cell Biol.*, Vol.77, pp. 299-309.

Leach, A. R., 2001, "Molecular modelling: principles and applications (Second edition)", *Prentice Hall.*

Lee, G. W., F. Melchior, M. J. Matunis, R. Mahajan, Q. Tian and P. Anderson, 1998, "Modification of Ran GTPase-activating Protein by the Small Ubiquitin-related Modifier SUMO-1 Requires Ubc9, an E2-type Ubiquitin-conjugating Enzyme Homologue", *The Journal Of Biological Chemistry*, Vol. 273, No. 11, pp. 6503-6507.

Lieberman, A. P., 2004, "SUMO, a ubiquitin-like modifier implicated in neurodegeneration", *Experimental Neurology*, Vol. 185, pp. 204-207.

Macauley, M. S., W. J. Errington, M. Okon, M. Schrpf, C. D. Mackereth, B. A.

Schulman and L. P. McIntosh, 2004, "Structural and dynamic independence of isopeptide-linked RanGAP1 and SUMO-1", *Journal of Biological Chemistry*, Vol. 279, pp. 49131-49137.

Mahajan, R., C. Delphin, T. Guan, L. Gerace and F. Melchior, 1997, "A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2", *Cell*, Vol. 88, pp. 97-107.

Mahajan, R., L. Gerace and F. Melchior, 1998, "Molecular Characterization of the SUMO-1 Modification of RanGAP1 and Its Role in Nuclear Envelope Association", *The Journal of Cell Biology*, Vol. 140, pp. 259270.

Martin, S., K. A. Wilkinson, A. Nishimune and J. M. Henley, 2007, "Emerging extranuclear roles of protein SUMOylation in neuronal function and dysfunction", *Nature Reviews Neuroscience.*, Vol. 8, pp. 948-959.

Matunis, M. J., E. Coutavas and G. Blobel, 1996, "A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex", *Journal of Cell Biology*, Vol. 135, pp. 1457-1470.

Matunis, M. J., J. Wu and G. Blobel, 1998, "SUMO-1 Modification and Its Role in Targeting the Ran GTPase-activating Protein, RanGAP1, to the Nuclear Pore Complex", *The Journal of Cell Biology*, Vol. 140, pp. 499509.

Melchior, F., L. Gerace, 1995, "Mechanisms of nuclear protein import ", *Current Opinion in Cell Biology*, Volume 7, Issue 3, Pages 310-318.

Melchior, F., 2000, "SUMO–nonclassical ubiquitin", *Annual Review of Cell and Developmental Biology*, Vol. 16, pp. 591-626.

Melchior, F., M. Schergaut and A. Pichler, 2003, "SUMO: ligases, isopeptidases and nuclear pores", *Trends in Biochemical Sciences*, Vol. 28, pp. 612-618.

Mo, Y. Y., Y. Yu, E. Theodosiou, P. L. Rachel Ee and W. T. Beck, 2005, "A role for Ubc9 in tumorigenesis", *Oncogene*, Vol. 24, pp. 2677-2683.

Mo, Y. Y. and S. J. Moschos, 2005, "Targeting Ubc9 for cancer therapy", *Expert Opin. Ther. Targets*, Vol. 9(6), pp. 1203-1216.

Moschos, S. J. and Y. Y. Mo, 2006, "Role of SUMO/Ubc9 in DNA damage repair and tumorigenesis", *J. Mol. Hist.*, Vol. 37, pp. 309-319.

Muller, S., C. Hoege, G. Pyrowolakis, S. Jentsch, 2001, "SUMO, ubiquitin's mysterious cousin", *Nature Reviews Molecular Cell Biology*, Vol. 2, pp. 202-210.

Muller, S., A. Ledl and D. Schmidt, 2004, "SUMO: a regulator of gene expression and genome integrity", *Nature Publishing Group*, Vol. 23, pp. 1998-2008.

Palancade, B. and V. Doye, 2008, "Sumoylating and desumoylating enzymes at nuclear pores: underpinning their unexpected duties", *Trends in Cell Biology*, Vol. 18, pp. 174-183.

Pedrioli, P. G. A., B. Raught, X. D. Zhang, R. Rogers, J. Aitchison, M. Matunis and R. Aebersold, 2006, "Automated identification of SUMOylation sites using mass spectrometry and SUMmOn pattern recognition software", *Nature Methods*, Vol. 3, pp. 533-539.

Pichler,A., A. Gast, J. S. Seeler, A. Dejean and F. Melchior, 2002, "The nucleoporin RanBP2 has SUMO1 E3 ligase activity", *Cell*, Vol. 108, pp. 109-120.

Pichler, A. and F. Melchior, 2002, "Ubiquitin-Related Modifier SUMO1 and Nucleocytoplasmic Transport", *Traffic*, Vol. 3, pp. 381-387.

Pichler, A., P. Knipscheer, H. Saitoh, T. K. Sixma and F. Melchior, 2004, "The RanBP2 SUMO E3 ligase is neither HECT- nor RING-type", *Nature Structural & Molecular Biology*, Vol. 14, pp. 984-991.

Pouladi, M. A., 2009, "Taking a SUMO off a TRP for bad conduct", *Journal of Clinical Investigations*, Vol. 119, pp. 2737-2744.

Pountney, D. L., Y. Huang, R.J. Burns, E. Haan, P.D. Thompson, P.C. Blumbergs, and W.P. Gai, 2003, "SUMO-1 marks the nuclear inclusions in familial neuronal intranuclear inclusion disease", *Experimental Neurology*, Vol. 184, pp. 436-446.

Rader, A. J., C. Chennubhotla, L. Yang and I. Bahar, 2006, "The Gaussian Network Model: Theory and Applications, Normal Mode Analysis. Theory and Applications to Biological and Chemical Systems", *Mathematical and Computational Biology Series, CRC Press, Taylor & Francis Group.*

Reverter, D., C. D. Lima, 2005, "Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex", *Nature*, Vol. 435, pp. 687-692.

Rush, M. G., G. Drivas, P. D'eustachio, 1996, "The small nuclear GTPase Ran: How much does it run?", *BioEssays*, Volume 18, Issue 2, pp. 103-112.

Saitoh, H., R. Pu, M. Cavenagh and M. Dasso, 1997, "RanBP2 associates with Ubc9p and a modified form of RanGAP1", *Cell Biology*, Vol. 94, pp. 3736-3741.

Saitoh, H., D. B. Sparrow, T. Shiomi, R. T. Pu, T. Nishimoto, T. J. Mohun and M. Dasso, 1998, "Ubc9p and the conjugation of SUMO-1 to RanGAP1 and RanBP2", *Current Biology*, Vol. 8, pp. 121-124.

Sampson, D. A., M. Wang and M. J. Matunis, 2001, "The Small Ubiquitin-like Modifier-1 (SUMO-1) Consensus Sequence Mediates Ubc9 Binding and Is Essential for SUMO-1 Modification", *The Journal Of Biological Chemistry*, Vol. 276, pp. 21664-21669.

Sarge, K. D. and O. K. P. Sarge, 2009, "Sumoylation and human disease pathogenesis", *Trends in Biochemical Sciences*, Vol. 34, pp. 200-205.

Sazer, S., 1996, "The search for the primary function of the Ran GTPase continues ",

Trends in Cell Biology, Volume 6, Issue 3, pp. 81-85.

Schwamborn, K., P. Knipscheer, E. V. Dijk, W. J. V. Dijk, T. K. Sixma, R. H. Meloen and J. P.M. Langedijk, 2008, "SUMO Assay with Peptide Arrays on Solid Support: Insights into SUMO Target Sites", *The Japanese Biochemical Society*, Vol. 144, pp. 39-49.

Schwarz, S. E., K. Matuschewski, D. Liakopoulos, M. Scheffner and S. Jentsch, 1998, "The ubiquitin-like proteins SMT3 and SUMO-1 are conjugated by the UBC9 E2 enzyme", *Cell Biology*, Vol. 95, pp. 560-564.

Seeler, J. S., A. Dejean, 2003, "Nuclear and unclear functions of SUMO", *Nature Reviews Molecular Cell Biology*, Vol. 4, pp. 690-699.

Smallridge, R., 2006, "The power of SUMO", *Molecular Cell Biology*, Vol. 7, pp. 2-3.

Song, J., L. K. Durrin, T. A. Wilkinson, T. G. Krontiris and Y. Chen, 2004, "Identification of a SUMO-binding motif that recognizes SUMO-modified proteins", *The Proceedings of the National Academy of Sciences*, Vol. 101, pp. 14373-14378.

Song, J., Z. Zhang, W. Hu and Y. Chen, 2005, "Small Ubiquitin-like Modifier (SUMO) Recognition of a SUMO Binding Motif", *The Journal Of Biological Chemistry*, Vol. 280, pp. 40122-40129.

Su, H. L., S. S. L. Li, 2002, "Molecular features of human ubiquitin-like SUMO genes and their encoded proteins", *Gene*, Vol. 296, pp. 65-73.

Tatham, M. H., E. Jaffray, O. A. Vaughan, J. M. P. Desterro, C. H. Botting, J. H. Naismith and R. T. Hay, 2001, "Polymeric Chains of SUMO-2 and SUMO-3 Are Conjugated to Protein Substrates by SAE1/SAE2 and Ubc9", *The Journal Of Biological Chemistry*, Vol. 276, pp. 35368-35374.

Tatham, M. H., Y. Chen and R. T. Hay, 2003, "Role of two residues proximal to the active site of Ubc9 in substrate recognition by the Ubc9-SUMO-1 thiolester

complex", *Biochemistry*, Vol. 42, pp. 3168-3179.

Tatham, M. H., S. Kim, E. Jaffray, J. Song, Y. Chen and R. T. Hay, 2005, "Unique binding interactions among Ubc9, SUMO and RanBP2 reveal a mechanism for SUMO paralog selection", *Nature Structural & Molecular Biology*, Vol. 12, pp. 67-74.

Verger, A., J. Perdomo, M. Crossley, 2003, "Modification with SUMO", *EMBO Reports*, Vol. 4, pp. 137-142.

Vertegaal, A. C. O., J. S. Andersen, S. C. Ogg, R. T. Hay, M. Mann and A. I. Lamond, 2006, "Distinct and Overlapping Sets of SUMO-1 and SUMO-2 Target Proteins Revealed by Quantitative Proteomics", *Molecular & Cellular Proteomics*, Vol. 5.12, pp. 2298-2310.

Vertegaal, A. C. O., 2007, "Small ubiquitin- elated modifiers in chains", *Biochemical Society Transactions*, Vol. 35, pp. 1422-1423.

Watts, F. Z., 2004, "SUMO modification of proteins other than transcription factors", *Seminars in Cell & Developmental Biology*, Vol. 15, pp. 211-220.

Wilkinson, K. A., A. Nishimune, J. M. Henley, 2008, "Analysis of SUMO-1 modification of neuronal proteins containing consensus SUMOylation motifs", *Neuroscience Letters*, Vol. 436, pp. 239-244.

Wilson, V. G. and D. Rangasamy, 2001, "Intracellular Targeting of Proteins by Sumoylation", *Experimental Cell Research*, Vol. 271, pp. 57-65.

Wu, F., Y. Y. Mo,2007, "Ubiquitin-like protein modifications in prostate and breast cancer", *Frontiers in Bioscience*, Vol. 12, pp. 700-711.

Yang, M., C. T. Hsu, C. Y. Ting, L. F. Liu and J. Hwang, 2006, "Assembly of a Polymeric Chain of SUMO1 on Human Topoisomerase I in Vitro", *The Journal Of Biological Chemistry*, Vol. 281, pp. 8264-8274.

Yeh, E. T. H., L. Gong, T. Kamitani, 2000, "Ubiquitin-like proteins: new wines in new bottles", *Gene*, Vol. 248, pp. 1-14.

Yunus, A. A., C. D. Lima, 2006, "Lysine activation and functional analysis of E2-mediated conjugation in the SUMO pathway", *Nature Structural & Molecular Biology*, Vol. 13, pp. 491-499.

Zhang, H., H. Saitoh and M. J. Matunis, 2002, "Enzymes of the SUMO Modification Pathway Localize to Filaments of the Nuclear Pore Complex", *Molecular And Cellular Biology*, Vol. 22, pp. 6498-6508.

Zhao, J., 2007, "Sumoylation regulates diverse biological processes", *Cell and Molecular Life Sciences*, Vol. 64, pp. 3017-3033.

Zhu, S., H. Zhang, M. J. Matunis, 2006, "SUMO modification through rapamycin-mediated heterodimerization reveals a dual role for Ubc9 in targeting RanGAP1 to nuclear pore complexes", *Experimental Cell Research*, Vol. 312, pp. 1042-1049.

REFERENCES NOT CITED

Hoeller D., C. Hecker and I. Dikic, 2006, "Ubiquitin and ubiquitin-like proteins in cancer pathogenesis", *Nature Reviews Cancer*, Vol. 6, pp. 776-788.

Kirkin, V., I. Dikic, 2007, "Role of ubiquitin- and Ubl-binding proteins in cell signaling", *Curr. Opin. Cell Biol.*, Vol. 19, pp. 199-205.

Matunis, M. J., C. M. Pickart, 2005, "Beginning at the end with SUMO", *Nature Structural & Molecular Biology*, Vol. 12, pp. 565-566

Nalepa, G., M. Rolfe and J. W. Harper, 2006, "Drug discovery in the ubiquitin-proteasome system", *Nature Reviews Drug Discovery*, Vol. 5, pp. 596-613.

Tang, Z., C. M. Hecker, A. Scheschonka and H. Betz, 2008, "Protein interactions in the sumoylation cascade-lessons from X-ray structures", *FEBS Journal*, Vol. 275, pp. 3003-3015.

Um, J. W., K. C. Chung, 2006, "Functional modulation of parkin through physical interaction with SUMO-1", *Journal of Neuroscience Research*, Vol. 84, pp. 1543-1554.

Watson, I. R., M. S. Irwin, 2006, "Ubiquitin and ubiquitin-like modifications of the p53 family", *Neoplasia.*, Vol. 8, pp. 655-666.

Wu, P.Y., M. Hanlon, M. Eddins, C. Tsui, R. S. Rogers, J. P. Jensen, M. J. Matunis, A. M. Weissman, C. Wolberger and C. M. Pickart, 2003, "A conserved catalytic residue in the ubiquitin-conjugating enzyme family", *EMBO Journal*, Vol. 22, pp. 5241-5250.