THE DYNAMICS UNDERLYING ENHANCEMENT OF E2 LIGASE ACTIVITY BY E3 LIGASES IN SUMOYLATION

by

Melda Tozluoğlu

B.S., Chemical Engineering, Boğaziçi University, 2006

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

2008

ACKNOWLEDGEMENTS

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

I would like to express my gratitude to my thesis advisor, Professor Türkan Haliloğlu, for her invaluable guidance and encouragement through each step of my thesis. I would also like to thank the members of my Examining Committee, Assistant Professor Burak Alakent and Professor Pemra Doruker Turgut, for reading and providing their improving comments on my thesis.

I would like to thank all the members of PRC, as colleagues and friends, who have helped me in the course of this project. Special thanks to Bülent Balta, for precious discussions.

ABSTRACT

THE DYNAMICS UNDERLYING ENHANCEMENT OF E2 LIGASE ACTIVITY BY E3 LIGASES IN SUMOYLATION

Covalent attachment of SUMO (Small Ubiquitin Like Modifier) to proteins, sumoylation, is a posttranslational modification that can alter intracellular localization, interactions with other proteins or lead to modifications by other post-translational modifiers. Defects in sumoylation pathway are related to many neurological diseases, such as Huntington's disease, Parkinson's disease and more. Additionally, sumoylation is a part of cancer related pathways. Similar to other ubiquitin like modifier (Ubl) conjugation mechanisms, the conjunction of SUMO to targtes involves three groups of enzymes: E1 ligase, Aos1/Ub2 hetero-dimer; E2 ligase, Ubc9; E3 ligases, one of which is RanBP2. Differing from the other Ubl conjunction paths, the E2 ligase, Ubc9, can function without the E3 enzymes but with lower reaction efficiency. One of the target proteins that can be efficiently sumoylated by Ubc9 only, is RanGAP1. Although there are suggested models, it is not clear to date how the E3 enzyme, RanBP2, enhances sumoylation. This thesis mainly aims to identify the conformational/configurational restrictions and allosteric effects that RanBP2 may have on Ubc9-SUMO complex to increase sumoylation rate. Along, the structural motion that drives Ubc9-SUMO complex into association with RanBP2 is also addressed. For this, Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes from Ubc9-SUMO-RanGAP1-RanBP2 crystal structure are studied by Molecular Dynamics (MD) simulations. The conformational dynamics are elaborated by various means to reflect the equilibrium and dynamic behavior of these complex structures. The results in general suggest that RanBP2 restricts the conformational space of Ubc9-SUMO complex and as well as the orientational space of its monomers with respect to each other. The differences in the network of interactions between Ubc9 and SUMO residues in RanBP2 bound and unbound states suggest the determinants of the restriction in the motion observed. The correlations between the

fluctuations of the residues associated with the catalytic activity and the residues that are responsible for the specific target recognition in Ubc9 are shown to be stabilized with RanBP2 binding. The comparative analysis of the dynamics with and without RanBP2 identifies a possible allosteric effect of RanBP2 binding on the mobility and flexibility of specific Ubc9 residues, Asp100 and Lys101, which are functional in target recognition. Additionally, it is seen that the dynamics of Ubc9-SUMO complex displays a pre-existing behavior for the binding to RanBP2. This may in general imply that the dynamics of structures set the sequence of events in the association with others to form complex structures.

ÖZET

SUMOLANMADA E2 LİGAZ AKTİVİTESİNİN E3 LİGAZLAR İLE ARTTIRILMASININ BELİRLEYİCİ DİNAMKLERİ

SUMO (Küçük Ubikuitin Benzeri Değiştirici - Small Ubiquitin Like Modifier) proteininin kovalent bağlanma ile proteinleri işaretlemesi, sumolanma, okunma sonrası bir değiştirmedir. Sumolanma proteinlerin hücre içi konumlarında değişikliklere, başka proteinlerle olan etkileşimlerde değişimlere, veya başka okunma sonrası değiştiriciler tarafından işaretlenmelere sebep olabilir. Sumolanma yollarında olan bozukluklar, Huntington hastalığı, Parkinson hastalığı ve benzerleri gibi birçok nörolojik hastalık ile ilintilendirilmiştir. Bunun yanı sıra sumolanma, kanser ile ilgili etkileşim yolları üzerinde de bulunmaktadır. Diğer ubikuitin benzeri değiştiricilerde olduğu üzere, hedef proteinlere SUMO bağlanması üç grup enzim aracılığı ile gerçekleşir: E1 ligaz, Aos1/Ub2 hetero-dimeri; E2 ligaz, Ubc9; E3 ligazlar, bir örnek RanBP2. Diğer ubikuitin benzeri değiştiricilerden farklı olarak, E2 enzimi, Ubc9, daha düşük bir verim ile, E3 enzimlerininden bağımsız olarak da sumolama işlevi gösterebilir. Sadece Ubc9 aracılığı ile işaretlenebilen hedef proteinlerden biri RanGAP1'dir. Önerilen modeller olsa dahi, günümüzde, E3 enzimi RanBP2'nin sumolanma hızını arttırma mekanizması tam olarak açıklanabilmiş değildir. Bu tezin ana hedefi, RanBP2 proteininin, Ubc9-SUMO kompleksi üzerinde oluşturduğu konformasyon/konfigürasyon kısıtlamalarını ve sumolanma hızını arttırıcı allosterik etkilerini belirlemektir. Aynı süreçte, Ubc9-SUMO kompleksini RanBP2 ile etkileşime girmeye itebilecek yapısal hareketler de gösterilmiştir. Bu çalışma için, Ubc9-SUMO ve Ubc9-SUMO-RanBP2 kompleksleri Ubc9-SUMO-RanGAP1-RanBP2 kristal yapısından alınmış, ve Moleküler Dinamik (MD) Benzetimleri ile incelenmiştir. Konformasyon dinamikleri, bu komplekslerin denge ve dinamik durumlarında davranışlarını yansıtabilmek adına çeşitli yöntemler ile çalışılmış ve detaylandırılmıştır. Sonuçlar genel olarak RanBP2 bağlanmasının Ubc9-

SUMO komplexinin konformasyon uzayını ve aynı zamanda iki zincirin birbirlerine göre bulunabilecekleri konumlarını kısıtladığını göstermektedir. Ubc9 ve SUMO'nun, RanBP2 ile bağlı ve bağsız durumlarındaki etkileşim ağlarında görülen değişiklikler, görülen hareket kısıtlamalarının kaynağı olarak önerilmektedir. Ubc9 üzerinde katalitik işlev ile ilgili bölgelerin ve belirli hedef proteinleri tanımadan sorumlu amino asitlerin, konumsal dalgalanmaları arasında gözlemlenen karşılıklı bağıntı, RanBP2 varlığında dengelenmiş ve sabitlenmiştir. RanBP2 ile bağlı ve bağsız konumlardaki dinamiklerin karşılıklı incelenmesi, RanBP2 bağlanmasının hedef protein tanımada işlevli Ubc9 bölgeleri, Asp100 ve Lys101 amino asitleri, üzerinde hareket ve esneklik kısıtlayıcı allosterik etkisini ortaya koymuştur. Bunun yanı sıra, Ubc9-SUMO kompleksi, RanBP2 bağlanması için kendinden sahip olduğu dinamikler göstermektedir. Bu durum, genel anlamda kompleks oluşumu sırasında olay sıralamasının, yapıların dinamiklerini aracılığı ile belirlendiğinin bir göstergesidir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii
ABSTRACT iv
ÖZET vi
LIST OF FIGURES ix
LIST OF TABLES x
LIST OF SYMBOLS/ABBREVIATIONS xi
1. INTRODUCTION 1
1.1. Biological Background 1
1.1.1. Ubiquitin Like Proteins Superfamily and SUMO-1 1
1.1.2. Sumoylation Mechanism 2
1.1.2.1. E2 ligase Ubc9 3
1.1.2.2. Sumoylation target RanGAP1 5
1.1.2.3. E3 ligases and RanBP2 5
1.1.3. Interactions and Functional Sites of Ubc9 6
1.1.3.1. SUMO binding 6
1.1.3.2. Target Recognition 7
1.1.3.3. RanBP2 binding 9
1.2. Objective and Methodology of the Study 10
1.2.1. The model structure 11
1.2.2. Reference Structures 11
1.2.3. Brief Summary of methods 12
1.2.4. Contributions to Current Literature 12
2. MATERIALS AND METHODS 14
2.1. Studied Structures of the Sumoylation Mechanism 14
2.2. Modeling of the Thioester Bond 17
2.3. Molecular Dynamics Simulations 18
2.3.1. Theoretical Background 18
2.3.2. The Molecular Dynamics Simulation Parameters 21
2.4. Clustering Analysis 21

2.5. Mean Square Fluctuations of Positions 22
2.6. Pseudo-Dihedral Angle Fluctuations 23
2.7. Virtual Bond Vector Correlations 24
2.8. Gaussian Network Model 25
2.9. Contact Map Construction 27
3. RESULTS AND DISCUSSION 29
3.1. Mobility and Flexibility of Structures from MD Simulation Trajectories 29
3.1.1. Root Mean Square Deviations from Initial Structures 29
3.1.1.1. Isolated Ubc9 and SUMO proteins 29
3.1.1.2. Ubc9-SUMO complex 31
3.1.1.3. Ubc9-SUMO-RanBP2 complex 34
3.1.2. Mean-square fluctuations of Ubc9 and SUMO in Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes 36
3.1.2.1. Individual Ubc9 and SUMO 37
3.1.2.2. Ubc9-SUMO and Ubc9-SUMO-E3 complexes 39
3.1.3. Pseudo - Dihedral Angle Fluctuations 42
3.1.4. Time Delayed Auto-Correlations 43
3.2. Differing Interactions of Functional Residues in RanBP2 bound and unbound Ubc9-SUMO complexes 51
3.2.1. Contact Maps 52
3.2.2. Functional and Structural Interactions between Ubc9 and SUMO residues 56
3.2.2.1. Orientation of Asp100 and Lys101 of Ubc9 with respect to the catalytic site 56
3.2.2.2. Lost interactions of structural significance in Ubc9-SUMO simulation 57
3.2.2.3. Gained interactions in Ubc9-SUMO simulation 59
3.3. Representative Structures and Time Distribution of Conformations by Cluster Analysis 62
3.3.1. Clusters of Isolated Ubc9 structure 62
3.3.2. Clusters of Ubc9 - SUMO complex 63
3.3.3. Clusters of Ubc9-SUMO-RanBP2 complex 68

3.3.4. Quaternary Structure Change and Ubc9 Mobile Fragments from Clusters of Conformations 70
3.4. Correlations between the Fluctuations 73
3.4.1. Correlations involving RanBP2 binding sites 74
3.4.1.1. Anchor-like behavior of Ubc9, residues Val27 to Glu42 74
3.4.1.2. Existing correlations among RanBP2 binding sites of Ubc9 and SUMO 74
3.4.1.3. New correlations among SUMO and RanBP2 binding sites of Ubc9 80
3.4.2. Correlations involving functional residues in target recognition . 80
3.4.3. Correlations from Gaussian Network Model Analysis 81
4. CONCLUSIONS AND FUTURE STUDIES 88
4.1. Conclusions 88
4.2. Future Studies 91
APPENDIX A: ALIGNMENT EFFECT on POSITIONAL MEAN SQUARE FLUCTUATIONS 92
APPENDIX B: CONTACT MAPS 95
APPENDIX C: CLUSTERING ANALYSIS 103
REFERENCES 104

LIST OF FIGURES

Figure 1.1. Structures of SUMO family members 2

Figure 1.2. Sumoylation Mechanism 4

Figure 1.3. Ubc9 contacts with SUMO 7

Figure 1.4. Ubc9 contacts with sumoylation target RanGAP1 8

Figure 1.5. Ubc9 contacts with E3 enzyme RanBP2 10

Figure 2.1. Crystal structure of the Ubc9-SUMO-RanGAP1-RanBP2 complex. 15

Figure 2.2. Structure of the Ubc9-SUMO complex. 15

Figure 2.3. Structure of the Ubc9-SUMO-RanBP2 complex. 16

Figure 2.4. Pseudo-dihedral Angle Model 23

Figure 2.5. Virtual Bond Vector 24

Figure 3.1. Isolated Ubc9 rmsd 30

Figure 3.2. Isolated SUMO rmsd 31

Figure 3.3. Ubc9-SUMO Complex rmsd 32

Figure 3.4. Individually Aligned Ubc9 and SUMO rmsds from Ubc9-SUMO complex 32

Figure 3.5. Snapshots form Ubc9-SUMO simulation at 5 ns and 12 ns 33

Figure 3.6. Ubc9-SUMO-RanBP2 Complex RMSD 34

Figure 3.7. Individually Aligned Ubc9, SUMO and RanBP2 RMSDs from Ubc9-SUMO-RanBP2 complex 35

Figure 3.8. Snapshots from Ubc9-SUMO-RanBP2 simulation at 6.7 and 44.7 ns 35

Figure 3.9. Orientation of SUMO in two complexes 36

Figure 3.10. Isolated SUMO Positional MSF 38

Figure 3.11. Isolated SUMO Positional MSF 38

Figure 3.12. Alignment effects on SUMO Positional MSF 39

Figure 3.13. Ubc9 Positional MSF upon Complex Formation 40

Figure 3.14. SUMO Positional MSF upon Complex Formation 41

Figure 3.15. Mean Pseudo-Dihedral Angle Fluctuations 43

Figure 3.16. Time Delayed Auto-correlations of Virtual Bond Vectors for Ubc9-SUMO-RanBP2 complex 44

Figure 3.17. Time Delayed Auto-correlations of Virtual Bond Vectors at Selected Delays for Ubc9-SUMO-RanBP2 complex 45

Figure 3.18. Normalized Time delayed Auto-correlations of Virtual Bond Vectors for Ubc9-SUMO-RanBP2 complex 46

Figure 3.19. Time Delayed Auto-correlations of Virtual Bond Vectors for Ubc9-SUMO complex 48

Figure 3.20. Time Delayed Auto-correlations of Virtual Bond Vectors at Selected Delays for Ubc9-SUMO complex 49

Figure 3.21. Normalized Time delayed Auto-correlations of Virtual Bond Vectors for Ubc9-SUMO complex 50

Figure 3.22. Virtual Bond Vector between Asp100 and Lys101 51

Figure 3.23. Ubc9-SUMO Contact Map at 5 ns, 7 $\mathring{A}$ 52

Figure 3.24. Ubc9-SUMO Contact Map at 10 ns, 7 $\mathring{A}$ 53

Figure 3.25. Ubc9-SUMO Contact Map at 58 ns, 7 $\mathring{A}$ 53

Figure 3.26. Ubc9-SUMO-RanBP2 Contact Map at 5 ns, 7 $\mathring{A}$ 54

Figure 3.27. Ubc9-SUMO-RanBP2 Contact Map at 10 ns, 7 $\mathring{A}$ 54

Figure 3.28. Ubc9-SUMO-RanBP2 Contact Map at 50 ns, 7 $\mathring{A}$ 55

Figure 3.29. Distance between Cys93 and Asp100 of Ubc9 56

Figure 3.30. Distance between Cys93 and Lys101 of Ubc9 57

Figure 3.31. Distance between Arg 63 of SUMO and Glu122 of Ubc9 58

Figure 3.32. Distance between Gln29 of SUMO and Gln111 of Ubc9 59

Figure 3.33. Distance between Gln92 of SUMO and Arg104 of Ubc9 60

Figure 3.34. Distance between Gln94 of SUMO and Glu122 of Ubc9 61

Figure 3.35. Distance between Gln92 of SUMO and Glu118 of Ubc9 61

Figure 3.36. Clusters of Ubc9 63

Figure 3.37. Best members for Clusters of Ubc9 64

Figure 3.38. Clusters of Ubc9-SUMO complex, at 2 Å cut-off 65

Figure 3.39. Clusters of Ubc9-SUMO complex, at 1.7 Å cut-off 66

Figure 3.40. Best members for Clusters of Ubc9-SUMO 67

Figure 3.41. Clusters of Ubc9-SUMO-RanBP2 complex, at 2 $\mathring{A}$ 69

Figure 3.42. Best members for Clusters of Ubc9-SUMO-RanBP2 69

Figure 3.43. SUMO motion in Ubc9-SUMO complex representative structures . 71

Figure 3.44. Alterations in Ubc9 mobility upon RanBP2 binding 72

Figure 3.45. Alterations in SUMO mobility upon RanBP2 binding 72

Figure 3.46. Fluctuation Correlations for Ubc9-SUMO complex, 0.5 to 58 ns . . 75

Figure 3.47. Fluctuation Correlations for Ubc9-SUMO complex, 0.5 to 6 ns . . 75

Figure 3.48. Fluctuation Correlations for Ubc9-SUMO complex, 6 to 12 ns . . . 76

Figure 3.49. Fluctuation Correlations for Ubc9-SUMO complex, 12 to 24 ns . . 76

Figure 3.50. Fluctuation Correlations for Ubc9-SUMO complex, 24 to 31 ns . . 77

Figure 3.51. Fluctuation Correlations for Ubc9-SUMO complex, 31 to 42 ns . . 77

Figure 3.52. Fluctuation Correlations for Ubc9-SUMO complex, 42 to 58 ns . . 78

Figure 3.53. Fluctuation Correlations for Ubc9-SUMO-RanBP2 complex, 0.4 to 50 ns 78

Figure 3.54. Fluctuation Correlations for Ubc9-SUMO-RanBP2 complex, 0.4 to 28 ns 79

Figure 3.55. Fluctuation Correlations for Ubc9-SUMO-RanBP2 complex, 28 to 50 ns 79

Figure 3.56. Ubc9-SUMO-RanBP2 Correlations of Motion for Cluster 1 82

Figure 3.57. Ubc9-SUMO-RanBP2 Correlations of Motion for Cluster 2 83

Figure 3.58. Ubc9-SUMO-RanBP2 Correlations of Motion for Cluster 1 83

Figure 3.59. Ubc9-SUMO-RanBP2 Correlations of Motion for Cluster 4 84

Figure 3.60. Ubc9-SUMO Correlations of Motion for Cluster 1 85

Figure 3.61. Ubc9-SUMO Correlations of Motion for Cluster 2 85

Figure 3.62. Ubc9-SUMO Correlations of Motion for Cluster 3 86

Figure 3.63. Ubc9-SUMO Correlations of Motion for Cluster 4 86

Figure 3.64. Ubc9-SUMO Correlations of Motion for Cluster 5 87

Figure 3.65. Ubc9-SUMO Correlations of Motion for Cluster 6 87

Figure A.1. Alignment effects on Ubc9 MSF (Ubc9-SUMO complex) 92

Figure A.2. Alignment effects on SUMO MSF (Ubc9-SUMO-RanBP2 complex) 93

Figure A.3. Alignment effects on Ubc9 MSF (Ubc9-SUMO-RanBP2 comlex) . 93

Figure A.4. Alignment effects on RanBP2 MSF (Ubc9-SUMO-RanBP2 comlex) 94

Figure B.1. Ubc9-SUMO Contact Map at 15 ns, 7 $\mathring{A}$ 95

Figure B.2. Ubc9-SUMO Contact Map at 20 ns, 7 $\mathring{A}$ 95

Figure B.3. Ubc9-SUMO Contact Map at 25 ns, 7 $\mathring{A}$ 96

Figure B.4. Ubc9-SUMO Contact Map at 30 ns, 7 $\mathring{A}$ 96

Figure B.5. Ubc9-SUMO Contact Map at 35 ns, 7 $\mathring{A}$ 97

Figure B.6. Ubc9-SUMO Contact Map at 40 ns, 7 $\mathring{A}$ 97

Figure B.7. Ubc9-SUMO Contact Map at 45 ns, 7 $\mathring{A}$ 98

Figure B.8. Ubc9-SUMO Contact Map at 50 ns, 7 $\mathring{A}$ 98

Figure B.9. Ubc9-SUMO Contact Map at 55 ns, 7 $\mathring{A}$ 99

Figure B.10. Ubc9-SUMO-RanBP2 Contact Map at 15 ns, 7 $\mathring{A}$ 99

Figure B.11. Ubc9-SUMO-RanBP2 Contact Map at 20 ns, 7 $\mathring{A}$ 100

Figure B.12. Ubc9-SUMO-RanBP2 Contact Map at 25 ns, 7 $\mathring{A}$ 100

Figure B.13. Ubc9-SUMO-RanBP2 Contact Map at 30 ns, 7 $\mathring{A}$ 101

Figure B.14. Ubc9-SUMO-RanBP2 Contact Map at 35 ns, 7 $\mathring{A}$ 101

Figure B.15. Ubc9-SUMO-RanBP2 Contact Map at 40 ns, 7 $\mathring{A}$ 102

Figure B.16. Ubc9-SUMO-RanBP2 Contact Map at 45 ns, 7 $\mathring{A}$ 102

Figure C.1. Clusters of Ubc9-SUMO-RanBP2 complex, at 1.7 $\mathring{A}$ 103

Figure C.2. Clusters of Ubc9-SUMO-RanBP2 complex, at 1.5 $\mathring{A}$ 103

LIST OF TABLES

Table 2.1. Structure Details and Simulation Lengths for Analyzed Complexes 16

Table 3.1. Time Spent at each Cluster for Ubc9-SUMO Simulations 68

Table 3.2. Time Spent at each Cluster for Ubc9 - SUMO - E3 Simulations . . 70

Table 3.3. Time Intervals for calculated Fluctuation Correlations 73

LIST OF SYMBOLS/ABBREVIATIONS

a_i	Acceleration of particle i
C^{α}	Alpha-carbon
CM_i	Auto-correlation of Virtual bond vector i
CO_{ij}	Orientational cross-correlation of positions i and j
$d_{i,j}$	Distance between residues i and j
F_i	Force acting on particle i
M_i	Virtual bond vector of particle i
m_i	Mass of particle i
N	Number of atoms
ns	nanosecond
R_i	Position of particle i
V	Potential energy
$\mathring{A}$	Angstrom
Γ	Kirchhoff matrix
ϕ	Pseudo-dihedral angle
τ	Time frame
A	Ala, Alanine
C	Cys, Cysteine
D	Asp, Aspartic acid
E	Glu, Glutamic acid
F	Phe, Phenylalanine
G	Gly, Glycine
GNM	Gaussian Network Model
H	His, Histidine
I	Ile, Isoleucine
K	Lys, Lysine
L	Leu, Leucine

M	Met, Methionine
MD	Molecular Dynamics
MSF	Mean squared fluctuations
N	Asn, Asparagine
NMR	Nuclear Magnetic Resonance
P	Pro, Proline
PDB	Protein Data Bank
PME	Particle Mesh Ewald
Q	Gln, Glutamine
R	Arg, Arginine
rmsd	Root mean squared deviation
S	Ser, Serine
SUMO	Small Ubiquitin Like Modifier
T	Thr, Threonine
Ub	Ubiquitin
V	Val, Valine
vdw	van der Waals
Ubl	Ubiquitin Like
W	Trp, Tryptophan
WT	Wild-type
Y	Tyr, Tyrosine
3-D	Three-dimensional

1. INTRODUCTION

1.1. Biological Background

1.1.1. Ubiquitin Like Proteins Superfamily and SUMO-1

Protein function is regulated by numerous mechanisms, one of which is post-translational modifications. The post-translational modifications can alter the activity, interactions with other proteins, degradation level or spatial distribution of the modified protein *in vivo*. Covalent binding of chemical groups, such as acetate, phosphate, lipids, sugars, or other proteins to the target protein is mostly carried out via enzymatic reactions (Melchior, 2000; Müller *et al.*, 2001). Ubiquitin (Ub) and Ubiquitin-like (Ubl) modifiers, which are proteins of approximately 100 amino-acids, are covalently bound to lysine residues on target proteins with analogous mechanisms (Capili and Lima, 2007a; Müller *et al.*, 2001). Ub and Ubl modification takes part in cellular processes including differentiation, apoptosis, cell cycle, and stress response (Bernier-Villamor *et al.*, 2002; Melchior, 2000; Müller *et al.*, 2001). SUMO-1 (Small ubiquitin-related modifier), also known as PIC1, UBL1, GMP1 , and Sentrin, is a member of Ubl superfamily (Boddy *et al.*, 1996; Mahajan *et al.*, 1997; Matunis *et al.*, 1996; Okura *et al.*, 1996; Shen *et al.*, 1996). SUMO-1 is a 101 amino acid protein with 18 per cent sequence identity to ubiquitin (Mahajan *et al.*, 1997). Despite the relatively low sequence similarity, SUMO-1 has the $\beta\beta\alpha\beta\beta\alpha\beta$ fold of the ubiquitin protein family. There is a high similarity in the 3-Dimensional structures of Ubiquitin and SUMO-1 (Bayer *et al.*, 1998).

Additional to SUMO1, there are 3 more SUMO proteins in mammals, SUMO-2 to -4. The sequence identities of SUMO-2 and SUMO-3 to SUMO-1 are 43 per cent and 42 per cent, respectively (Capili and Lima, 2007b). There is also high structural similarity between the family members, available structures of SUMO-1, -2 and -3 reveal the same $\beta\beta\alpha\beta\beta\alpha\beta$ fold (Figure1.1).

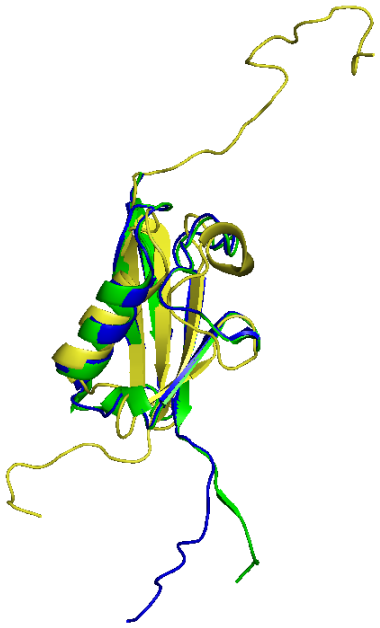

Figure 1.1. Structures of SUMO-1 (yellow) (PDB ID: 1A5R), SUMO-2 (blue) (PDB ID: 2IO0) and SUMO-3 (green) (PDB ID: 2IO1) proteins. Structures are visualized and superimposed using program PyMol (DeLano, 2002).

There are at least 100 different proteins reported to be targets for SUMO conjunction (Panse *et al.*, 2004; Wohlschlegel *et al.*, 2004; Zhau *et al.*, 2004; Zhou *et al.*, 2004), where SUMO is attached to consensus motifs on targets with the aid of E1, E2 and E3 enzymes (Section 1.1.2). Similar to ubiquitination, in addition to mono-sumoylation, poly-SUMO chains can be formed on targets. SUMO-2 and SUMO-3 chain formation on targets have been demonstrated *in vivo* and *in vitro*, whereas poly-SUMO-1 chains are shown to be formed *in vitro* (Tatham *et al.*, 2001; Yang *et al.*, 2006).

Modification by SUMO can have a variety of biological consequences. Modified proteins can change intracellular localization, their interactions with other proteins or modifications by other post-translational modification can be altered. In addition, there is an evidence indicating disorder in neuronal SUMO conjunction is related to many neurological diseases, such as Huntington's disease, Parkinson's disease and more (Martin *et al.*, 2007; Zhao, 2007). Additionally, regulation of protein activity by sumoylation has roles in tumorigenesis (Mo *et al.*, 2005).

1.1.2. Sumoylation Mechanism

The mechanism for SUMO conjunction is analogous to conjunction mechanisms of Ub and other Ubls, involving consequent activity of E1, E2 and in most cases, E3 enzymes. The dedicated E1 is a heterodimer of Aos1 and Uba2 proteins for SUMO (Melchior, 2000). Ubc9 is the E2 enzyme of the sumoylation pathway (Bernier-Villamor *et al.*, 2002) and there are several E3 enzymes, one of which is RanBP2.

SUMO, which is produced as an inactive precursor, needs to be cleaved at the C-terminus for maturation. The cleavage of SUMO produces the C-terminal Gly-Gly motif, which is conserved in Ub/Ubl family (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. At the second step of sumoylation, the SUMO thioester is transferred to the active cysteine of the E2 enzyme (Melchior, 2000). In contrast to several E2 enzymes used in ubiquitin conjugation, Ubc9 is the single known E2 enzyme for sumoylation (Bernier-Villamor *et al.*, 2002). 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 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). SUMO conjunction is a reversible process, isopeptidases known as SENPs (SUMO/sentrin specific protease) de-conjugate SUMO from modified targets (Martin *et al.*, 2007; Zhao, 2007). Figure 1.1.2 gives a schematic representation of the SUMO conjunction mechanism.

1.1.2.1. E2 ligase Ubc9. E2 ligase Ubc9 is conserved from yeast to human (Melchior, 2000). It is essential protein for cell viability (Melchior, 2000; van Waardenburg *et al.*, 2006). Ubc9 can recognize a consensus Sumoylation motif existent in most of the

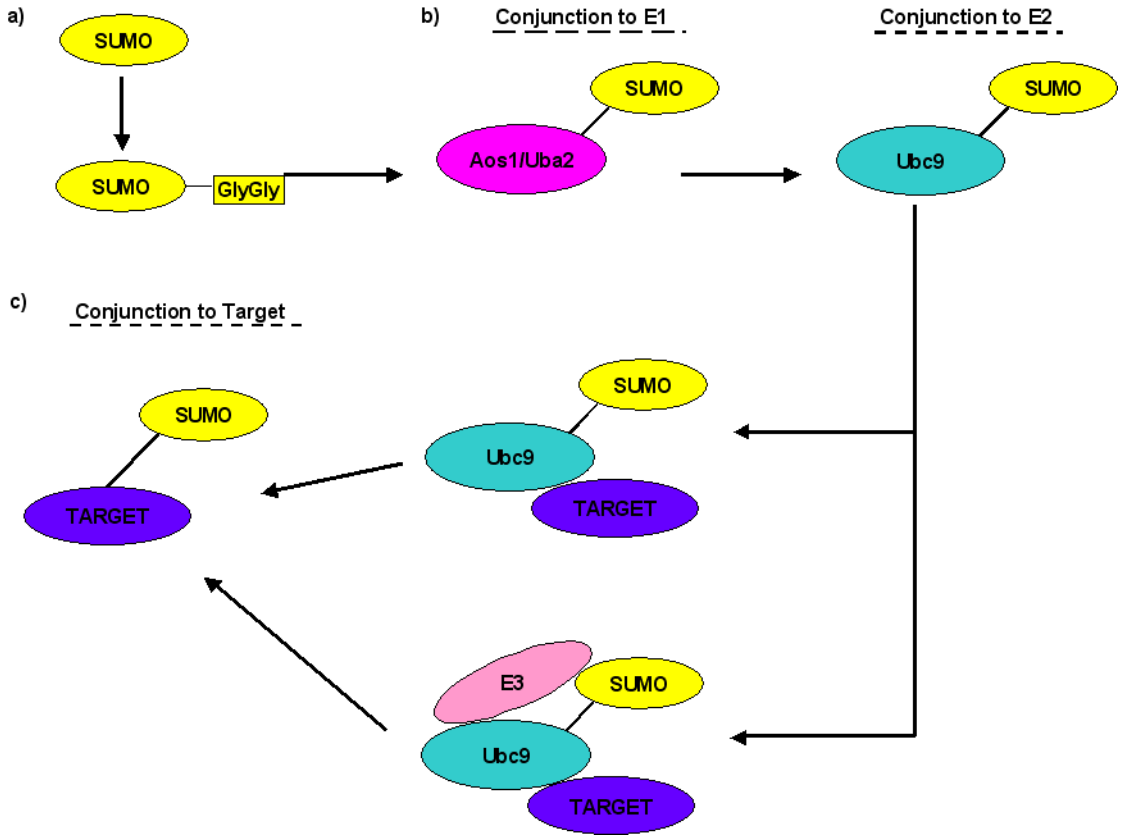

Figure 1.2. Sumoylation Mechanism. (a) SUMO is cleaved at C-terminus. (b) Mature SUMO is activated and transferred on the active cysteine of E1 enzyme, Aos1/Uba2 heterodimer. The SUMO thioester is then transferred to the active cysteine of E2 enzyme, Ubc9. (c) SUMO is transferred to the target. The conjunction can be mediated by a dedicated E3 enzyme. Alternatively Ubc9 can directly transfer SUMO to the target.

known SUMO targets. Thus, Ubc9 partly plays a role in target recognition. SUMO modification motif is a "Ψ-K-x-D/E" tetrapeptide motif, where Ψ represents a hydrophobic residue, K is the SUMO acceptor lysine, x is any amino acid and D/E is an acidic residue (Bernier-Villamor *et al.*, 2002). In addition to covalently bound SUMO, a non-covalent complex between Ubc9 and SUMO can also be formed. The binding region for this complex partly coincides with E1 and E3 interaction sites (Capili and Lima, 2007b). Ubc9 also interacts with E3 enzymes during transfer of SUMO to targets, (Reverter and Lima, 2005) details of which are given in Section 1.1.3.

1.1.2.2. Sumoylation target RanGAP1. RanGAP1 is a Ran-GTPase activating protein. It is highly concentrated in complex with RanBP2 at the cytoplasmic periphery of the Nuclear Pore Complexes (NPC). Sumoylation of RanGAP1 is essential for its localization to the NPC. Additionally, RanGAP1 mediated GTP hydrolysis by Ran is required for nuclear import, and inhibition of RanGAP1 localization to NPC also inhibits nuclear protein import. (Bischoff *et al.*, 1994; Mahajan *et al.*, 1997) Therefore, sumoylation of RanGAP1 is a key factor in the integrity of nuclear protein import mechanism. The SUMO consensus motif "LKSE" of RanGAP1 resides between residues 525-528. Apart from the consensus motif, RanGAP1 has a second contact surface with Ubc9. This additional contact surface, which is thought to be responsible for the higher efficiency of modification compared to other substrates, includes the residues of the two helices between 511-522 and 555-566 (Bernier-Villamor *et al.*, 2002).

1.1.2.3. E3 ligases and RanBP2. The known E3 ligases for Ub/Ubl conjunction can be classified into two main categories, RING type and HECT type E3 ligases. RING type E3 ligases mediate transfer of Ub/Ubl to the target by binding both to E2-Ubl, via a zinc domain, and to target proteins. HECT type E3 ligases are "true enzymes" in the sense that they first transfer the Ub/Ubl on themselves from the E2 enzyme, forming a thioester bond, and then transfer the modifier on the target (Pichler *et al.*, 2002; Reverter and Lima, 2005). RanBP2 is a protein of the Nuclear Pore Complex (NPC), and forms a stable complex with sumoylated RanGAP1 (Mahajan *et al.*, 1997). RanBP2 is one of the E3 ligases identified for sumoylation, but does not fall into either of the RING

or HECT type E3 ligase categories. RanBP2 works for a subgroup of sumoylation targets, indicating substrate specificity determination. A fragment of RanBP2 between residues 2633-2761, consisting of IR1-M-IR2 domains, is sufficient for E3 activity *in vivo* and *in vitro*. More specifically, IR1-M and M-IR2 constructs are also functional and IR1-M is the catalytic core domain (Pichler *et al.*, 2002; Pichler *et al.*, 2004; Reverter and Lima, 2005). Although it is possible that full length RanBP2 binds to targets and functions as an adaptor, the activity of the RanBP2 fragment indicates RanBP2 exerts its catalytic effect by altering Ubc9's properties, rather than by mediating target interactions (Pichler *et al.*, 2002). The available structure of SUMO-RanGAP1-Ubc9-RanBP2 complex supports this idea and suggests that E3 enzyme RanBP2 increases reaction efficiency by reducing the conformational flexibility of the Ubc9-SUMO complex. The existence of an E3 enzyme in the complex may align the thioester in position, and prevent non-productive conformations of Ubc9-SUMO (Reverter and Lima, 2005)

The available 3-Dimensional structures of complexes SUMO-RanGAP1-Ubc9-RanBP2, and Ubc9-RanGAP1 hetero-tetramer reveal strong interactions between RanGAP1 and Ubc9. Ubc9 can also facilitate sumoylation of RanGAP1 without an E3 enzyme. It is a matter of debate weather RanBP2 exerts E3 activity for RanGAP1 sumoylation in the complex of Ubc9-SUMO-RanGAP1-RanBP2 or it only maintains the complex at the NPC (Bernier-Villamor *et al.*, 2002; Reverter and Lima, 2005).

1.1.3. Interactions and Functional Sites of Ubc9

1.1.3.1. SUMO binding. E2 enzyme Ubc9 can form both covalently and non-covalently bound complexes with SUMO (Capili and Lima, 2007b; Reverter and Lima, 2005). In the non-covalent complex formed between Ubc9 and SUMO, the 3-Dimensional structure reveals direct interactions between seven Ubc9 residues and eleven SUMO residues, with a cut off of 3.8 $\AA$ between the closest atoms in the respective residues: Arg13, Arg17, Lys18, Phe22, Gly23, Val27, and Lys49 and Lys25, Gly28, Gln29, Arg63, Glu67, Gly81, Glu83, Asp86, Val87, Glu89, and Tyr91, respectively. In contrast, a NMR study with chemical shift perturbations identified 20 SUMO-1 residues (Ile27-Ser32, Val38, Arg63-Ile71, and Met82-Val90) and 25 Ubc9 residues (Ser7, Ala10, Lys14,

His20, Phe22, Val27, Thr35, Asn37, Ala44, Lys49, Leu60, Met62, Leu63, Glu99, Ile109, Lys110, Ile112, Leu113, and Ser158) (Capili and Lima, 2007b; Liu *et al.*, 1999; Tatham *et al.*, 2003). The binding sites of Ubc9 with non-covalently bound SUMO overlap with those sites of Ubc9 that interacts with E3 enzyme RanBP2. On the other hand, covalent and non-covalent bound SUMO utilize different surfaces on Ubc9, thus, a single Ubc9 can interact with two SUMO molecules at a time, one being covalently attached to active cysteine, the other being non-covalently attached. It is also suggested the non-covalent binding between SUMO and Ubc9 may have significance in SUMO chain formation (Capili and Lima, 2007b).

The covalent binding occurs by transfer of SUMO form E1 enzyme (Aos1/Uba2) to Ubc9. In this complex, the C-terminal glycine residue of SUMO is attached to the active cysteine (Cys93) of Ubc9. In the covalently bound complex, several Ubc9 residues make contacts with SUMO. Glu122 of Ubc9 makes a hydrogen bond with Arg63 of SUMO. Ubc9 Arg104 in Ubc9 contacts Gln92 in SUMO and glutamine Gln111 of Ubc9 contacts with the main chain atoms of Gln29 in SUMO. Thr95 and Gly96 of SUMO contact with Ser95 of Ubc9. Additionally, Arg85 of Ubc9 in the structurally significant HPN motif (discussed below) makes contacts with C- terminal di-Gly motif of SUMO (Reverter and Lima, 2005). These contacts can be visualized in Figure 1.3.

1.1.3.2. Target Recognition. Target recognition is partially accomplished by E2 enzyme Ubc9 in sumoylation pathway. Ubc9 can recognize the "Ψ-K-x-D/E" consensus SUMO conjugation motif existing in most of the sumoylation target proteins. A complex between sumoylation target RanGAP1 and Ubc9 reveals the interactions directing target recognition by Ubc9 (Bernier-Villamor *et al.*, 2002). The first residue of the consensus motif Ψ (hydrophobic residue)) is Leu523 on RanGAP1 in this complex. Leu523 has van der Waals (vdw) contacts with Ubc9 residues Pro128, Ala129, Gln130 and Ala131. The second residue of the motif is the acceptor lysine of SUMO conjunction, and corresponds to Lys524 in RanGAP1 structure. Asp127, Pro128, Ala129 and Tyr87 in Ubc9 form a groove to fit the acceptor lysine. Asp127 and Lys524 are within a hydrogen bonding distance. Lys526 is also located at hydrogen bonding distance to the

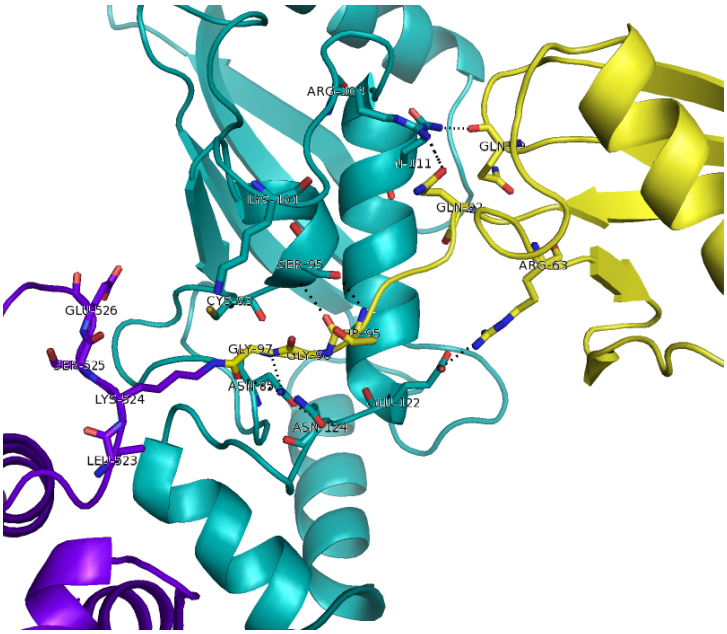

Figure 1.3. Ubc9 contacts with SUMO. SUMO-1 is represented in yellow, Ubc9 blue. The residues of sumoylation target RanGAP1 are also represented in the figure, in purple.

active cysteine Cys93 of Ubc9, where it can attack the thioester between C-terminus of SUMO and Cys93 of Ubc9 during conjunction. The non-conserved third residue of the motif is Ser525 in the complex. This residue serves as to create a distance between second and forth residues of the motif , Lys524 and Glu526 in RanGAP1. The last residue of the motif is an aspartic acid or a glutamic acid residue; that is Glu526 in RanGAP1. This residue is in hydrogen bonding distance with Ser89, Thr91 and Lys74 of Ubc9. There are also vdw interactions between Glu526 of target and Tyr87 of Ubc9 (Bernier-Villamor *et al.*, 2002). The interactions are displayed in Figure 1.4. The interaction surface for RanGAP1 and Ubc9 complex is not exclusive to the sumoylation motif; there are additional contacts between these two proteins in the crystal structure. The two helices of RanGAP1, Ser509-Met520 and Lys553-Thr564, interact with the helix, Ala131-Asn140, of Ubc9.

Biochemical and mutational analysis indicates rather catalytic then structural roles for residues Asn85, Tyr87 and Asp127. Additionally, Asp100 and Lys101 are oriented such that they can interact with an approaching target. Mutation of Asp100 and Lys101 also reduces SUMO conjunction to targets by 2.5 fold (Tatham *et al.*, 2003;

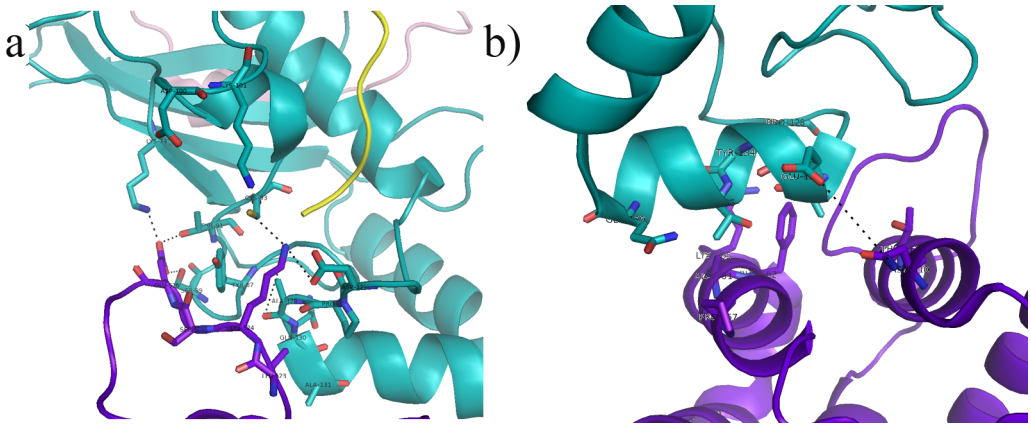

Figure 1.4. Ubc9 contacts with sumoylation target RanGAP1. Ubc9 is represented in blue and RanGAP1 in purple a) Ubc9-RanGAP1 contacts in the consensus sumoylation motif. b) Ubc9-RanGAP1 additional binding surface.

Yunus and Lima, 2006).

Two other conserved motifs of Ubc9 are also thought to have structural significance; the His-Pro-Asn (HPN) motif between residues 83-85, and Glu-Pro-Asn (EPN) motif at residues 122-124. Especially the HPN motif maintains hydrogen-bonding networks, taking part in orienting the SUMO C-terminal Gly-Gly motif. EPN motif is located closer to the SUMO contact regions (Bernier-Villamor *et al.*, 2002; Wu *et al.*, 2003).

Mutational analysis revealed severe functional defects for mutation of Ubc9 at residues Asn85, Tyr87, Asp100, Pro128, Ala129, Ala131 and Tyr134. To a lesser extent, mutations of residues Lys74, Asp127, Glu132 and Thr135 also reduced activity (Bernier-Villamor *et al.*, 2002; Yunus and Lima, 2006). Mutations of Glu132 and Tyr134, on the other hand, do not affect SUMO conjunction to other targets such as p53 or IκBα significantly. These residues take part in the contacts between the additional binding surface of RanGAP1 and Ubc9, which is not conserved in other targets. These mutations may correlate with functionality of the SUMO conjunction mechanism for this specific target (Bernier-Villamor *et al.*, 2002; Yunus and Lima, 2006).

1.1.3.3. RanBP2 binding. The contact regions of Ubc9 and IR1-M domain fragment of RanBP2 are revealed by the available 3-Dimensional structure of the Ubc9-SUMO-RanGAP1 (carboxy-terminal domain of RanGAP1)-RanBP2 (IR1 - M fragment) complex. Starting from the C-termini of IR1, RanBP2 contacts the Ubc9 residues Arg8, Pro69, Pro105, Ala106 and the N-terminal Ubc9 helix. The following loop of RanBP2 makes contacts with the SUMO-Ubc9 interface, including the residues of N-terminal Ubc9 helix and loop between N-terminal β-sheets of SUMO (Ile27-Glu33). The acidic residues of RanBP2 between residues Glu2671 and Asp2676 contact Arg13, Arg17 and Lys30 of Ubc9. Hydrophobic residues of RanBP2 between Phe2677-Leu2684 pack on the β-sheet residues of Ubc9 between Lys30-Asn40. Finally, Leu2688, Tyr2689 and Leu2690 of RanBP2 contacts Phe22 and β-sheets around residues Asn40-Leu60 of Ubc9. Gly47 and Gly55 are in hydrogen bonding distance with Tyr2689 (Reverter and Lima, 2005). The interactions are displayed in Figure 1.5.

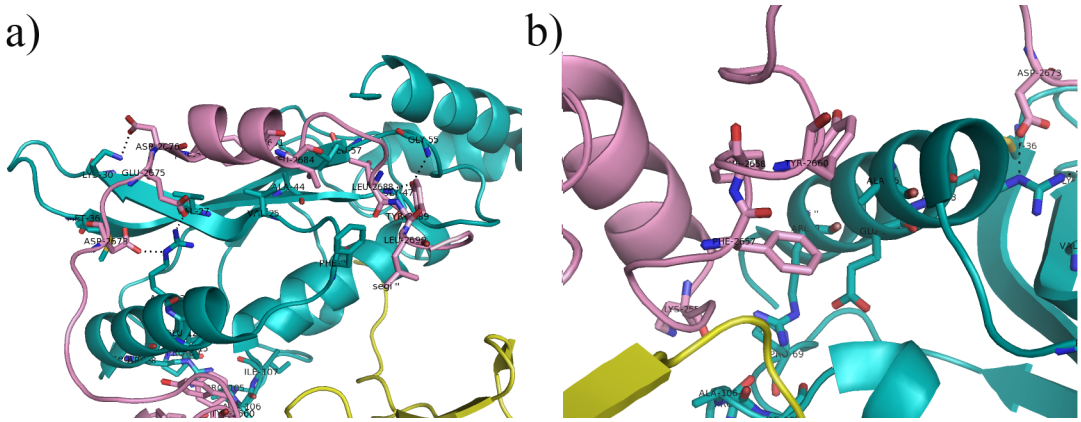

Figure 1.5. Ubc9 contacts with E3 anzyme RanBP2. Ubc9 is represented in blue, RanBP2 in pink and SUMO in yellow. a) Ubc9-RanBP2 contact region b) Ubc9-RanBP2 contacts at the interface between Ubc9-SUMO and RanBP2.

Biochemical analysis also revealed a set of residues in Ubc9, Arg17, Phe22-Asn31, Met36, Asn37, Asn40-Cys75, Glu78-Phe82, Gly115, Ile116, Phe155-Ala156, as interacting with RanBP2 (Tatham *et al.*, 2005). Mutations of residues Phe22, Val25, Val27, Leu57 and Lys59 resulted in severe impairment of RanBP2 binding, whereas mutations of residues Glu42, Lys48, Glu54 and Arg61 had moderate effect (Tatham *et al.*, 2005). With respect to the catalytic effects of RanBP2, Ile4 was identified together with Val25 and Leu57 of Ubc9 as RanBP2 interacting residues (Pichler *et al.*, 2004). The mutations of RanBP2 residues that significantly impaired function in sumoylation included

Leu2651, Leu2653 and Phe 2657. Mutations with moderate effects were also identified, Pro2654, Cys2659, Asp2676 and Phe2677 (Pichler *et al.*, 2004).

1.2. Objective and Methodology of the Study

The mechanism of SUMO conjunction has many aspects, which can stand as separate research areas alone. This study mainly focuses on the dynamics of the covalently bound Ubc9-SUMO complex and Ubc9-SUMO in complex with E3 enzyme, RanBP2. The E2 enzyme, Ubc9, can function without the guidance of a E3 enzyme. To this end, the objective is to clarify the structural and dynamic determinants of the increased reaction efficiency in the RanBP2 bound complex, and the motion that derive Ubc9-SUMO complex into forming a complex with E3.

The E3 enzyme, RanBP2, is thought to have an allosteric effect on Ubc9-SUMO complex that enhances the efficiency of SUMO transfer from Ubc9 to targets. The limitation in the available conformational space of Ubc9-SUMO complex with the bound RanBP2 may prevent some non-productive conformations (Pichler *et al.*, 2002; Reverter and Lima, 2005) and affect the sumoylation activity. To account for these propositions, we aim to carry our molecular dynamic simulations for Ubc9-SUMO complex with and without bound RanBP2 and analyze the MD trajectories exhaustively to identify the interactions responsible for the high effectiveness of RanBP2 bound complexes. This is also to address the potential allosteric effect of RanBP2 on these functional residues.

1.2.1. The model structure

The 3-Dimensional X-ray structure is available for Ubc9-SUMO-RanGAP1 (target)-RanBP2 (E3 enzyme) complex (Reverter and Lima, 2005). The PDB code for the structure is 1Z5S. Functional fragments of sumoylation target RanGAP1 and E3 enzyme RanBP2 are crystallized in this complex and SUMO is transferred from Ubc9-SUMO thioester to RanGAP1-SUMO isopeptide bond. Two of the structures to be analyzed are extracted from this complex: Ubc9-SUMO and Ubc9-SUMO-RanBP2. It is suggested that Ubc9 and SUMO do not alter their conformation to a high extent upon

conjunction of SUMO to target (Reverter and Lima, 2005). Since the focus of the study is on the steps prior to transfer of SUMO to target, the target structure is eliminated from the complexes. With the similar purposes, the bond between RanGAP1 and SUMO is transferred to the Ubc9 active cysteine. The molecular dynamic simulations are run for 50 and 58 ns for Ubc9-SUMO-RanBP2 and Ubc9-SUMO complexes respectively.

1.2.2. Reference Structures

The dynamics of E3 bound and non-bound complexes are analyzed comparatively. To be able to observe the extent of dynamic changes more clearly, simulations for isolated Ubc9 and isolated SUMO proteins are also run for 32.5 and 35 ns, respectively. The PDB code is 1A3S for isolated Ubc9 and is 1A5R for isolated SUMO. These trajectories are also analyzed along the complex structures described above.

1.2.3. Brief Summary of methods

Molecular dynamics (MD) simulations are run on all structures using *MD package Amber 8*. The conformations in time from MD trajectories are analyzed to study the dynamics of the structures with respect to mean-square fluctuations and correlation between the fluctuations. The pseudo-dihedral angle fluctuations and time-delayed auto-correlations of virtual bond vectors reflect information about the flexibilities of the structures. The interactions between the residues throughout the simulations are followed and tried to be characterized with respect to the function.

In the course of the analysis of the trajectories, the conformations throughout are clustered to identify unique conformers of the structure. In the clustering, the root mean square deviation is used as a similarity measure. The conformations and the distributions of the members of defined clusters in time are examined for the characterization of distinct conformations spanned during simulations. All the analysis conducted is carried out on the basis of the whole simulations as well as for the time intervals suggested by the clustering analysis.

Additional to the dynamic analysis, both crystal structures and selected snapshots from the MD trajectories are analyzed by Gaussian Network Models (GNM). Eliminating the noise originating from the atomistic simulations, this step of analysis further emphasis some interactions, which are not apparent from the direct analysis of the conformational ensemble of the simulations.

1.2.4. Contributions to Current Literature

In this study, how the RanBP2 binding may reduce the conformational space of Ubc9-SUMO complex in favour to increase the sumoylation activity is searched. The analysis of the molecular dynamics simulation trajectories by various means reveal that the orientation of SUMO with respect to the Ubc9 is different between E3 bound and unbound states. The implication of this on the conformational dynamics of both Ubc9 and SUMO are elaborated and the differences in residue interaction networks of the two complexes are characterized with respect to the RANBP2 binding.

Specifically, the residues Asp100 and Lys101, which are functional in target recognition, are observed to have more restricted mobility/flexibility in the RanBP2 bound complex. Also, more stabilized correlated fluctuations between the residues associated with RanBP2, including the latter residues, and of the catalytic residues are observed. The effect of E3 binding on the the additional binding site of Ubc9 provide evidence for the E3 ligase activity of RanBP2 on the sumoylation target RanGAP1; this is a matter of debate in the current literature (Bernier-Villamor *et al.*, 2002; Reverter and Lima, 2005).

Furthermore, the analysis of the dynamics Ubc9-SUMO complex reveals some structural motion of the complex structure, which may mark the Ubc9-SUMO complex for interaction with RanBP2.

2. MATERIALS AND METHODS

2.1. Studied Structures of the Sumoylation Mechanism

The structures studied in this work are the structures of un-complexed Ubc9 protein, un-complexed SUMO protein, Ubc9-SUMO complex and Ubc9-SUMO - RanBP2 complex.

The structure of human Ubc9 is available with the PDB code 1A3S and human SUMO-1 is available with PDB code 1A5R. The molecular dynamics simulations of these proteins provided the data to be taken as the basis in deducing the effects of RanBP2 binding to Ubc9-SUMO complex.

The crystal structure of Ubc9 (E2 enzyme)-SUMO-RanGAP1 (sumoylation target)-RanBP2 (E3 enzyme) complex is available with the PDB code 1Z5S (Figure 2.1). The modeled structures for the intermediate complexes Ubc9-SUMO (Figure 2.2 (a)) and Ubc9-SUMO-RanBP2 (Figure 2.3) are extracted from this complex structure. In the available crystal, SUMO is bound to the acceptor lysine of target protein RanGAP1. The major interests of this study are on the steps prior to conjunction of SUMO to the target proteins. Therefore, RanGAP1 is eliminated from the structures, and the thioester bond between Ubc9 active cysteine and SUMO C - terminal glycine is modeled (Figure 2.2 (b)). The dynamics of the structures are studied via the trajectories obtained from Molecular Dynamics Simulations.

The structures and the simulations are summarized in Table 2.1. The details of bond modeling are explained in Section 2.2 and the parameters of molecular dynamics simulations are given in Section 2.3.

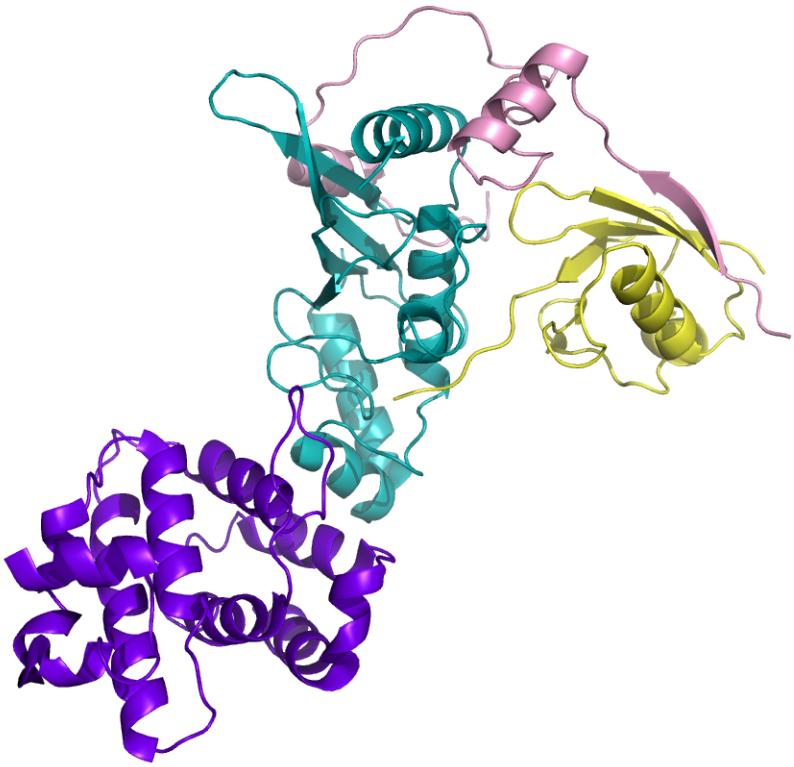

Figure 2.1. Crystal structure of the Ubc9-SUMO-RanGAP1-RanBP2 complex. Ubc9 (blue) is the SUMO E2 enzyme, SUMO-1 (yellow) is the studied SUMO family member, RanGAP1 (purple) is the target and RanBP2 (pink) is an E3 enzyme.

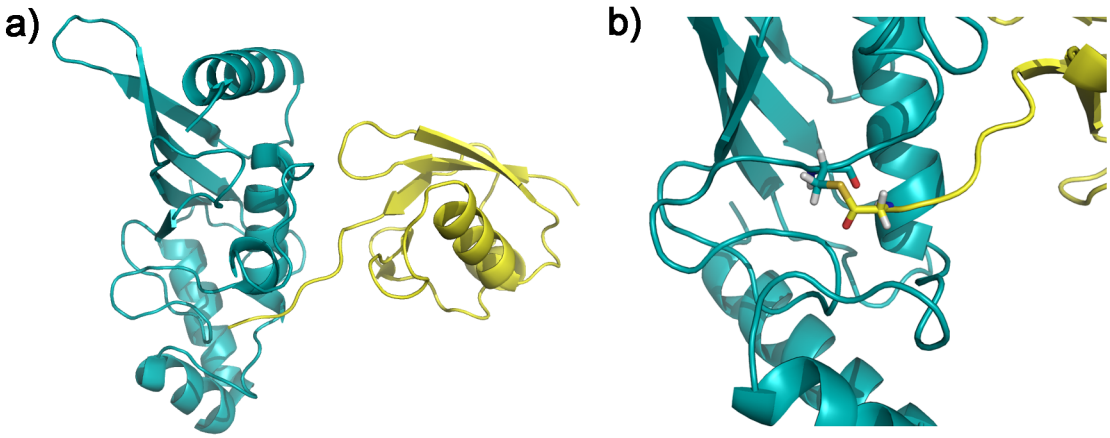

Figure 2.2. a) Structure of the Ubc9(blue)-SUMO(yellow) complex. The structure is extracted from the crystal structure Ubc9-SUMO-RanGAP1-RanBP2. b) The thioester bond between Ubc9 Cys93 and SUMO Gly97. The structure is a snapshot from MD simulations of Ubc9-SUMO complex at 1 ns.

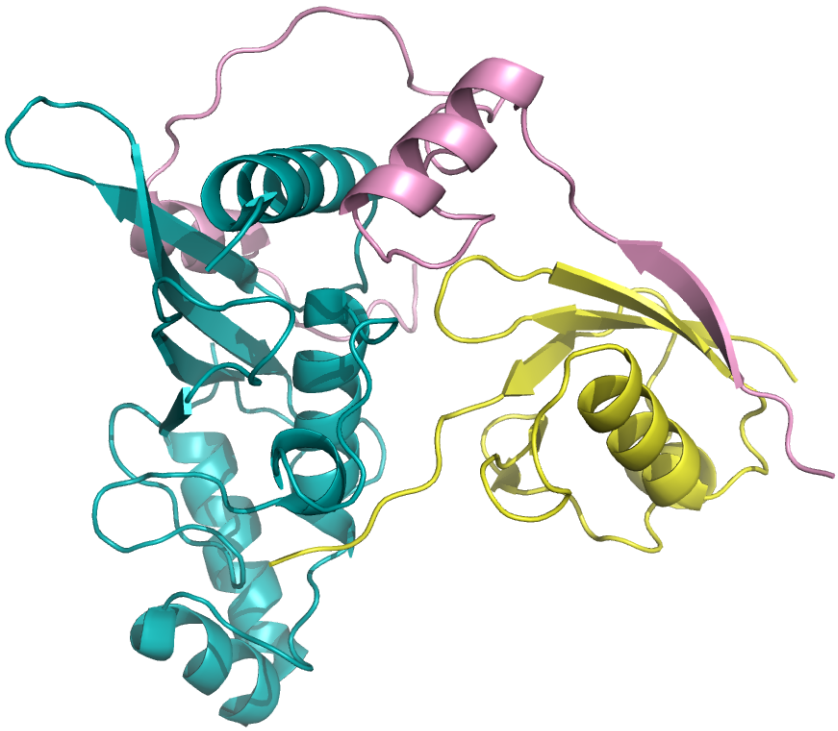

Figure 2.3. Structure of the Ubc9(blue)-SUMO(yellow)-RanBP2(pink) complex. The structure is extracted from the crystal structure of Ubc9-SUMO-RanGAP1-RanBP2 complex.

Table 2.1. Structure Details and Simulation Lengths for Analyzed Complexes

Structure Elements	Source PDB	Modification	Simulation Length
Ubc9	1A3S	No	32.5 ns
SUMO	1A5R	No	50 ns
Ubc9-SUMO	1Z5S	Yes	58 ns
Ubc9-SUMO-RanBP2	1Z5S	Yes	50 ns

2.2. Modeling of the Thioester Bond

The bond to be modeled in the Molecular dynamics simulations is a thioester bond between Cys93 of Ubc9 and C-terminal Gly97 of SUMO. For defining the bond between these residues, first residue types with desired atoms are defined for both residues. The parameters are either taken from the classical residue types in Amber, or generated using *General Amber Force Field* (gaff) in conjunction with antechamber program (Case *et al.*, 2004; Case *et al.*, 2005; Wang *et al.*, 2004). Along the way, the program Gaussian is used for geometry minimization.

For the Cys residue in thioester bond, the atom types and parameters, the point charges, van der Waals parameters, are taken from the residue type "Cyx", defined for Cysteines taking part in disulfide bond formation. The residue Gly97 is an C-terminal residue and Amber package automatically includes the -COOH group to this residue. For definition of the thioester bond, this addition inappropriate. For the Gly residue taking part in the thioester bond, a new residue type "Glb", standing for Glycine-bonded, is defined, which does not include the -COOH group. The atom types and parameters, the point charges, van der Waals parameters, are taken from the residue type "Gly" for this new residue type.

In the crystal structure, the residues are not in a distance close enough for a covalent bond. For gaff and antechamber to be able to recognize the bond between these defined residues, it is necessary to reduce the distance. For this purpose, Gaussian program is used. Using the Gaussian method (Becke, 1988; Becke, 1993a; Becke, 1993b; Lee *et al.*, 1988) B3LYP/6-31+G**, on the di-peptide constructed from the generated residues, the residues are brought into close distance. The necessary parameters are generated by gaff and antechamber using the minimized peptide.

The residues Cyx93 and Glb97 are bonded by tleap in Amber. Stretch, bend and torsion parameters generated by the method described above are used for the defined bond.

2.3. Molecular Dynamics Simulations

2.3.1. Theoretical Background

The methodology of Molecular Dynamics Simulations is based on Newton's second law,

$$F_i = m_i a_i \qquad or \qquad \frac{F_i}{m_i} = \frac{d^2 x_i}{dt^2} \tag{2.1}$$

where F_i is the force exerted on the particle i, m_i is its mass and a_i is its acceleration. Starting with the knowledge of the force on each atom, determination of the acceleration of each atom in the system is possible. Integration of the second form of the equation yields a trajectory that describes the positions, velocities and accelerations of the particles as they vary with time. The method is deterministic; meaning that knowing the positions and velocities of each atom, prediction of the state of the system at any time, future or past, is possible. Using this methodology, successive configurations of the system can be generated. In intermolecular interactions, the force applied on each particle changes whenever the particle itself of any other particle in interaction with it changes its position. This feature is implemented to the simulation with the use of continuous potential, in which the motions of all particles are coupled together (Leach, 2001). At this stage, since molecular systems generally consist of a vast number of particles, it becomes impossible to find the properties of such a complex systems analytically. Therefore, MD simulation integrates the equations by using numerical methods.

The force exerted on each particle at any time during simulation, F_i is calculated from a forcefield in MD simulations. The forcefield defines the potential energy of a system as a function of the atomic positions/coordinates, and the F_i are obtained from

derivatives of the potential function.

$$V_i(\mathrm{R}_1, ..., \mathrm{R}_N) = \sum_{bonds} \frac{k_i}{2}(l_i - l_{i,0})^2 + \sum_{angles} \frac{k_i}{2}(\theta_i - \theta_{i,0})^2 + \sum_{torsions} \frac{V_n}{2}(1 + cos(nw - \gamma)) + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right) \quad (2.2)$$

The potential energy of particle i as a function of all R_j, is defined in terms of interactions between bonded atoms, bond angle and torsional angle potentials, and electrostatic and van der Waals interactions between non-bonded atoms. The first term in the equation describes the interaction of pairs of bonded atoms, where l_i is the bond length. The second term is similarly the summation over all the angles in the molecule modeled using a harmonic potential, where θ_i is the angle between the three successive atoms. Torsional potential describes the change in energy when a bond rotates, and is depicted with the third term in the equation. The fourth contribution in the equation is for the non-bonded atoms, which are separated by at least three atoms. The non-bonded interactions are defined by two different potentials. The former one is the Lennard-Jones 12-6 potential function that accounts for van der Waals interactions, whereas the latter one is the Coulomb potential for electrostatic interactions (Leach, 2001). From the potential function, the F_i are generated as:

$$\mathrm{F}_i = - \bigtriangledown V_i(\mathrm{R}_1, ..., \mathrm{R}_N) = -\frac{\partial V(\mathrm{R}_1, ..., \mathrm{R}_N)}{\partial \mathrm{R}_i} \quad (2.3)$$

Knowing the state of the structure at time t, the positions, velocities and accelerations are approximated by Taylor series (Equation 2.4), where a is the acceleration

(2nd derivative), b is the 3rd derivative etc (Leach, 2001).

$$\begin{aligned} r(t+\delta t) &= r(t) + v(t)\delta t + \tfrac{1}{2}a(t)\delta t^2 + ... \\ v(t+\delta t) &= v(t) + a(t)\delta t + \tfrac{1}{2}b(t)\delta t^2 + ... \\ a(t+\delta t) &= a(t) + b(t)\delta t + ... \end{aligned} \tag{2.4}$$

The most common algorithm used in integration is Verlet algorithm. Positions and accelerations at a time t and positions from time $(t-\delta t)$ are used to calculate new positions at time $(t+\delta t)$. Then the velocities can be obtained from the difference in positions. From generated values, new positions can be obtained successively (Leach, 2001). Giving the appropriate equations:

$$\underbrace{r(t+\delta t) = r(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2 \;;\; r(t-\delta t) = r(t) - v(t)\delta t + \frac{1}{2}a(t)\delta t^2}$$

$$r(t+\delta t) = 2r(t) - r(t-\delta t) + \tfrac{1}{2}a(t)\delta t^2 \tag{2.5}$$

$$v(t) = \tfrac{\mathrm{r}(t+\delta t)+r(t-\delta t)}{2\delta t}$$

As can be seen from the structure of the algorithm, to be able to start a MD simulation and generate configurations of the system, the initial state should be defined. Either experimental or theoretical inputs can be used at this stage. In addition to the positions of the elements of the network, the initial velocities should also be defined.

Experimental inputs are generally used for the atomistic coordinates of a structure, in the form of X-ray or NMR structure. The structure is subjected to energy minimization before the starting the simulation.

The initial velocities are generated by theoretical methods. A commonly used

method is randomly selecting the initial velocities from a Maxwell-Boltzman distribution, for the specified temperature (Leach, 2001).

$$p(v_{ix}) = \left(\frac{m_i}{2\pi k_B T}\right)^{\frac{1}{2}} exp - \frac{1}{2}\frac{m_i v_{ix}^2}{k_B T} \tag{2.6}$$

2.3.2. The Molecular Dynamics Simulation Parameters

Molecular dynamics (MD) simulations are run on all structures using *MD package Amber 8* (Case *et al.*, 2004; Case *et al.*, 2005). During simulation, the forcefield ff03 was used, with explicit solvent, using truncated octahedron box and TIP3P water model (Jorgensen *et al.*, 1983). Periodic boundary conditions at constant pressure are preferred, with isotropic position scaling. Electrostatic energy of a periodic box is calculated using The Particle Mesh Ewald (PME) method (Essman *et al.*, 1995). The nonbonded cutoff is selected as 9 $\mathring{A}$. The SHAKE algorithm is used and bond interactions involving H-atoms are omitted (Ryckaert *et al.*, 1977).After the minimization, the velocities are generated first at 10 K from a Maxwellian distribution, the temperature was gradually raised to, and maintained at 300 K by the Berendsen weak-coupling algorithm (Berendsen *et al.*, 1984). The simulations are run at 0.5 ns steps, the energies and temperatures are recorded every 0.01 picosecond and the coordinates of the simulation are recorded at every 0.1 picosecond.

2.4. Clustering Analysis

MD is used to generate a high lumber of different conformations of the structures, many of which are close conformations. The large number of conformations should be sampled to select distinct representative conformations. In this work k-means clustering is used for the sampling.

In k-means clustering, a subset of the members of the data set to be clustered re selected randomly, and are assigned as centroid of distinct clusters. Using a similarity measure, the distance of each member the data set to each constructed cluster is

calculated. The member of the data set is attributed to the cluster, to which it is the closest. When the members of a cluster are updates, a new member is added to that cluster, the centroid is recalculated as the mean of its members. With this iterative procedure, the data set is grouped into clusters according to the selected similarity measure (Alpaydın, 2004).

In this study, the data set to be clustered are the generated conformations, and the similarity measure is the similarity matrix constructed with the root mean square deviation from the cluster centroid (Equation 2.7). The number of generated clusters is determined by setting a threshold for the rmsd. Clusters are generated using 2 $\mathring{A}$, 1.7 $\mathring{A}$ and 1.5 $\mathring{A}$. The clustering is performed using MMTSB Toolset's *kclust* utility (Feig *et al.*, 2004).

$$rmsd = \sqrt{\frac{\sum_{i=1}^{N}(d_i^2)}{N}} \tag{2.7}$$

2.5. Mean Square Fluctuations of Positions

The mean square fluctuations of positions are calculated as the deviation from the initial structure, averaged through the selected regions of simulation (Equation 2.8). The conformations are aligned before calculation of fluctuations, the alignment can be done over the whole structure, as well as a pre-defined subgroups of atoms in the structure, such as SUMO chain in Ubc9-SUMO structure. While defining the time period of simulations to be included in calculations, the equilibration period, defined from the root mean square deviations of residues from initial structure, are eliminated.

$$\Delta R_i = \langle R_i(t) - R_i^0 \rangle \tag{2.8}$$

where ΔR_i is the mean square fluctuation of residue i, and R_i^0 is the initial position of residue i. The brackets represent time averages over recorded snapshots. The results are represented in the form of B-factors, which is $(8/3)\pi^2$ times the mean square fluctuation.

The normalized orientational cross-correlations between residue pairs are defined as

$$CO_{i,j} = \frac{\langle \Delta R_i \Delta R_j \rangle}{\langle \Delta R_i^2 \rangle^{1/2} \langle \Delta R_j^2 \rangle^{1/2}} \tag{2.9}$$

where ΔR_i is the fluctuation in the position vector R_i of site i and ΔR_j is the fluctuation in the position vector R_j of site j. The cross-correlations vary in the range [-1, 1] with the lower and upper limit indicating fully anti-correlated and correlated fluctuations, respectively. The mean square fluctuation correlations are calculated for the total length of the simulations, excluding the equilibration periods. Additionally, the correlations in the time intervals defined by clustering analysis are calculated explicitly.

2.6. Pseudo-Dihedral Angle Fluctuations

The pseudo-dihedral angles are defined on the basis of virtual bonds between α carbons. Schematic presentation of the virtual bond model for a protein segment between backbone units: $C^{\alpha}_{(i-2)}$ and $C^{\alpha}_{(i+1)}$ is given in Figure 2.4. ϕ_i is the rotational angle of the i^{th} virtual bond, defined by the respective locations of the four backbone units $C^{\alpha}_{(i-2)}$,$C^{\alpha}_{(i-1)}$, $C^{\alpha}_{(i)}$, $C^{\alpha}_{(i+1)}$.

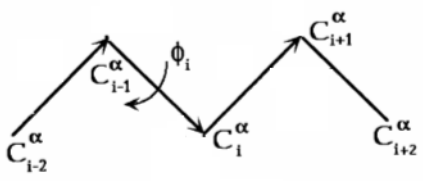

Figure 2.4. Pseudo-dihedral Angle Model

The means of fluctuations of pseudo-dihedral angles from the mean value during simulation are utilized for determining the flexibilities of the bonds (Equation 2.10).

$$\Delta\phi_i = \langle \phi_i(t) - \overline{\phi_i} \rangle \tag{2.10}$$

2.7. Virtual Bond Vector Correlations

The virtual bond vectors are defined as the normalized vector M_i from the α carbon $C^{\alpha}_{(i-1)}$ of residue *i-1* to the α carbon $C^{\alpha}_{(i)}$ of residue i (Figure 2.5).

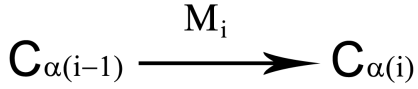

Figure 2.5. Virtual Bond Vector

The normalized time-delayed auto-correlations of virtual bond vectors are defined as

$$CM_i(\tau) = \langle M_i(t) M_i(t+\tau) \rangle \tag{2.11}$$

where $M_i(t)$ is the virtual bond vector i at time t and $M_i(t+\tau)$ is virtual bond vector i at time $t+\tau$. The brackets represent time averages over recorded snapshots. The virtual bond vectors are unit vectors by definition, therefore there is no need additional normalization in calculation of correlations in Equation 2.11. The auto-correlations are in the range [-1, 1] with the lower and upper limit indicating fully anti-correlated and correlated virtual bond positions, respectively. $\tau = 0$ gives the equal-time auto-correlations, which is 1 for all virtual bond vectors. The correlations are calculated for several time delays (τ), starting from 0 and ending at 30 ns. The highest value of the time delay (30 ns) is selected to be higher than half simulation times, which are 50 and 58 ns for Ubc9-SUMO-RanBP2 and Ubc9-SUMO complexes, respectively.

2.8. Gaussian Network Model

Elastic network models are models constructed with the aim of studying the dynamic behaviors of proteins. In elastic network models the C^{α} of amino acids constructing the protein are taken as nodes, and each node is connected to the neighboring nodes, within a defined cutoff distance, with a spring. The cutoff is usually defined as 7.0 $\mathring{A}$, which corresponds to the first coordination shell radius in protein structures (Rader *et al.*, 2006). This model does not include detailed chemical information, dynamics are defined only using the topology of a structure, but it makes it possible to observe collective motions in large structures (Chennubhotla *et al.*, 2005; Rader *et al.*, 2006). Another advantage of Elastic Network Models is that they require less computational time compared to methods like molecular dynamics, while providing information on the large scale motions.

The equilibrium position vector and the instantaneous position vector of the i^{th} node of the network can be used to define its fluctuations as:

$$\Delta R_i = R_i - R_i^0 \tag{2.12}$$

Then the fluctuations in the distance vector between two residues i and j can be obtained using their individual fluctuations (Bahar *et al.*, 1997):

$$\Delta R_{ij} = R_{ij} - R_{ii}^0 = \Delta R_j - \Delta R_i \tag{2.13}$$

Assuming isotropic and Gaussian fluctuations, and decomposing the position vector into its components, the following potential can be defined for the network composed of N nodes:

$$V_{GNM} = \frac{\gamma}{2}\left[\sum_{i,j}^{N} \Gamma_{ij}[(\Delta X_i - \Delta X_j)^2 + (\Delta Y_i - \Delta Y_j)^2 + (\Delta Z_i - \Delta Z_j)^2]\right] \tag{2.14}$$

Where Γ in Equation 2.14 defines the connectivity (Kirchhoff) matrix, and ij^{th} element is defines as in Equation 2.15:

$$\Gamma_{ij} = \begin{cases} -1 & if\ i \neq j\ and\ R_{ij} \leq\ r_c \\ 0 & if\ i \neq j\ and\ R_{ij} >\ r_c \\ -\sum_{j,j\neq i} & if\ i = j \end{cases} \tag{2.15}$$

In Equation 2.14, γ represents the force constant of the springs of the network. It is taken to be uniform for all springs, indicating no distinction between bonded and non-bonded residues inside the cutoff distance.

The main interest in constructing the model is on mean square fluctuations of a single residue, and correlated fluctuations of residues. The formulations are given respectively in Equations 2.16 and 2.17.

$$\langle \Delta R_i . \Delta R_i \rangle = \langle \Delta X_i^2 \rangle + \langle \Delta Y_i^2 \rangle + \langle \Delta Z_i^2 \rangle \tag{2.16}$$

$$\langle \Delta R_i . \Delta R_j \rangle = \langle \Delta X_i \Delta X_j \rangle + \langle \Delta Y_i \Delta Y_j \rangle + \langle \Delta Z_i \Delta Z_j \rangle \tag{2.17}$$

Including the isotropic and Gaussian probability distributions, after necessary calculations and substitutions, the fluctuations can be defined by Equations 2.18 and 2.19

$$\langle \Delta R_i^2 \rangle = \frac{3k_B T}{\gamma} (\Gamma^{-1})_{ii} \tag{2.18}$$

$$\langle \Delta R_i . \Delta R_j \rangle = \frac{3k_B T}{\gamma} (\Gamma^{-1})_{ij} \tag{2.19}$$

At this level, the connectivity matrix is decomposed into three matrices: U for eigenvectors u_i, Λ for eigenvalues λ_i and the transpose of U (Equation 2.20)

$$\Gamma = U \Lambda U^T \tag{2.20}$$

In this representation, the eigenvalues give the frequencies of individual modes (defining fast or slow modes), and eigenvectors gives the shapes of the modes. Fast modes will point out residues that may be functionally important, while slow (low frequency) modes will give clues about the global motion (Rader *et al.*, 2006).

2.9. Contact Map Construction

The contact maps are plots of contact matrices, generated from the pairwise distances of C^α of residues. The maps demonstrate the residues that are in contact distance, according to a pre-defined cut-off distance. The distances between residue

pairs for a snapshot are calculated as:

$$d_{i,j} = |R_i - R_j| \tag{2.21}$$

where R_i is the position of C_i^α of residue i, R_j is the position of C_j^α of residue j and $d_{i,j}$ is the distance between these atoms.

Selecting a cut-off distance, the contact matrix is constructed. The definition of elements of the contact matrix is given in Equation 2.22.

$$Cont_(i,j) = \begin{cases} 1 & if\ d_{i,j} \leq\ cut-off \\ 0 & if\ d_{i,j} >\ cut-off \end{cases} \tag{2.22}$$

3. RESULTS AND DISCUSSION

3.1. Mobility and Flexibility of Structures from MD Simulation Trajectories

3.1.1. Root Mean Square Deviations from Initial Structures

Root-mean-square deviation (rmsd) of a conformation is calculated by aligning its alpha-carbon backbone to the initial structure of MD simulation and then calculating the deviation of each atom from its initial position as averaged over all atoms. 500 snapshots per nanosecond are used in calculations.

Rmsd values are used not only to reflect an overall view of the extent of motion of the protein in the course of simulations, but also to determine the equilibration period. In this section, isolated Ubc9 is shown to be a rigid protein as compared to isolated SUMO. The equilibration periods for all structures, isolated Ubc9, isolated SUMO, Ubc9-SUMO complex, and Ubc9-SUMO-RanBP2 complex, are determined.For isolated SUMO only, a subgroup of residues is determined for the alignment of the generated conformations, for the following calculations. The change in the orientation of the chains with respect to each other is detected in Ubc9-SUMO complex. RanBP2, nevertheless, is shown to eliminate this orientation change in Ubc9-SUMO-RanBP2 complex.

3.1.1.1. Isolated Ubc9 and SUMO proteins. The isolated Ubc9 has an rmsd of around 2 $\AA$ throughout the 32.5 ns simulation. There are no significant fluctuations in rmsd except for the slight increase around 9 and 20 ns of the simulations (rmsd below 2.5 $\AA$) (Figure 3.1). Rmsd of Ubc9 does not show significant changes in the rmsd values in a time correlated manner.

The simulations of isolated Ubc9 are to provide a basis for the comparison of

Ubc9-SUMO and Ubc9-SUMO-E3 complexes. The changes in the behavior of Ubc9 by the complex formation with SUMO and SUMO-E3 are followed throughout the simulations. The equilibration period is takes as 2.5 ns in the analysis of Ubc9 MD trajectories.

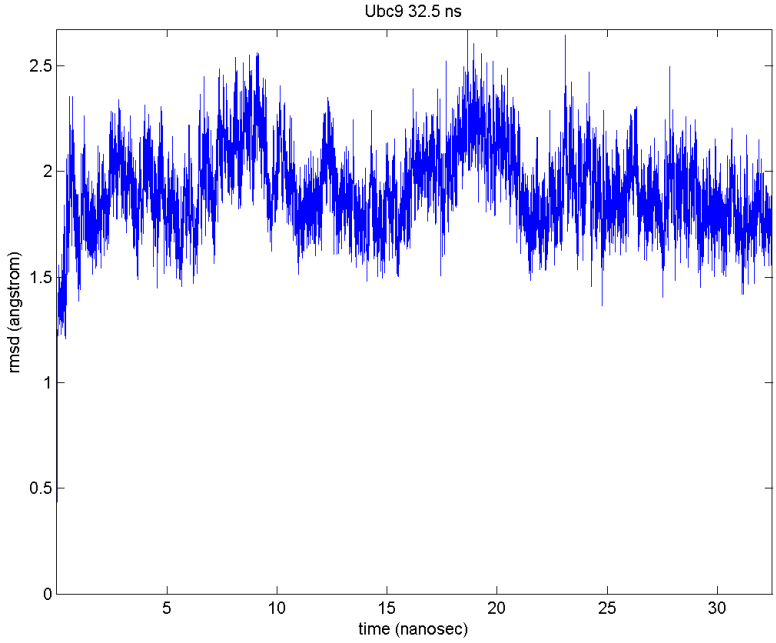

Figure 3.1. Isolated Ubc9 Root Mean Square Deviations from the initial structure throughout the trajectory. The stabilization of the protein corresponds to 2.5 ns

SUMO in its complex structures does not have its mobile N-terminus tail. On the other hand, this region together with the C-terminus HSTV residues cleaved at maturation exists in its isolated structure (PDB code: 1A5R). The high mobility of this N-terminus tail is reflected in the rmsd values of the isolated SUMO simulations (Figure 3.2, red). As the high mobility of this loop region overweight the fluctuations, the rmsd values are also calculated without this very highly mobile region. The N-terminal loop region, first 21 residues (from -1 to 19), and also four C-terminal residues are excluded in the analysis. This is also to have a correspondence between SUMO in isolated and in complex states.

The rmsd of truncated SUMO is significantly smaller, yet still being around 3 $\AA$. Nevertheless, the core region of the protein is significantly less mobile when compared to the full length SUMO. The course of the rmsd values suggests that first 4-5 ns is the

equilibration period. The trajectory of 5-35 ns, which reflects a dynamic equilibrium in the fluctuations, is utilized in the present analysis.

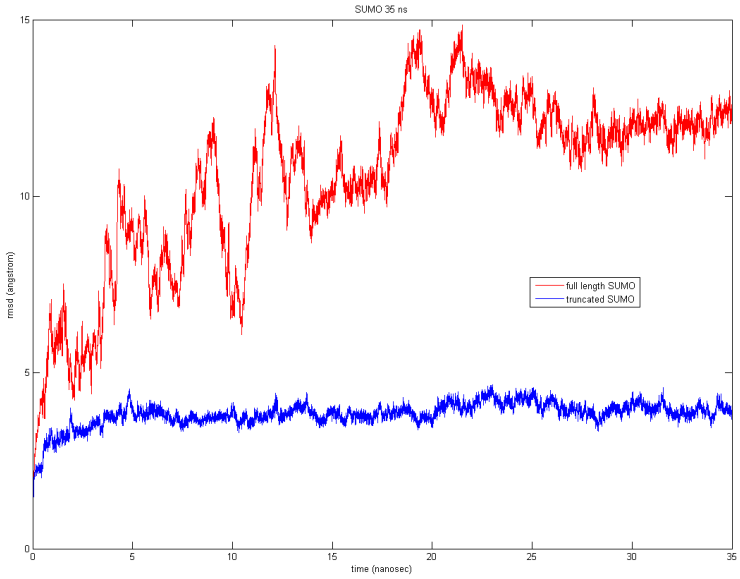

Figure 3.2. Isolated SUMO Root Mean Square Deviations from the initial structure throughout the trajectory.The rmsd for full length SUMO (red) and the rmsd with elimination of the N-terminal unstructured tail (blue).

3.1.1.2. Ubc9-SUMO complex. The Ubc9-SUMO complex structure displays a large amplitude change in its rmsd values between 5 and 12 ns. (Figure 3.3). This is due to the fact that the Ubc9-SUMO complex structure is extracted from a complex structure with additional members, namely the sumoylation target, RanGAP1, and the E3 enzyme, RanBP2, (PDB code: 1Z5S). It is thus expected to observe longer equilibration periods and high rmsd values from the initial structure. The high rmsd values could be due to the internal relaxation of individual proteins and as well as the relative motion of the two proteins with respect to each other in the absence of other components of the complex structure. To elaborate the source of the deviation, the Ubc9 and SUMO chains are aligned individually and then the rmsd values are recalculated (Figure 3.4).

It is seen that both chains of Ubc9-SUMO complex are quite stable with relatively small jumps between 25-40 ns. This indicates that the deviation is mainly caused by

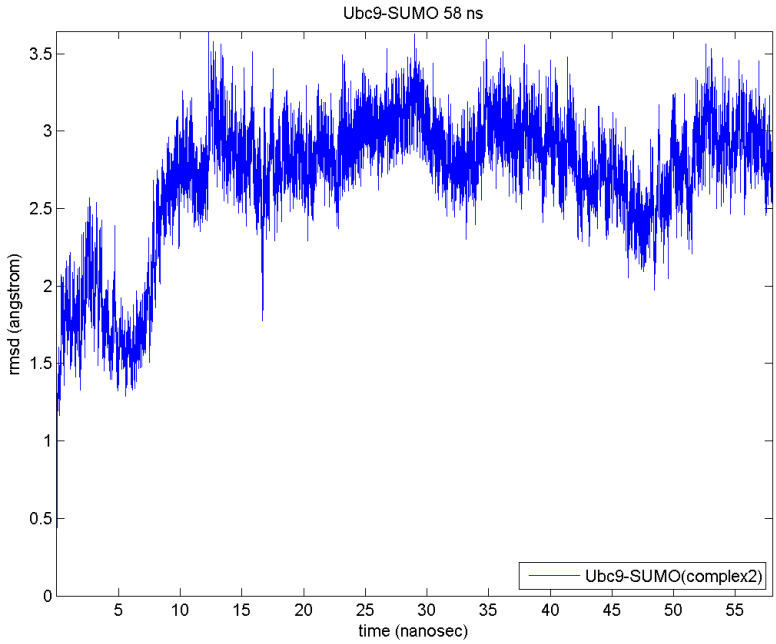

Figure 3.3. Ubc9-SUMO rmsd from the initial structure throughout the trajectory. The increase between 5 and 12 ns indicates a significant change in structures of the proteins. The origin of this jump is shown to be the change in quaternary structure.

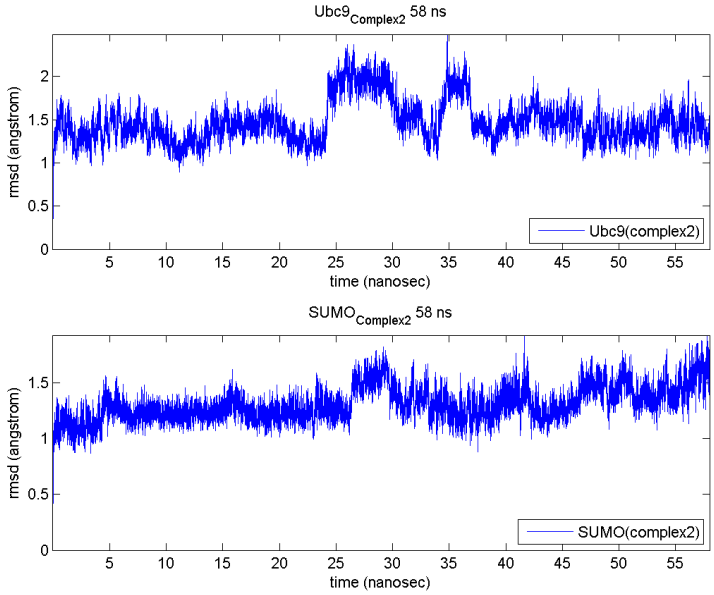

Figure 3.4. Individually Aligned Ubc9 and SUMO from Ubc9-SUMO complex, rmsd from the initial structures throughout the trajectory. The sudden increase in the rmsd is not observed in rmsd values generated from individually alignment of chains.

an alteration in quaternary structure of the complex between 5 and 12 ns, rather than the changes in the secondary or tertiary structures.

The snapshots of the trajectory points out the change in SUMO orientation in the complex (Figure 3.5). In the figure The structures of the complex at 5 ns (green) and 12 ns (blue) are displayed. The structures are aligned via the Ubc9 structure and color coding is applied to SUMO only, to demonstrate the SUMO motion clearly. While there are fluctuations in its position; SUMO, once moved, does not go back to its orientation in its crystal state during the 58 ns simulation. The jumps in the rmsd values of individually aligned chains originate from the motion of specific loops, the details of which will be discussed in following sections.

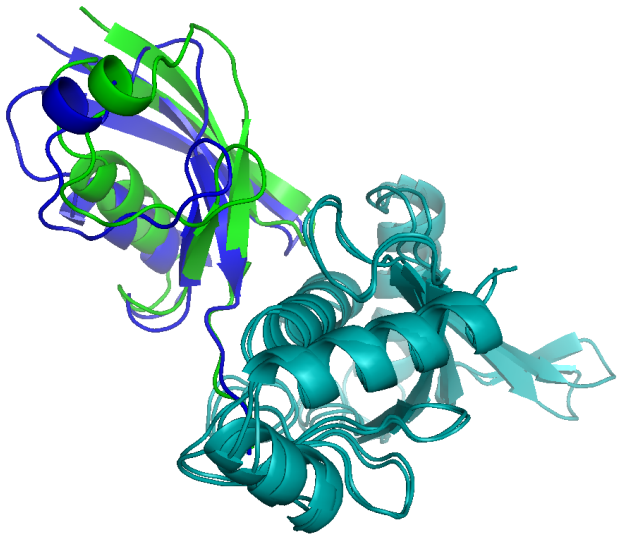

Figure 3.5. Snapshots form Ubc9-SUMO simulation at 5 ns and 12 ns. The structure at 5 ns and at 12 ns are represented in green and blue, respectively. The structures are aligned via Ubc9 chains, and color coding is applied to SUMO chains only.

The Ubc9-SUMO complex is expected to have a long equilibration interval. From the profile of rmds values in time, initial 12 ns is taken as the equilibration period. Nevertheless, as the change in the quaternary structure with the absence of E3 is also of interest, the initial period after 0.5 ns is also incorporated for some analysis, as will be discussed below.

3.1.1.3. Ubc9-SUMO-RanBP2 complex. The dynamics of Ubc9-SUMO-RanBP2 complex, as compared to Ubc9-SUMO complex, does not display jumps in positions of some residues with respect to the initial structure (Figure 3.6). Similarly, the individually aligned chains in the complex structure does not reveal significant jumps (Figure 3.7) during the 50 ns trajectory. The equilibration period is considered here as 0.4 ns.

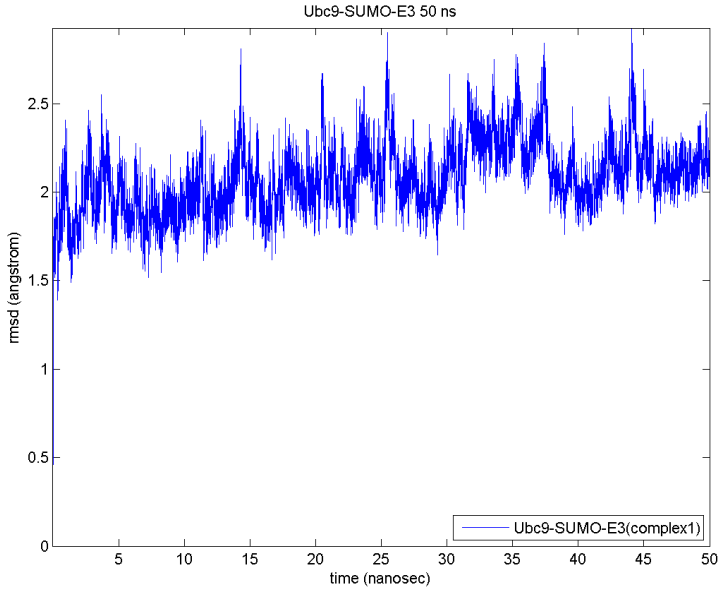

Figure 3.6. Ubc9-SUMO-RanBP2 Root Mean Square Deviations from the initial structure throughout the trajectory.

Snapshots from the MD trajectory also show that in the Ubc9-SUMO-RanBP2 complex, SUMO does not go through the orientational change observed in Ubc9-SUMO complex. The loop regions are mobile, but the orientation of the secondary structural elements are more stable (Figure 3.8). In fact, although not significant, the change in the orientation seen for SUMO in Ubc9-SUMO RanBP2 complex is in the opposite direction of the motion of SUMO in Ubc9-SUMO complex (Figure 3.9).

The spanned conformations of the covalently bound Ubc9-SUMO are limited in the existence of RanBP2. This is reflected in the profile of the rmds values. The conservation of the orientation of SUMO with respect to Ubc9 with the simulations of Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes is consistent with the proposed role of RanBP2; the restriction in the conformational/configurational space of the Ubc9-

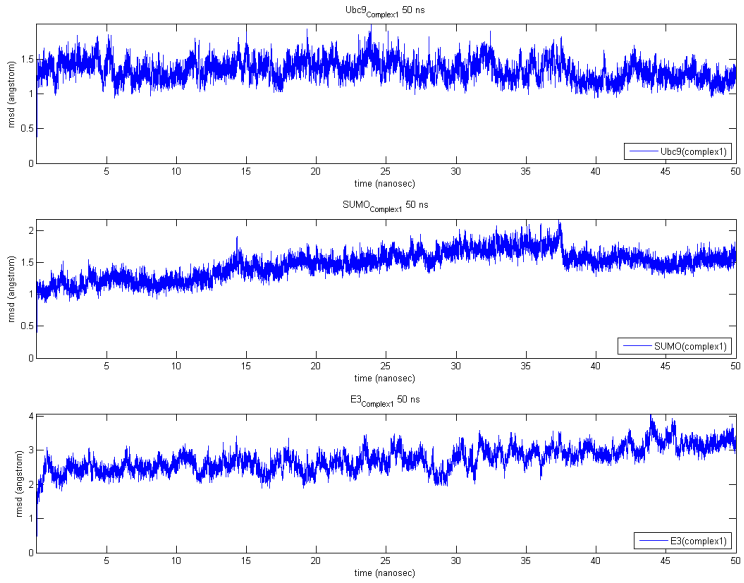

Figure 3.7. Individually Aligned Ubc9 (Upper lane), SUMO (Middle Lane) and RanBP2 (Lower Lane) rmsds from Ubc9-SUMO-RanBP2 complex.

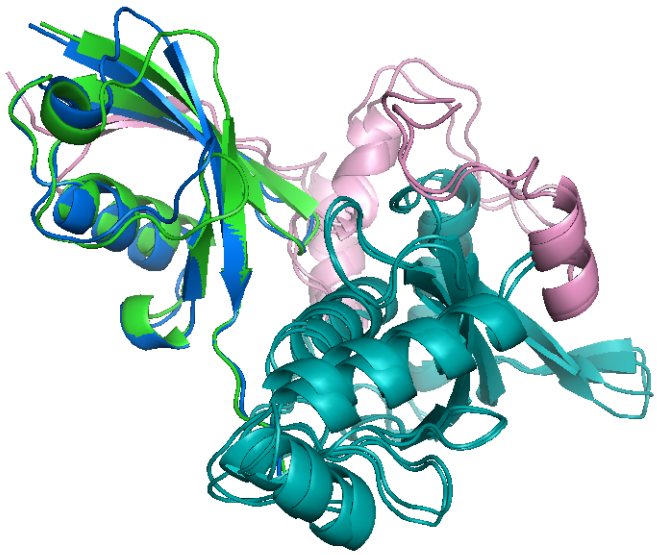

Figure 3.8. Snapshots from Ubc9-SUMO-RanBP2 simulation at 6.7 and 44.7 ns. Ubc9 is represented in light blue, RanBP2 is represented in pink in both snapshots. The structure of SUMO at 6.7 ns and at 44.7 ns are represented in green and blue, respectively. The structures are aligned via Ubc9 chains.

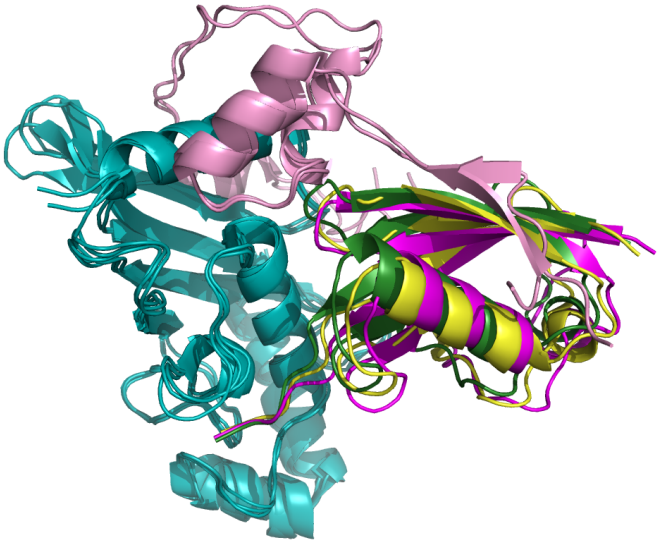

Figure 3.9. Orientation of SUMO snapshots of Ubc9(blue)-SUMO(magenta)-RanBP2(pink) (10 ns) and Ubc9(blue)-SUMO(green) (15 ns) complexes. The structures are aligned via Ubc9 chains.

SUMO complex and thus prevention of the non-productive conformations of Ubc9-SUMO (Reverter and Lima, 2005). This is further investigated with respect to the interactions of residues and correlations between the fluctuations of residues.

3.1.2. Mean-square fluctuations of Ubc9 and SUMO in Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes

The mean-square fluctuation of each residue in Ubc9-SUMO complex structure is calculated within a time window of 57.5 ns and 49.6 ns, excluding the initial 0.5 and 0.4 ns with and without RanBP2, respectively. The peaks in the distribution of fluctuations suggest mobile regions of the structure, whereas the minimum fluctuating regions may point to the hinge sites of the structure. Further, the differences in the fluctuations with and without RanBP2 reveals the residues that are affected from binding and may provide clues for the underlying mechanism of Ubc9 and RanBP2 functioning. For simplicity, Ubc9-SUMO-RanBP2 complex is referred as "complex1" and Ubc9-SUMO complex is referred as "complex2" in figure legends.

In this section of the study, Ubc9 residues between Val27 and Met39 are shown to have increased mobility upon SUMO binding, in Ubc9-SUMO complex. This mobility is again reduced with the RanBP2 binding in Ubc9-SUMO-RanBP2 complex. This may be related to the motivation of RanBP2 binding to Ubc9-SUMO complex; this will be discussed below. The binding of RanBP2 to Ubc9-SUMO complex is also shown to reduce the mobility of the region between Glu98 and Asp102 of Ubc9. This region contains residues Asp100 and Lys101, which are functional in target recognition. This region is far from RanBP2 binding sites and thus the reduced mobility suggests an allosteric effect of RanBP2 on Ubc9 that enhances the sumoylation efficiency. The mean square fluctuations point out an overall stabilization of SUMO protein in both Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes.

3.1.2.1. Individual Ubc9 and SUMO. The fluctuations of residues of SUMO are calculated for both full-length protein and for a subgroup of residues, eliminating 21 N-terminal residues (from -1 to 19) and 4 C-terminal residues form calculation. The alignment of the conformations of the MD trajectory is affected by the mobility of N-terminus, and results for full length SUMO are misleading (Figure 3.10). The fluctuations calculated by excluding these mobile residues are used in the discussion of the following sections. As seen, SUMO is highly mobile; with the peaks around residues His35, Lys45-Gln55, and Glu85. The experimental temperature factors are not available for the comparison with the calculated mean-square fluctuations. The available experimental structure of SUMO is determined by NMR (Bayer *et al.*, 1998).

Unlike SUMO, Ubc9 is a rather rigid protein with only peaks around residues His20, Lys30, Pro80- Gly90, Asp100, and Leu120-Asn140. All these regions correspond to the binding sites of Ubc9 either with RanBP2 or target proteins, as will be elaborated below. The effect of E3 ligase, RanBP2, binding on these sites are of interest here. The experimental temperature factors of Ubc9 do not exactly overlap the results of simulations, but the envelope is similar with same regions indicated as mobile (Figure 3.11).

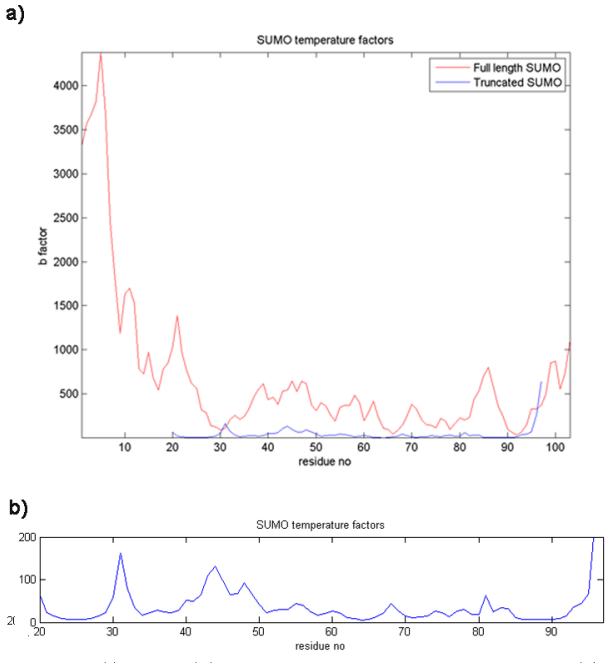

Figure 3.10. Isolated Positional Mean Square Fluctuations of SUMO protein. a) Red line is the MSF calculated from full length SUMO, blue line is the result of calculations using the truncated region only. b) Close view of Positional MSF for truncated SUMO.

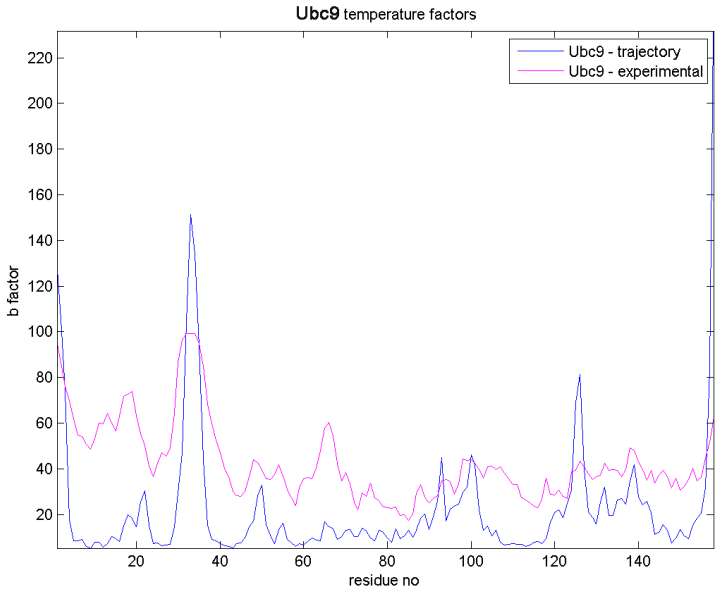

Figure 3.11. Isolated SUMO Positional Mean Square Fluctuations. The experimental positional MSF(magenta) and positional MSF calculated from trajectory.

3.1.2.2. Ubc9-SUMO and Ubc9-SUMO-E3 complexes. While calculating the mean-square fluctuations for the complexes, the change in the quaternary structure, i.e., the orientations of the two structures of the complex with respect to each other, affect the calculated fluctuations due to improper alignment of the structures. Especially in Ubc9-SUMO complex, the small size and high overall mobility of SUMO causes improper alignment of SUMO residues resulting in misleading values for the fluctuations. This is coped by aligning each chain individually (Figure 3.12). The figures presenting the fluctuations of residues with the alignment of each individual chain and the full complex structure for the systems of Ubc9-SUMO and Ubc9-SUMO-E3 can be found in Appendix A. In the comparison between different systems, the fluctuations based on the alignment of individual chains are utilized.

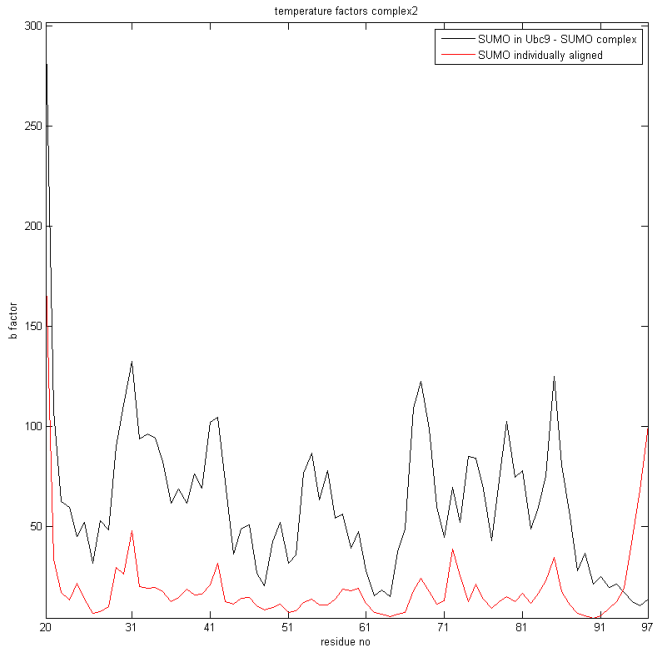

Figure 3.12. Alignment effects on SUMO Positional Mean Square Fluctuations. The positional msf calculated by alignment of the whole complex (black) and calculated by alignment of individual the chain upon itself (red).

Upon binding to SUMO, residues Val27 to Met39 in Ubc9 gain high mobility. These residues mark the first and second β-sheets of Ubc9, and the loop in between. The β-sheets serve as binding sites for E3 enzyme, RanBP2 (Reverter and Lima, 2005). This region of Ubc9 possesses already high fluctuations in the isolated state, but the

mobility is further enhanced in SUMO bound state. Following E3 binding, the fluctuations are reduced back to the values in the unbound state of Ubc9 (Figure 3.13). High mobility of this region is also responsible for the jumps observed in the rmsd profile of Ubc9 in complex with SUMO (Figure 3.13). The change in the mobility of this site, which serves as a binding site for E3 enzyme, RanBP2, should obviously have a functional significance: The enhanced mobility may mark the SUMO bound Ubc9 for E3 interaction. This notion is further supported by the correlations between the fluctuations of residues in Section 3.4.

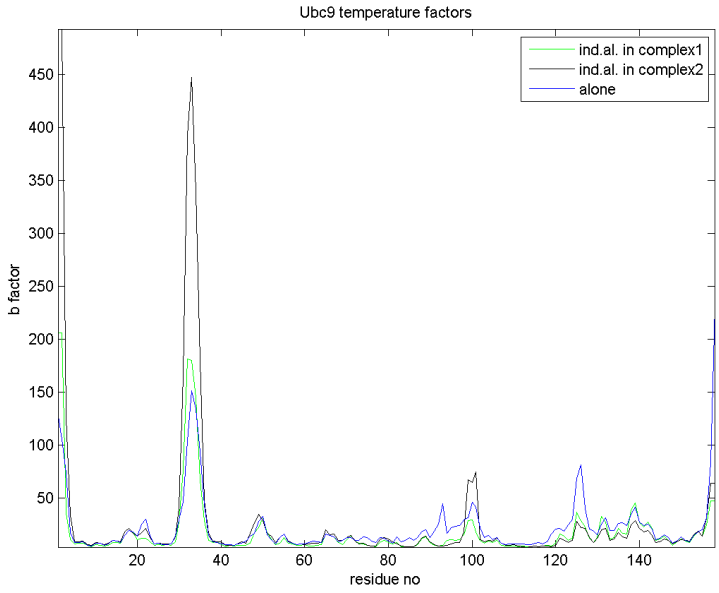

Figure 3.13. Ubc9 positional Mean Square Fluctuations upon Complex Formation are displayed. MSF in Ubc9 - SUMO -E3 complex (green), in Ubc9-SUMO complex (black) and in unbound form (blue) demonstrate the differential effects of each bound chain on Ubc9. Values in complexes are calculated by individual alignment of Ubc9.

In addition to residues from Val27 to Met39, residues between Glu98 and Asp102 of Ubc9 have increased mobility with SUMO complexation, compared to its unbound state. Similar to the β-sheet regions, the fluctuations of this loop are lowered upon binding of the E3 enzyme, RanBP2. This loop has a functional role in the sumoylation mechanism; residues Asp100 and Lys101 take part in target recognition (Tatham *et al.*, 2003; Yunus and Lima, 2006). The E3 enzyme, RanBP2 does not have direct contacts with this loop. On the whole, this result is consistent with the suspected allosteric effect of RanBP2 on Ubc9 target recognition (Pichler *et al.*, 2002; Reverter

and Lima, 2005).

The fluctuations of active cysteine Cys93 of Ubc9 is restricted in both complexes, with and without E3 as compared to unbound state of Ubc9. This is expected as this region is covalently bound to SUMO. Additionally, the residues from Asn124 to Pro128 also have reduced fluctuations compared to unbound Ubc9 in both complexes. The loop harbors residues Asp127 and Pro 128, which are in contact with the consensus sumoylation tetra-peptide motif (Bernier02). This loop is close to the active cysteine and covalently bound SUMO C-terminus in both Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes. Therefore, the restriction in the fluctuations of this loop may be a result of SUMO binding alone, independent of RanBP2 binding. The active site and the loop harboring functional residues in target recognition are less mobile in Ubc9-SUMO complex relative to the unbound Ubc9. The E3 enzyme, RanBP2 binding does not supply further stability to these regions. The ability of SUMO binding to alter the fluctuations of these active sites may be correlated with the ability of Ubc9 to function without the aid of an E3 enzyme.

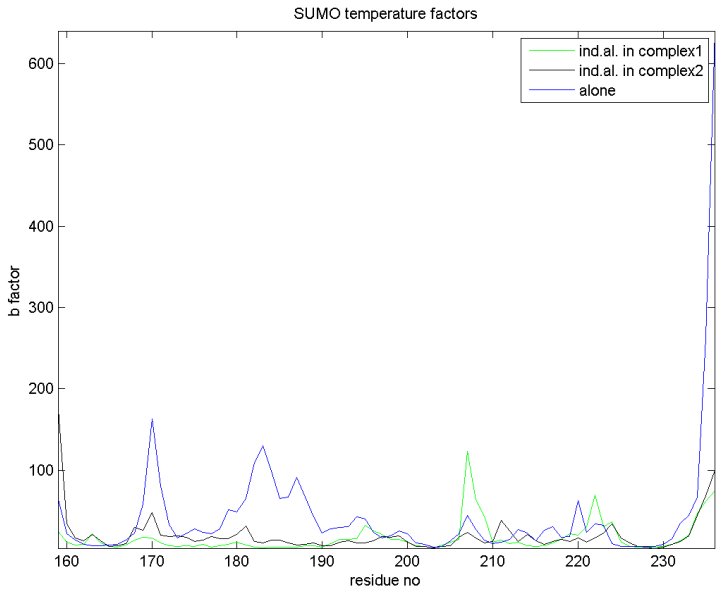

Figure 3.14. SUMO Positional Mean Square Fluctuations upon Complex Formation. MSF in Ubc9-SUMO-E3 complex (green), in Ubc9-SUMO complex (black) and in unbound form (blue). Values in are calculated by individual alignment of SUMO.

The mean-square fluctuations of SUMO are significantly reduced compared to

the unbound states in both complexes. (Figure 3.14) The two peaks, where the fluctuations are higher in Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes relative to the unbound state of SUMO are Gln69 and Glu84. To the present knowledge, these regions do not seem to have a functional role related to the steps of sumoylation mechanism within the scope of this study. Nevertheless, these residues could be taken as plausible functional sites proposed by the present work.

3.1.3. Pseudo - Dihedral Angle Fluctuations

The mean values of pseudo-dihedral angle fluctuations reflect the flexible regions of a protein involved in its dynamics. Here, the pseudo-dihedral fluctuations of Ubc9-SUMO complex, with E3, RanBP2, and without E3 highlight the changes in the flexibility between the two.

The comparison of the pseudo-dihedral fluctuations between Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes do not reveal significant changes in the flexibility of Ubc9. On the other hand, SUMO experiences the decrease in its flexibility (Figure 3.15) at the two main regions with the binding of Ubc9-SUMO to RanBP2. The region starting from residues Glu33-Ile34 and ending at residues Tyr41-Tyr42 of SUMO have reduced flexibility in Ubc9-SUMO-RanBP2 complex, as compared to Ubc9-SUMO complex. These residues construct the β - sheets of SUMO that pack with RanBP2 N -terminal β-sheet. The fluctuations of pseudo-dihedral angles point out limited changes in the flexibility of Ubc9 with RanBP2 binding to Ubc9-SUMO complex. The main changes are observed in SUMO flexibility when RanBP2 binds to Ubc9-SUMO complex. Together with the orientation change in Ubc9-SUMO structure, implied by the rmsd values in the simulations, suggest that RanBP2 mainly affects the dynamics of SUMO. In sumoylation mechanism, RanBP2 protein limits the mobility and reduces the flexibility of SUMO.. The present dynamic analyses of virtual rotational angles of the structures define these conformational restrictions as positioning SUMO to an efficient orientation for conjunction.

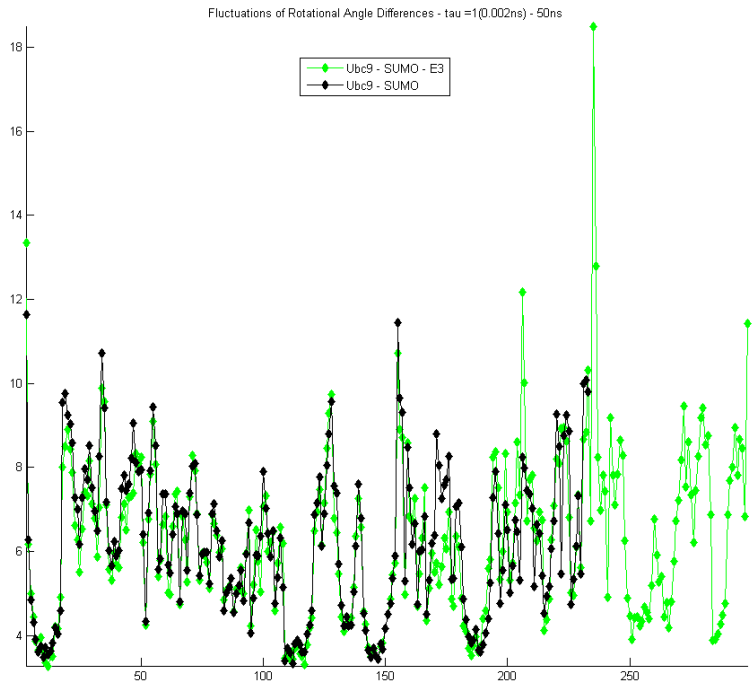

Figure 3.15. Mean Pseudo-Dihedral Angle Fluctuations of Ubc9-SUMO (black) and Ubc9-SUMO-RanBP2 (green) complexes.

3.1.4. Time Delayed Auto-Correlations

The auto-correlations of virtual bond vectors in time are monitored at various time delays. The equilibrium values that these correlations decay and the correlation times reflecting how fast those equilibrium values are reached differ in Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes.

In Ubc9-SUMO-RanBP2 complex, the correlations of virtual bonds are not lost although the time delay is increased up to 30 ns. (Figures 3.16 and 3.17) For almost all virtual bond vectors, the time delayed auto-correlations are above 0.9. When the auto-correlations with time delays up to 30 ns are normalized, it is seen the correlation times are low (Figure 3.18). The virtual bond vectors do not have oreintational mobility and the memory with respect to their orientations for these bond vectors are preserved throughout the trajectory.

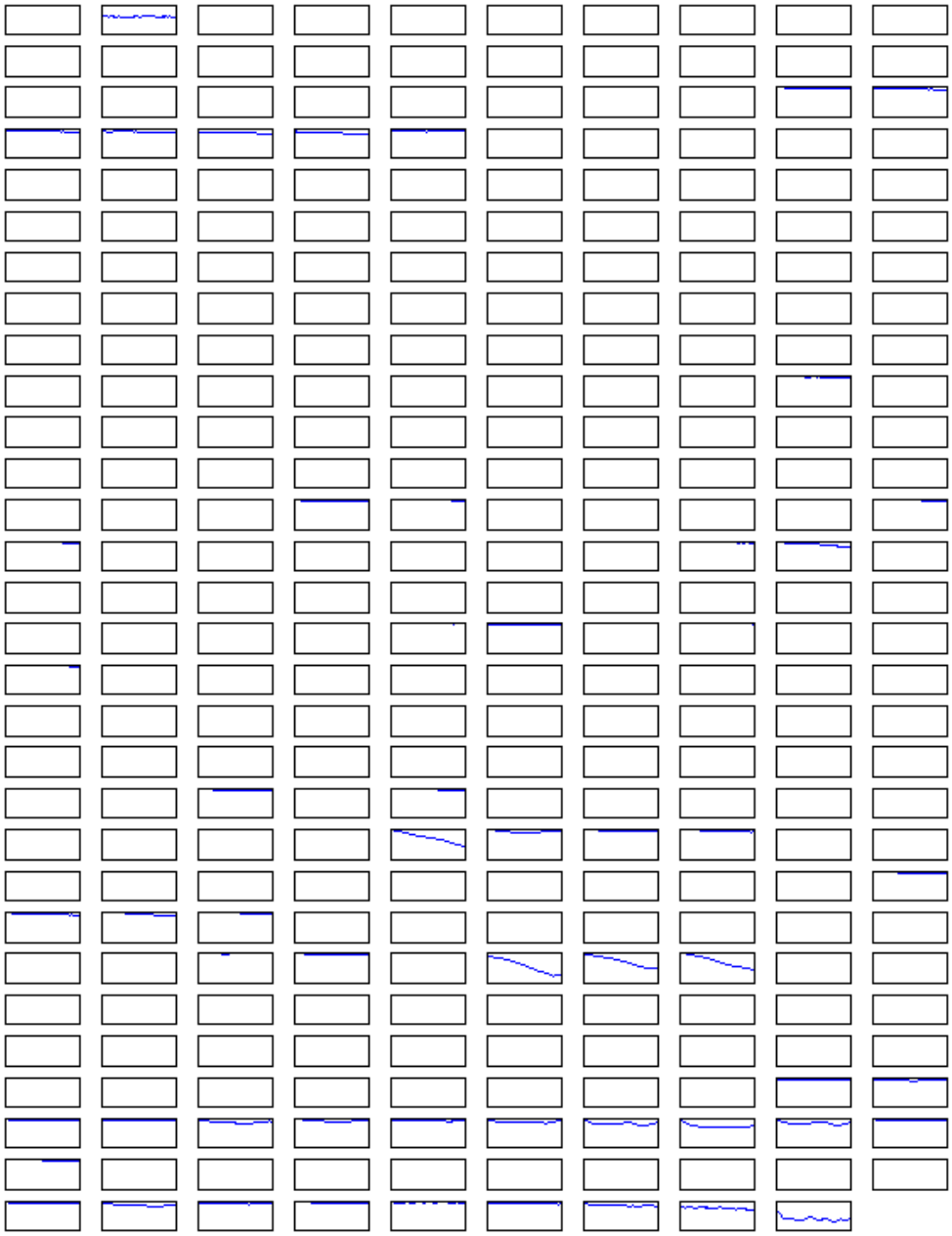

Figure 3.16. Time delayed auto-correlations of Virtual Bond Vectors for Ubc9-SUMO-RanBP2 complex. The correlations with increasing time delay are plotted for each virtual bond vector in the boxes.

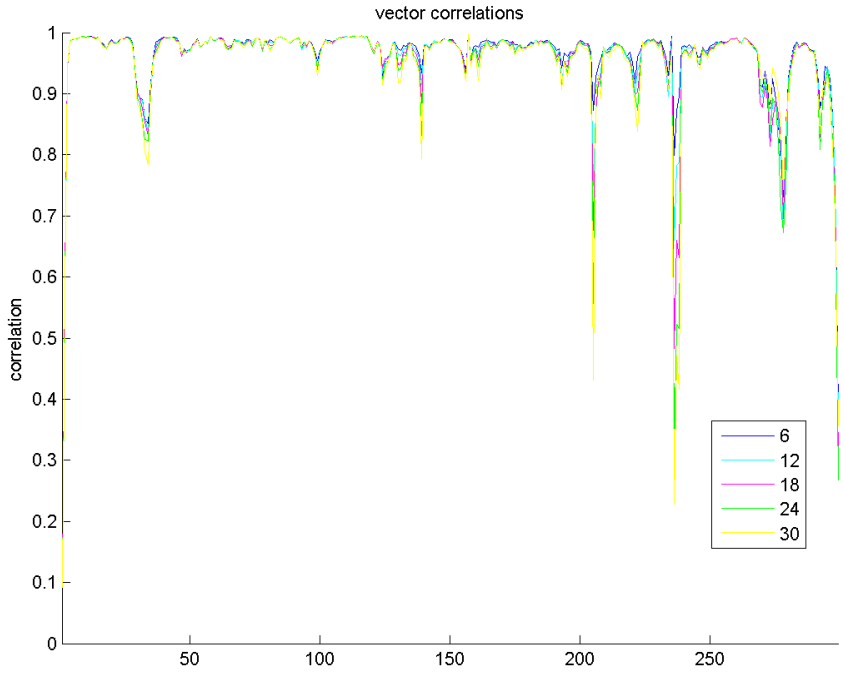

Figure 3.17. Time delayed auto-correlations of Virtual Bond Vectors at selected delays for Ubc9-SUMO-RanBP2. Auto-correlations of virtual bond vectors at 6 ns (blue), 12 ns (cyan), 18 ns (pink), 24 ns (green) and 30 ns (yellow) are represented.

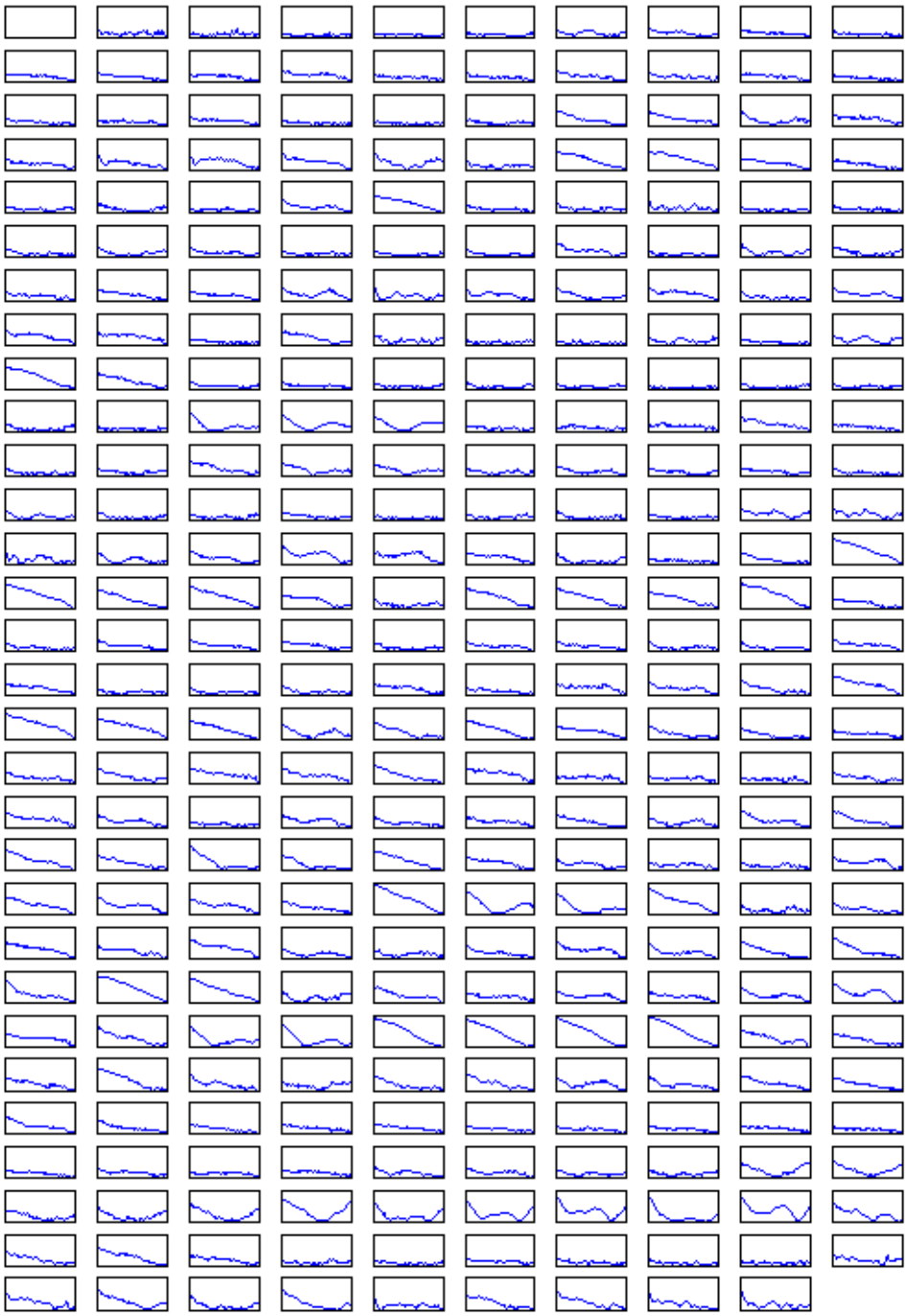

Figure 3.18. Normalized time delayed auto-correlations of Virtual Bond Vectors for Ubc9-SUMO-RanBP2 complex. The correlations with increasing time delay are plotted for each virtual bond vector in the boxes.

In Ubc9-SUMO complex, the bond vectors lose their the orientational correlations as compared to Ubc9-SUMO-RanBP2 complex (Figure 3.19 and 3.20).The virtual bond vectors have higher orientational freedom as compared to Ubc9-SUMO-RanBP2 complex. The changes between the allowed conformations of the virtual bonds take place in longer times, which leads to higher correlation times, i.e., lower relaxation rates (Figure 3.21).

Additional to the general behavior of the virtual bonds in complexes, specific residues proposed to be affected by RanBP2 binding in the previous sections are further investigated here. Asp100 and Lys101 of Ubc9 are functional in recognition of the targets (Tatham *et al.*, 2003; Yunus and Lima, 2006). The mobilities and orientational freedoms of these two residues and near neighboring residues are shown to be reduced upon RanBP2 binding in Sections 3.1.2.2 and 3.2.2.1. The auto-correlations of the virtual bond vectors between Glu99-Asp100 and Asp100-Lys101 in Ubc9-SUMO and in Ubc9-SUMO-RanBP2 complexes openly demonstrates that the orientational space of these bond vectors is quite reduced for the latter (Figure 3.22). This also suggests that RanBP2 binding to Ubc9-SUMO complex aligns these residues to a position favoring the transfer of SUMO from Ubc9 to target proteins. Together with the fact that these residues are far from RanBP2 binding sites on Ubc9, the limitations exerted on Asp100 and Lys101 of Ubc9 can be suggested as an allosteric effect of RanBP2 binding, increasing the efficiency of sumoylation reaction, on Ubc9-SUMO complex.

Other than the residues around Asp100, Figures 3.17 and 3.20 point out the region around Lys30 of Ubc9 with altered flexibility upon RanBP2 binding. This region corresponds to the highly mobile region of Ubc9-SUMO complex, which has reduced mobility upon RanBP2 binding. The virtual bonds around Gly68 of SUMO have low auto-correlation values in both Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes. This residue is at a flexible loop region of SUMO.

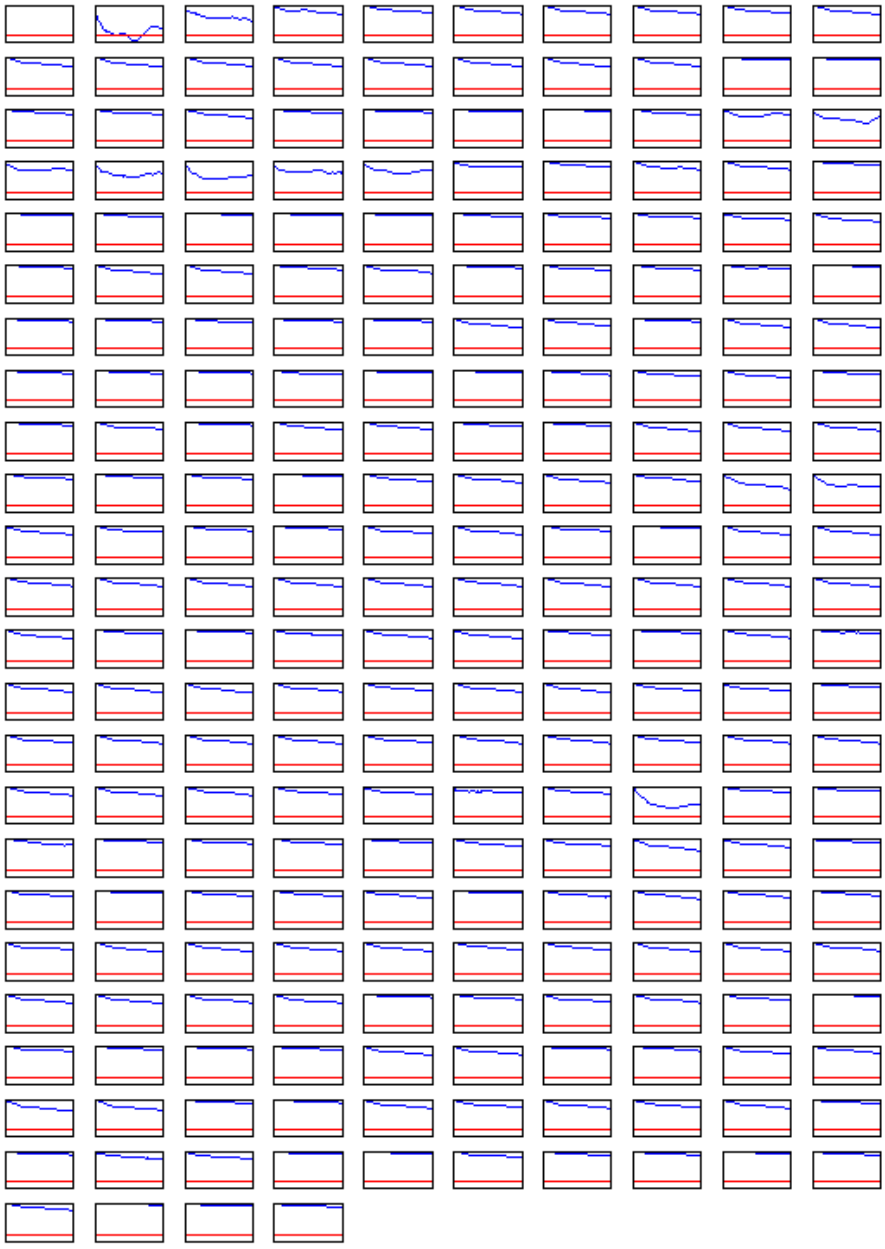

Figure 3.19. Time delayed auto-correlations of Virtual Bond Vectors for Ubc9-SUMO complex. The correlations with increasing time delay are plotted for each virtual bond vector in the boxes. Red lines represent the zero lines.

Figure 3.20. Time delayed auto-correlations of Virtual Bond Vectors at selected delays for Ubc9-SUMO. Auto-correlations of virtual bond vectors at 6 ns (blue), 12 ns (cyan), 18 ns (pink), 24 ns (green) and 30 ns (yellow) are represented.

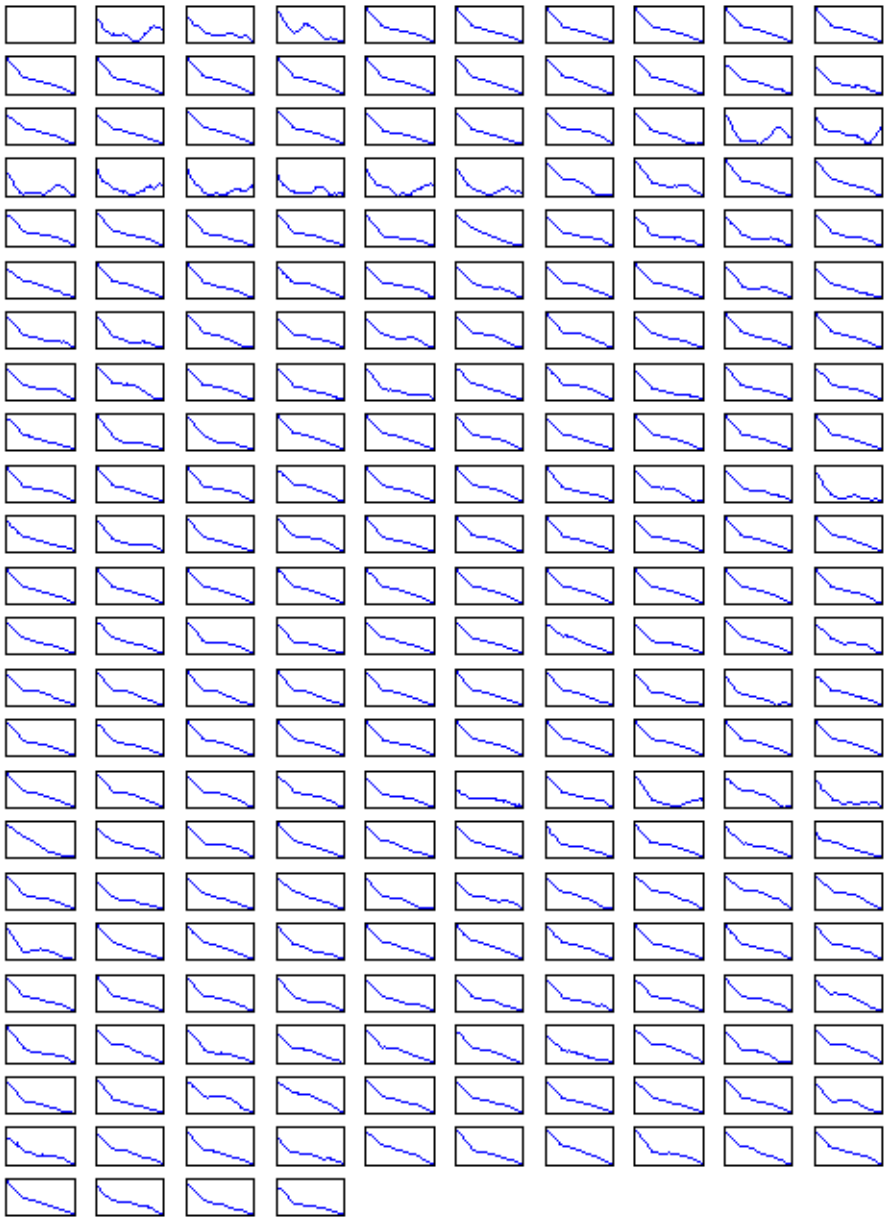

Figure 3.21. Normalized time delayed auto-correlations of Virtual Bond Vectors for Ubc9-SUMO complex. The correlations with increasing time delay are plotted for each virtual bond vector in the boxes.

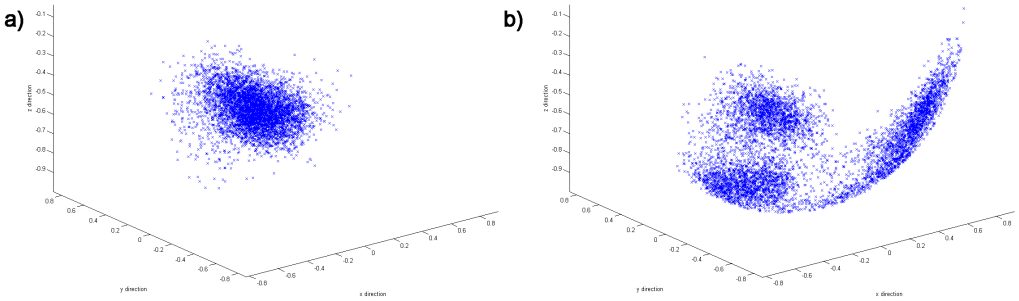

Figure 3.22. Virtual Bond Vector between Asp100 and Lys101. The orientations of the vector in a) Ubc9-SUMO-RanBP2 complex and b) Ubc9-SUMO complex demonstrate the restricted flexibility of the bond in the RanBP2 bound complex.

3.2. Differing Interactions of Functional Residues in RanBP2 bound and unbound Ubc9-SUMO complexes

In the proposed models, E3 enzyme, RanBP2, binding enhances the sumoylation rates by allosteric effects, preventing unproductive orientations of proteins (Pichler *et al.*, 2002; Reverter and Lima, 2005). The observed conformational restriction induced upon Ubc9 and SUMO by RanBP2 binding in the present simulations is in accord with this model. To determine the interactions underlying the effectiveness of favored conformations is the next step of analysis. The contact maps of the conformations throughout the simulations are plotted to monitor the changes in the interacting residues. The interactions revealed with the contact maps, together with known Ubc9-SUMO interaction sites, are followed throughout the trajectories of Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes. The changes in the interaction distances are tried to be functionally characterized.

The discussed functional Ubc9 residues Asp100 and Lys101 are shown to move away from the catalytic site of the complex, which is Cys93 of Ubc9. Among the residue pairs suggested to form hydrogen bonds, three interactions are shown to have differing dynamics in Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes. Additionally, new potential hydrogen bonds are identified; this interactions may help SUMO retain its new orientation with Ubc9.

3.2.1. Contact Maps

The contact maps are calculated for the conformations of Ubc9-SUMO and Ubc9-SUMO-RanBP2 structures, sampled at every 5 ns and for the representative structures selected by the clustering analysis (Representative structure selection is discussed in detail in Section 3.3). Among the regions that show fluctuations in the contact distances, two regions are selected for further analysis. The maps are given in Appendix B.

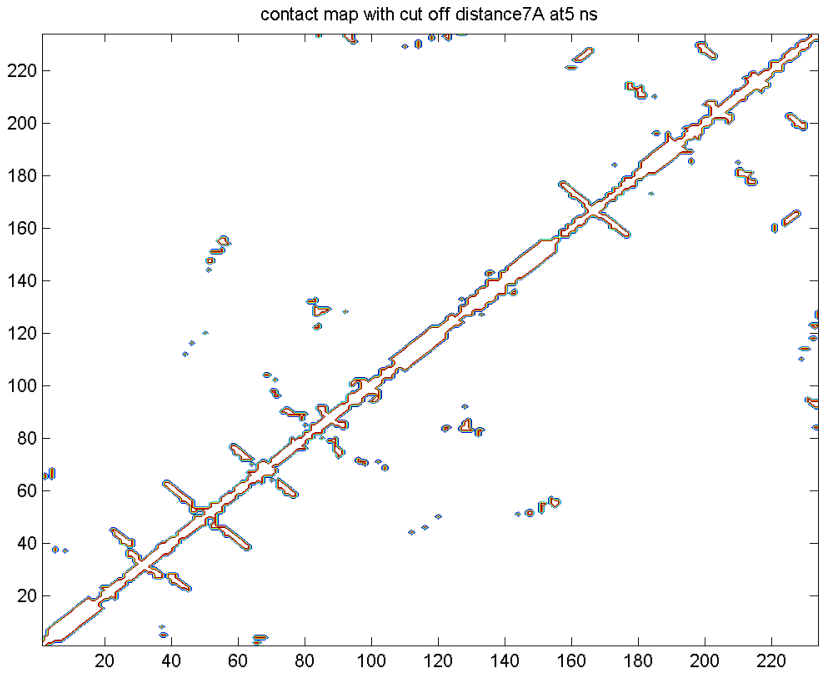

Figure 3.23. Ubc9-SUMO Contact Map at 5 ns, with cut-off distance 7$\mathring{A}$

The first group includes the residues preceding the catalytic sites of Ubc9, and C-terminal tail region of SUMO. The distances among these residues fluctuate throughout the simulation. These sites on Ubc9 and SUMO are close to the catalytic sites on both proteins, thus any interactions arising among these residues might be of importance.

The second group of residues includes the regions between Ser70 - Lys74 of Ubc9 and Ile96-Lys101 of Ubc9. The residues of interest harbor Lys74, functional in binding to consensus sumoylation motif, and Asp100 and Lys101, functional in target recognition (Bernier-Villamor *et al.*, 2002; Tatham *et al.*, 2003; Yunus and Lima, 2006)).

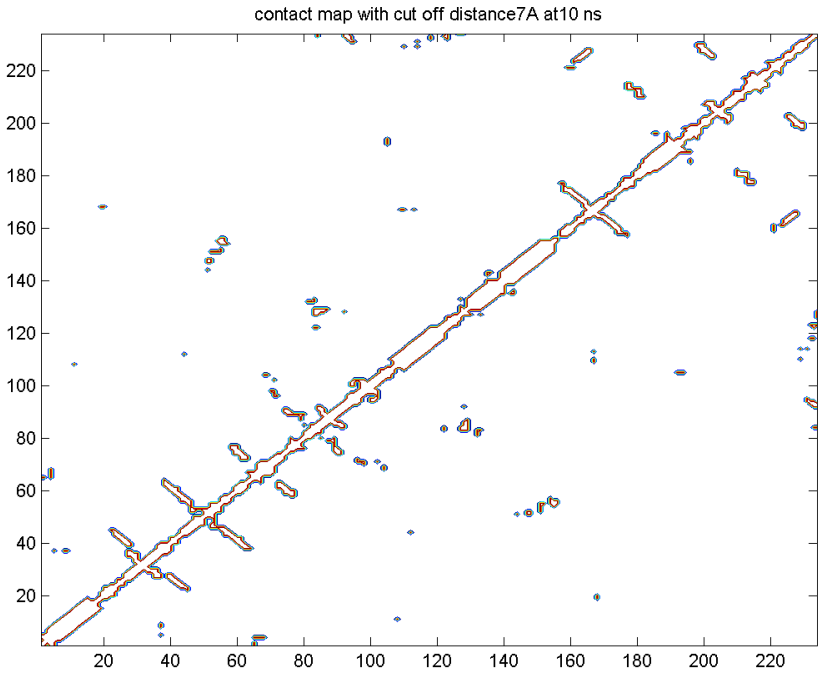

Figure 3.24. Ubc9-SUMO Contact Map at 10 ns, with cut-off distance 7 $\mathring{A}$

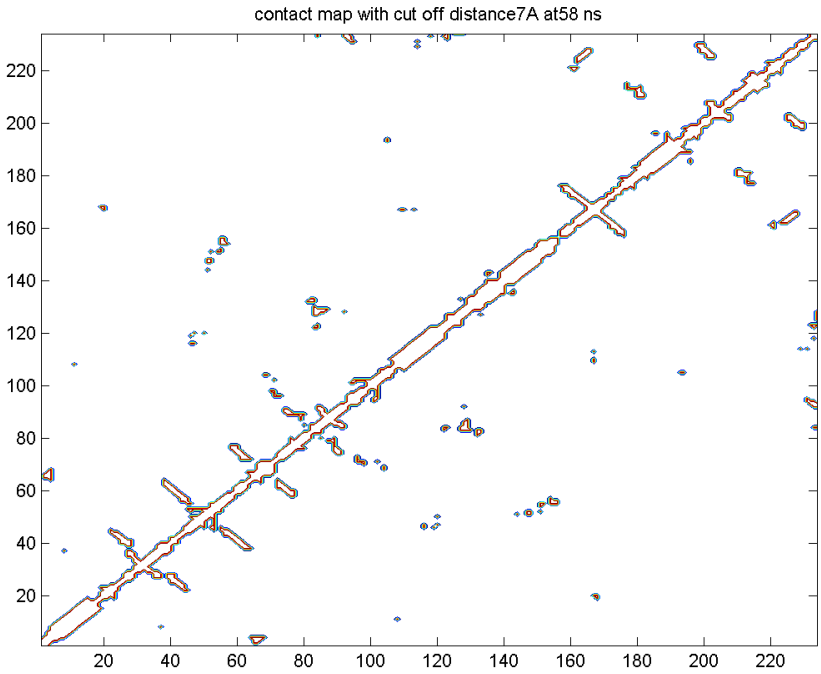

Figure 3.25. Ubc9-SUMO Contact Map at 58 ns, with cut-off distance 7 $\mathring{A}$

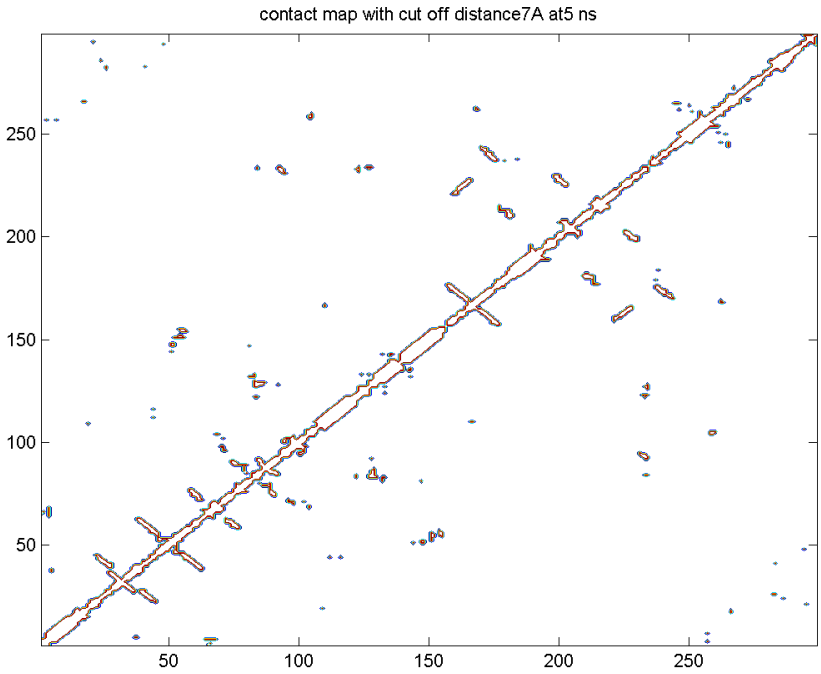

Figure 3.26. Ubc9-SUMO-RanBP2 Contact Map at 5 ns, with cut-off distance 7 $\mathring{A}$

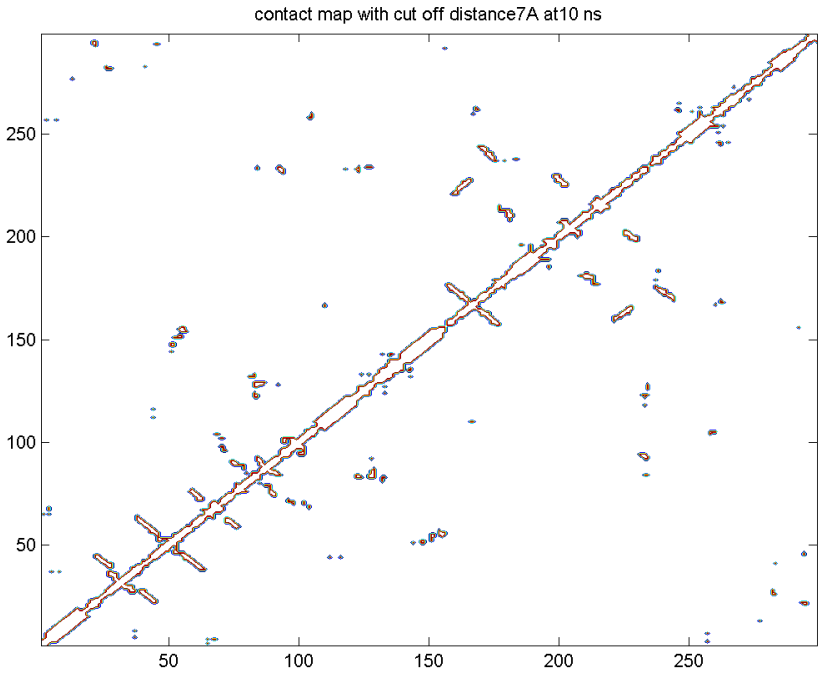

Figure 3.27. Ubc9-SUMO-RanBP2 Contact Map at 10 ns, with cut-off distance 7 $\mathring{A}$

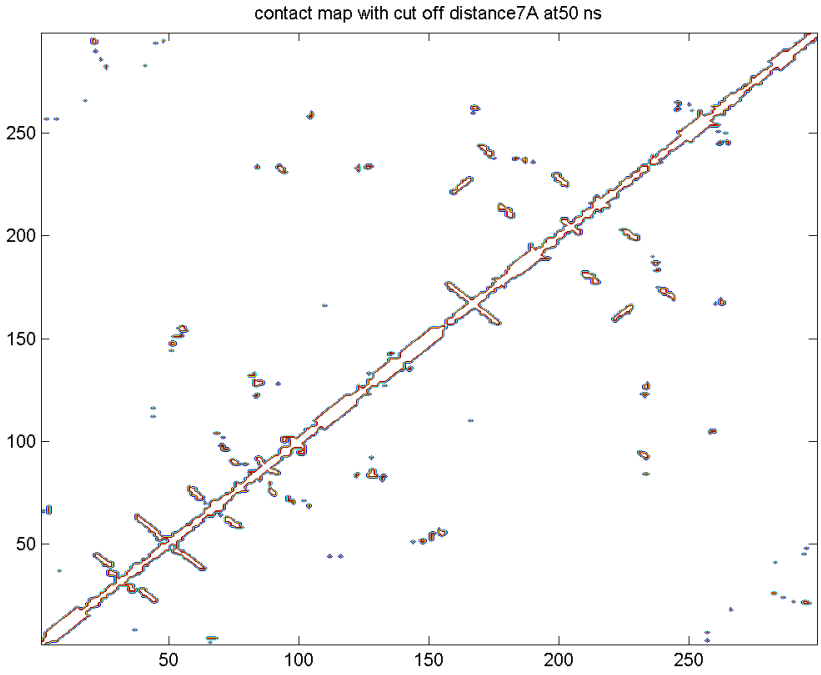

Figure 3.28. Ubc9-SUMO-RanBP2 Contact Map at 50 ns, with cut-off distance 7 $\mathring{A}$

Further investigation of these regions does not reveal potential strong interactions such as hydrogen bonds between these regions, but the change in distance implies aletrations in orientations of related residues. Since residues Asp100 and Lys101 are thought to be responsible in interactions with the target (Tatham *et al.*, 2003; Yunus and Lima, 2006), their orientation with respect to the reaction site should be of importance. Taking together the observations of the increase in mean square fluctuations of residues Asp100 and Lys101 with SUMO binding and the decrease of the mean-square fluctuations with RanBP2 binding, it is decided to monitor the distances between Asp100 and Lys101 and active cysteine.

Additional to these two groups of residues, during the simulation of Ubc9-SUMO complex, residues Lys18-Phe22 of Ubc9 and residues Gly28-Ser32 of SUMO come into contact distance. This decrease in the distance between the two regions is initially observed at the time interval corresponding to the orientation change of SUMO, and is not lost throughout the 58 ns simulation. The loop on Ubc9 includes E3 enzyme, RanBP2, binding site and the loop on SUMO corresponds to the loop between the β-sheets, which pack with RanBP2 β-sheet. In the existence of RanBP2, instead of

interacting with each other, these sites interact with the binding sites on E3. . The interactions arising in Ubc9-SUMO complex, between Ubc9 and SUMO residues that are significant in RanBP2 binding for both proteins, imply functional importance. On the other hand, no potential interactions strong enough to have a structural role are identified in this region.

3.2.2. Functional and Structural Interactions between Ubc9 and SUMO residues

3.2.2.1. Orientation of Asp100 and Lys101 of Ubc9 with respect to the catalytic site.

Asp100 and Lys101 of Ubc9 display increased mobility when Ubc9 forms a complex structure with SUMO. Nevertheless, the binding of RanBP2 reduces this mobility, as can be seen from Figure 3.13. The distances of Asp100 and Lys101 to the active cysteine Cys93 are given in Figures 3.29 and 3.30, respectively.

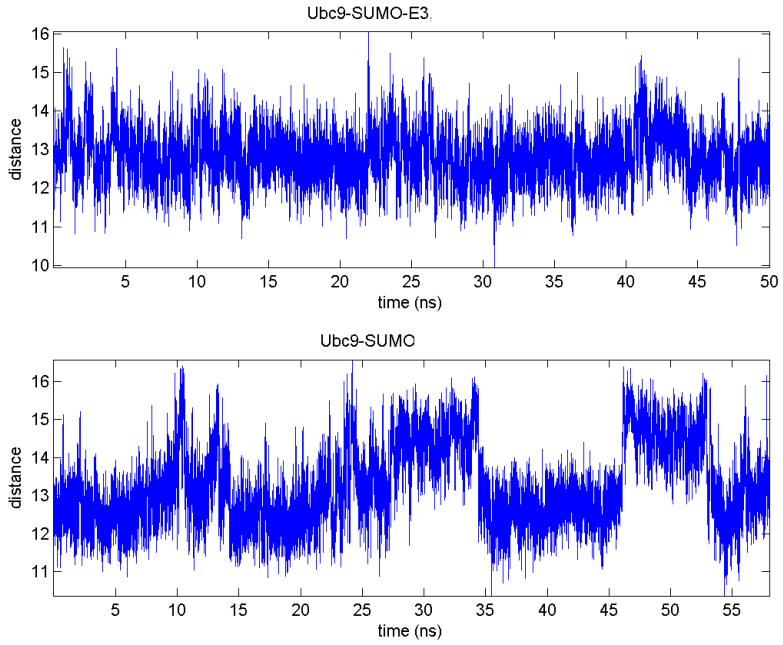

Figure 3.29. Distance between center of masses of residues Cys93 and Asp100 of Ubc9 throughout the trajectories. Upper lane is the distance for Ubc9-SUMO-E3 complex, and lower lane is the distance for Ubc9-SUMO complex

The fluctuations of the distances clearly show that both residues change their positions with respect to the thioester bond in Ubc9-SUMO complex and RanBP2

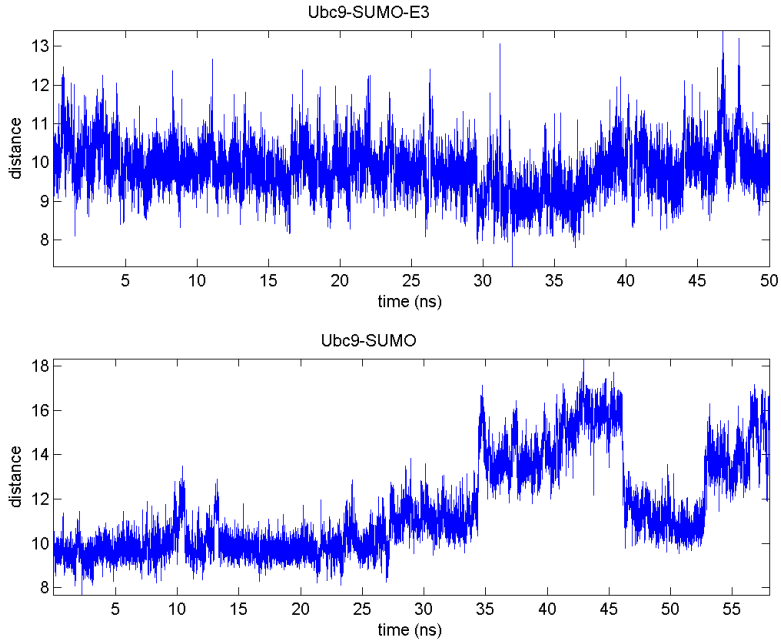

Figure 3.30. Distance between center of masses of residues Cys93 and Lys101 of Ubc9 throughout the trajectories. Upper lane is the distance for Ubc9-SUMO-E3 complex, and lower lane is the distance for Ubc9-SUMO complex

fixes the distance between these residues. While Asp100 moves away from the SUMO bound active cysteine in short time intervals, the increase in the distance is more evident between Lys101 and C93. In short, the residues Asp100 and Lys101, functional in recognizing the approaching targets, are fixed upon RanBP2 binding. This restriction in the fluctuations with respect to the catalytic sites should be one of the factors increasing the reaction efficiency. This is an indication of an allosteric effect of RanBP2 on Ubc9, as these residues are at the opposite surface of Ubc9 with respect to RanBP2 binding sites.

3.2.2.2. Lost interactions of structural significance in Ubc9-SUMO simulation. Among the interactions between Ubc9 and SUMO, the interactions between Arg63 of SUMO - Glu122 of Ubc9, Gln29 of SUMO - Gln111 of Ubc9 and Gln92 of SUMO - Arg104 of Ubc9 (Figure 1.3) have differences in their stabilities in Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes.

Residues Arg63 of SUMO and Glu122 of Ubc9 are in hydrogen bonding distance in the crystal structure. The hydrogen bonding distance is preserved in the simulations of Ubc9-SUMO-RanBP2, but this interaction is unstable and lost in Ubc9-SUMO complex after approximately 42 ns (Figure 3.31). One of the interactions of structural significance between SUMO and Ubc9 is lost when Ubc9-SUMO complex functions alone without E3 enzyme, RanBP2.

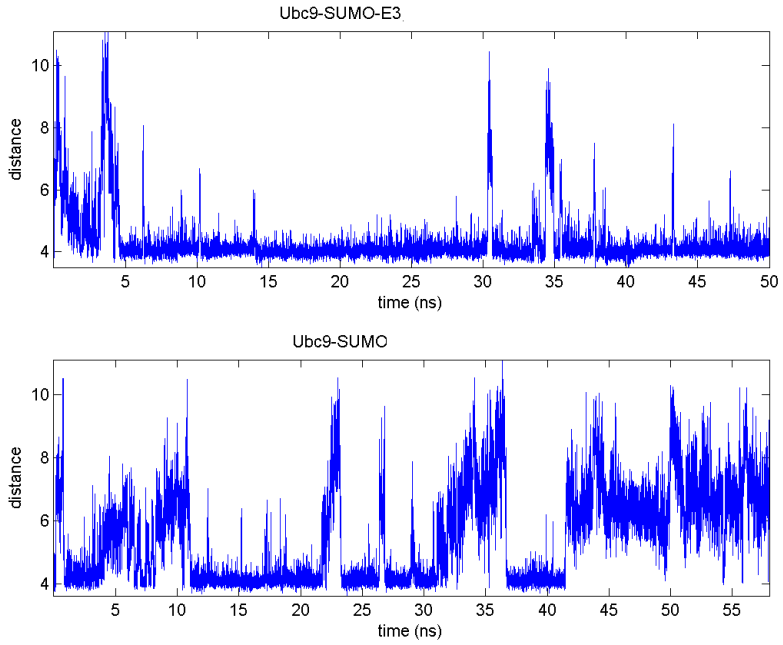

Figure 3.31. Distance between carbon atoms of residues Arg 63 of SUMO and Glu122 of Ubc9 throughout the trajectories. Upper lane is the distance for Ubc9-SUMO-E3 complex, and lower lane is the distance for Ubc9-SUMO complex

Residues Gln29 of SUMO and Gln111 of Ubc9 are also in hydrogen bonding distance in the crystal structure. The interaction can be seen in the Ubc9-SUMO-RanBP2 simulations (Figure 3.32).Although the distance is slightly longer and the angle is slightly distorted for a classical hydrogen bond, there is an interaction between these residues. On the other hand, the distance between these two residues increases to approximately 9 $\mathring{A}$ at the initial steps of the simulation of Ubc9-SUMO and does not decrease below 8 $\mathring{A}$ thereafter. This interaction with a potential structural role is lost and never regained in during the 58 ns simulations of Ubc9-SUMO complex structure.

Residues Gln92 of SUMO - Arg104 of Ubc9 are proposed to take part in formation

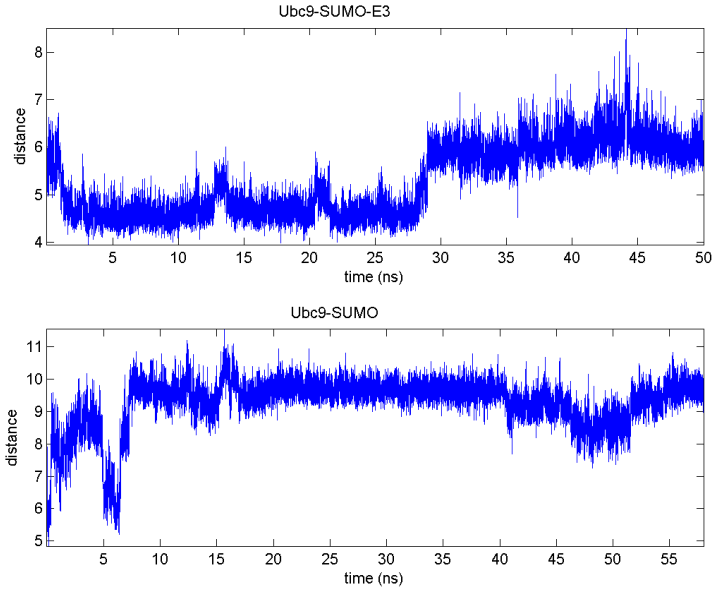

Figure 3.32. Distance between α carbon atoms of residues Gln29 of SUMO and Gln111 of Ubc9 throughout the trajectories. Upper lane is the distance for Ubc9-SUMO-RanBP2 complex, and lower lane is the distance for Ubc9-SUMO complex

of a hydrogen bond in the crystal structure of Ubc9-SUMO-RanGAP1 (target)-RanBP2 (E3 enzyme). While the crystal structure suggests such an interaction, the dynamics reveals that these residues spend a small fraction of simulation time in an appropriate orientation and distance for interaction (Figure 3.33). The probability of obtaining a conformation favoring the interaction between these two sites is higher in Ubc9-SUMO-RanBP2 complex as compared to Ubc9-SUMO complex, but the distance is too high to suggest any preserved interaction. To this end, these contact is not indicative of a contact that could be encountered in the average structure.

3.2.2.3. Gained interactions in Ubc9-SUMO simulation. Investigation of the residues around the catalytic regions on both Ubc9 and SUMO proteins as proposed by the contact maps reveals two new-formed interactions in Ubc9-SUMO complex, in the region preceding the catalytic sites.

Residues Glu122 of Ubc9 and Gln94 of SUMO have a distance of approximately

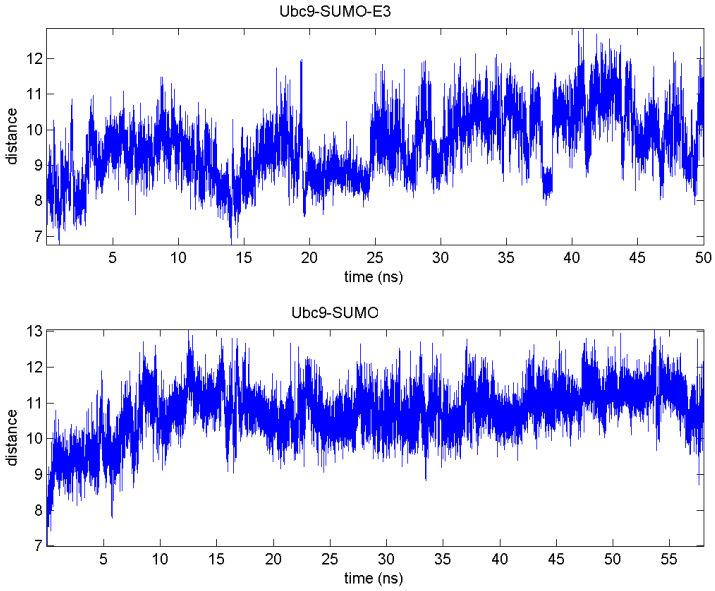

Figure 3.33. Distance between atoms C and N, holding the potential partners of a hydrogen bond, of residues Gln92 of SUMO and Arg104 of Ubc9. Upper lane is for Ubc9-SUMO-RanBP2 complex, and lower lane is for Ubc9-SUMO complex

7 $\mathring{A}$ between potentially interacting atoms in Ubc9-SUMO-RanBP2 complex structure (Figure 3.34). The same distance for Ubc9-SUMO complex structure is below 4 $\mathring{A}$ for the majority of the simulation time. The residue Glu122 is supposed to have a hydrogen bond with Arg63 of SUMO. This bond does not exist for Ubc9-SUMO, but instead Glu122 makes another contact with residue Gln94 within 4 $\mathring{A}$.

Glu118 of Ubc9 and Gln92 of SUMO are at 5.5 $\mathring{A}$ distance in Ubc9-SUMO-RanBP2 complex structure throughout the simulations. Ubc9-SUMO keeps this distance at 4.5 $\mathring{A}$ (Figure 3.35). This strong interaction is formed after SUMO changes its position in the Ubc9-SUMO complex. The change in position of SUMO is persistent after 10 ns. The simulations suggest that the E3 enzyme, RanBP2, binding eliminates this conformational change of SUMO. The timing, orientation and strength of the interaction between Glu118 of Ubc9 and Gln92 of SUMO suggest that this interaction may help SUMO stay in this altered position reached around 10 ns. The fact that the interaction is not seen in the RanBP2 bound complex also supports this model, where the interaction Glu118-Gln92 stabilizes the SUMO conformation that is eliminated by RanBP2 binding.

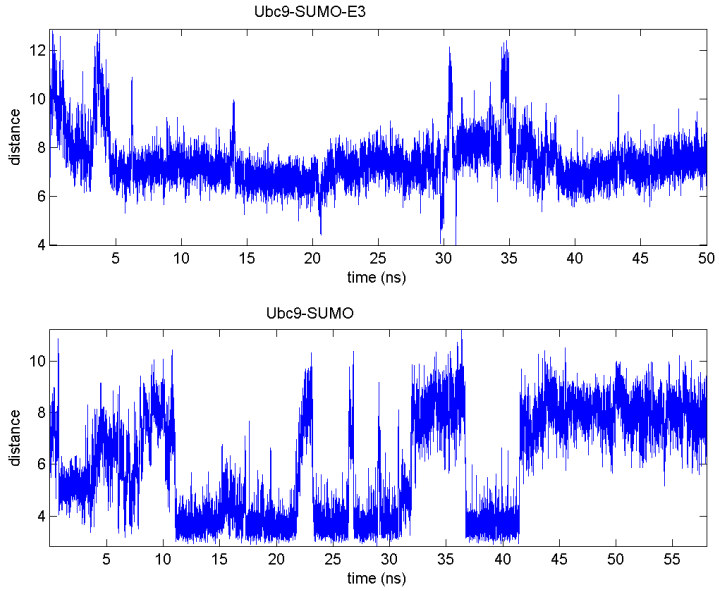

Figure 3.34. Distance between atoms C and N, holding the potential partners of a hydrogen bond, of residues Gln94 of SUMO and Glu122 of Ubc9. Upper lane is for Ubc9-SUMO-RanBP2 complex, and lower lane is for Ubc9-SUMO complex

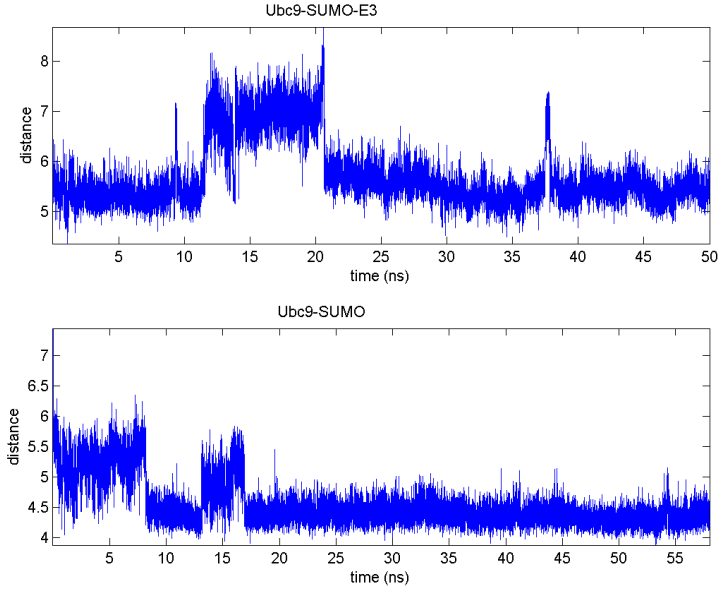

Figure 3.35. Distance between alpha C of residues Gln92 of SUMO and Glu118 of Ubc9. Upper lane is for Ubc9-SUMO-RanBP2 complex, and lower lane is for Ubc9-SUMO complex

Ubc9-SUMO does not have exactly the same interaction network as its crystal structure. New formed interactions may stabilize the complex in a different conformation than the conformation seen in the crystal structure (Glu118 - Gln92). Nevertheless, some of the lost interactions are replaced with interactions of neighboring sites (Glu122-Gln94). New interactions may also stabilize the orientation of C terminal SUMO tail, affecting the efficiency of the reaction without an E3 enzyme.

3.3. Representative Structures and Time Distribution of Conformations by Cluster Analysis

Cluster analysis groups the conformations of structures according to the similarity between the conformations, measured by rmds. The number of clusters and the distribution of cluster best members at different rmsd values provide information about the extent of mobility, as well as the unique conformers, of the complex structures studied. This reduces the conformational space sampled throughout the trajectory. The distribution of cluster members in time provides information on sequential conformation changes.. The simulation time each cluster spans obviously correlates with the probability of the ensemble of conformations it represents.

The cluster analysis revealed the rigidity of isolated Ubc9, relative to the complexes. The conformations generated from the simulation of Ubc9-SUMO complex formed clusters distributed to a well defined time intervals throughout, The mobility of the region around residue Lys30 of Ubc9 and the orientation change of SUMO are highlighted outcomes. Conformations of Ubc9-SUMO-RanBP2 complex leads to a smaller number of clusters with similar percentages of occupied simulation time.

3.3.1. Clusters of Isolated Ubc9 structure

The trajectory of the unbound Ubc9 cannot be divided into more than one cluster until the rmsd is reduced to 1.5 $\mathring{A}$. The rmsd of each snapshot from the centroid of related cluster is plotted against time in Figure 3.36 (a), members of Cluster 1, Cluster 2 and Cluster 3 are color coded, respectively, in black, red and blue. The rmsd values

between the respective centroids are around 1 $\mathring{A}$. Ubc9 is a rather rigid and stable protein throughout the simulation, excluding the loop regions (Figure 3.37).

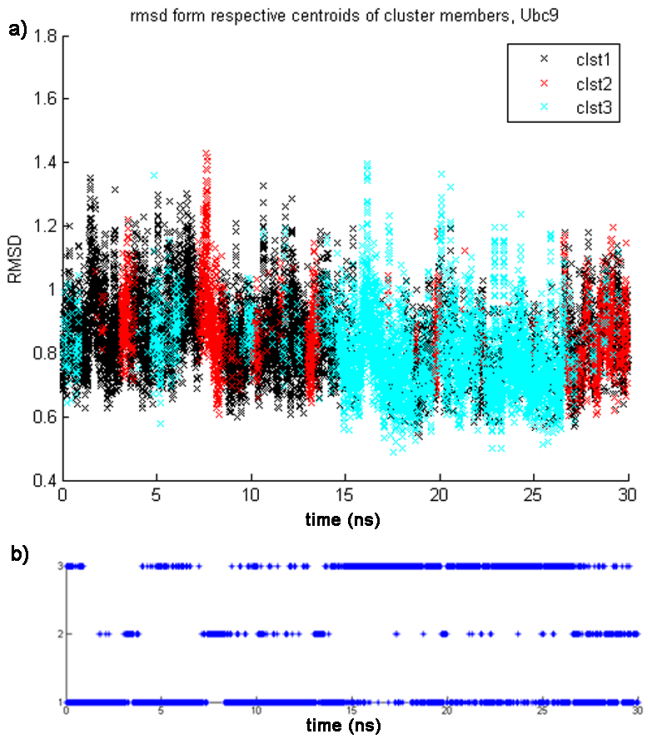

Figure 3.36. Clusters of Ubc9 at 1.5 $\mathring{A}$ cut-off. a) Members of Cluster 1, Cluster 2 and Cluster 3 are given in black, red and blue respectively. b) Time distribution of members of each cluster.

3.3.2. Clusters of Ubc9 - SUMO complex

Clustering the trajectory of Ubc9-SUMO complex with 2 $\mathring{A}$ rmsd results in formation of four clusters (Figure 3.38). The time distribution of cluster members clearly suggests the orientation change of SUMO during the simulation, starting at 6 ns. After the orientational change, lesser number of clusters are formed. Additionally, the time period between 25 to 30 ns has high rmsd. Thus, for more detailed sampling, the analysis is continued with more representative structures obtained by clustering with

Figure 3.37. Best member for Clusters of Ubc9 at 1.5 $\AA$ cut-off. Best member of Cluster 1, Cluster 2 and Cluster 3 are given in black, cyan, blue and magenta, respectively.

lower rmds values.

When the rmsd value is lowered to 1.7 $\AA$, the number of clusters increases to six (Figure 3.39). As expected, the interval from 24 to 31 ns is defined as an additional cluster at this rmsd value. Decreasing the rmds value further leads to several additional clusters, which is also not desired. Thus, the analysis is preferred to be carried out at rmsd value of 1.7 $\AA$.

The interesting feature of the Ubc9-SUMO clusters is that the defined clusters also separate the trajectory into distinct time intervals. After SUMO motion starting at 6 ns, the clusters have members belonging to almost non-overlapping time intervals. The cluster occupying approximately final 10 ns has members in the transition region of SUMO orientation, and Cluster 4 also has members at the end of the simulation. This indicates that the structure is under the dynamic equilibrium.

The simulations suggest that the Ubc9-SUMO complex structure has conformational changes occupying long time intervals. The structure goes through ordered

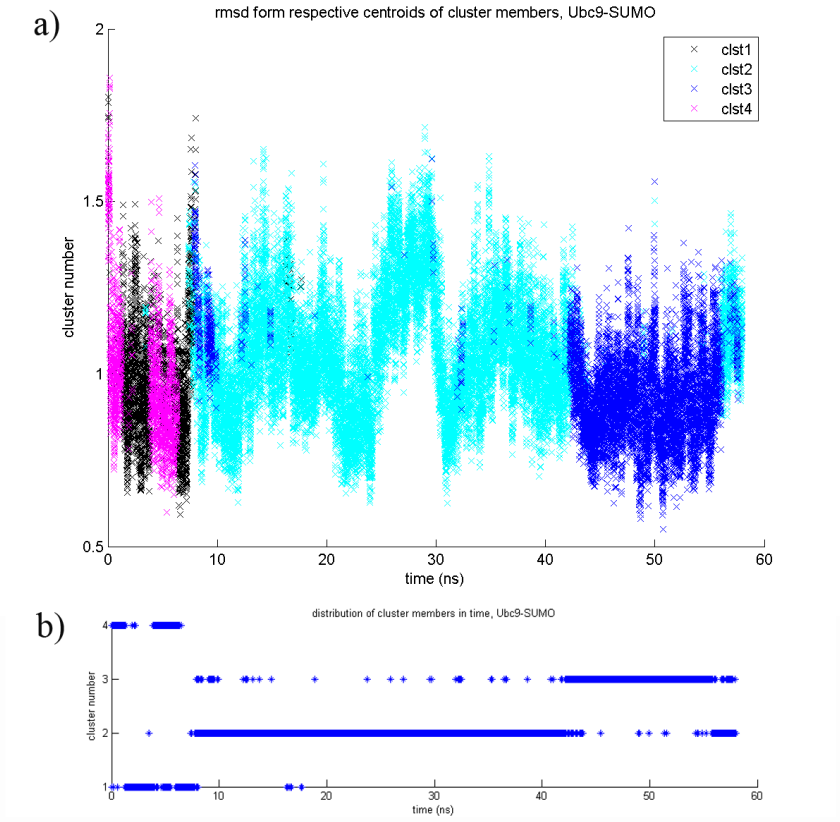

Figure 3.38. Clusters of Ubc9-SUMO complex at 2 $\mathring{A}$ cut-off. a) Members of clusters are represented as black, cyan, blue, magenta, in numerical order b) Time distribution of members of each cluster.

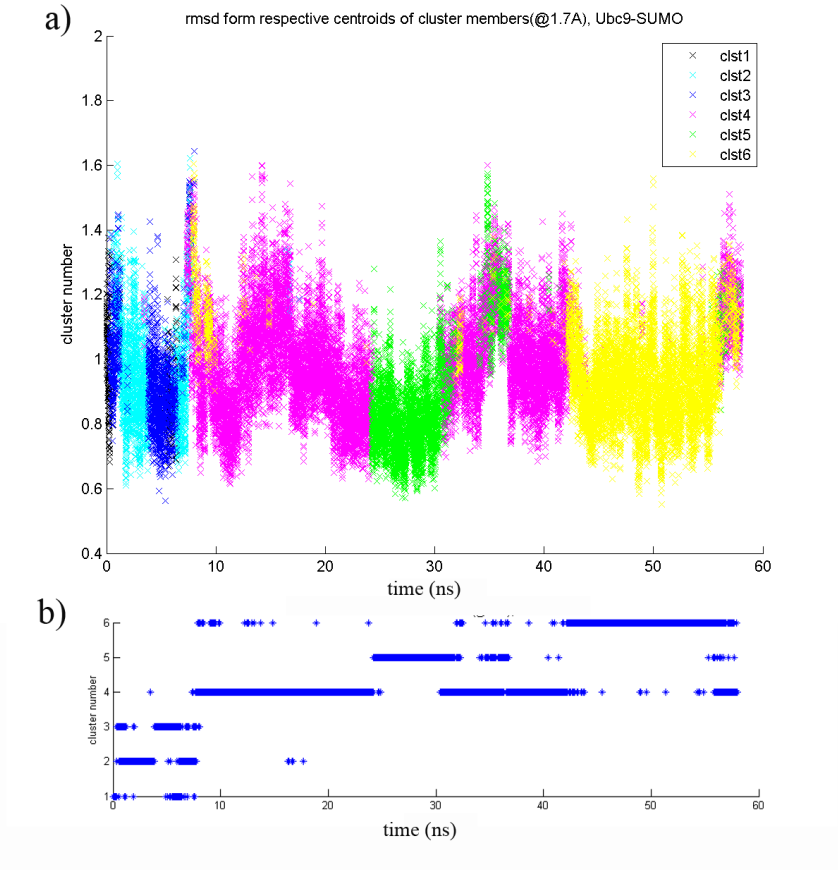

Figure 3.39. Clusters of Ubc9-SUMO complex at 1.7 $\mathring{A}$ cut-off. a) Members of clusters are represented as black, cyan, blue, magenta, green and yellow, in numerical order b) Time distribution of members of each cluster.

motions, instead of spanning all possible conformations at all times. Clusters 4 and 6 display that, SUMO does not go back to its position in the crystal structure once it changes its orientation. Cluster 5 signifies the mobility of the region around residue Lys30 of Ubc9 (Figure 3.40). This region has binding sites with RanBP2 and its mobility can also be observed from positional MSF, as discussed in Section 3.1.2. Cluster analysis displays that the mobility of this region specifically increases after SUMO changes its initial orientation, which would be prevented by RanBP2 binding. This also suggests that the mobility may mark SUMO bound Ubc9 for RanBP2 binding.

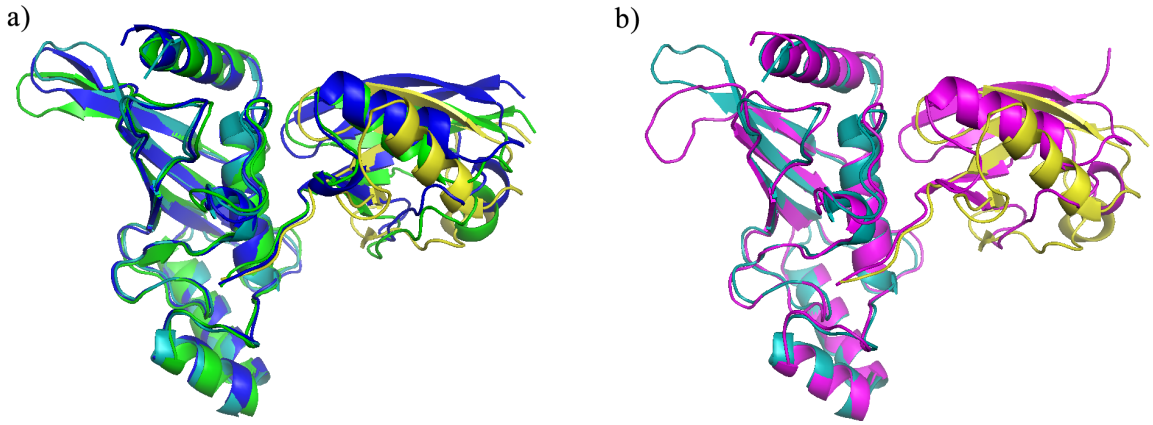

Figure 3.40. Best member for Clusters of Ubc9-SUMO at 1.7 $\AA$ cut-off. The structures are aligned via Ubc9 chains, Ubc9 is in light blue and SUMO is in yellow in the crystal structure. a) Best member of Cluster 4 is in blue and best member of Cluster 6 is in green. The orientation change of SUMO is evident in both structures. b) Best member of Cluster 5 is represented in magenta. The mobility of the region from Val27 to Met39 of Ubc9 is represented.

The clusters with higher percentages of the simulation time represent more probable conformations (Table 3.1). In the majority of the simulation time, (86.68 per cent) Ubc9-SUMO complex attains a conformation belonging to Cluster 4, 5 or 6. All the three clusters represent SUMO protein in its altered conformation with respect to the crystal structure. This may indicate SUMO would most probably have this orientation in the Ubc9-SUMO complex, *in vivo*.

Table 3.1. Time Spent at each Cluster for Ubc9-SUMO Simulations

Cluster	Percentage of Time Spent
1	1.42
2	7.06
3	4.82
4	48.34
5	13.24
6	25.12

3.3.3. Clusters of Ubc9-SUMO-RanBP2 complex

Clustering of the conformations of Ubc9-SUMO-RanBP2 with rmsd of $\mathring{A}$ yields four clusters. The members of the clusters are scattered into two time spans (Figure 3.41).

Members of Cluster 1 and 2 are distributed in the first 28 ns of the trajectory, whereas the members of Cluster 3 and 4 are distributed from 28 to 50 ns. When the best members of each cluster are superimposed, the difference in the packing between the β-sheet of RanBP2 and β-sheet of SUMO can be seen (Figure 3.42). This change in the conformation is the reason for the separation of the trajectory mainly into the two time windows. Reducing the rmsd values for the clustering yields high number of clusters. The clusters define the conformations spanned during the transition between two conformations, where N-terminus of RanBP2 fragment is fully packed with SUMO β-sheets, and the terminal residues of RanBP2 fragment are unpack, forming a mobile loop (Figures C.1 and C.2).

The time the complex spends as a member of each cluster is given in Table 3.2. The percentages of time covered by each cluster are close during the simulation. Especially if the values are considered in the context of two time windows identified by the β-sheets, the distribution is almost the same. From 0 to 28 ns, Clusters 1 and 2 are equally distributed and from 28 to 50 ns, Clusters 3 and 4 are equally distributed.

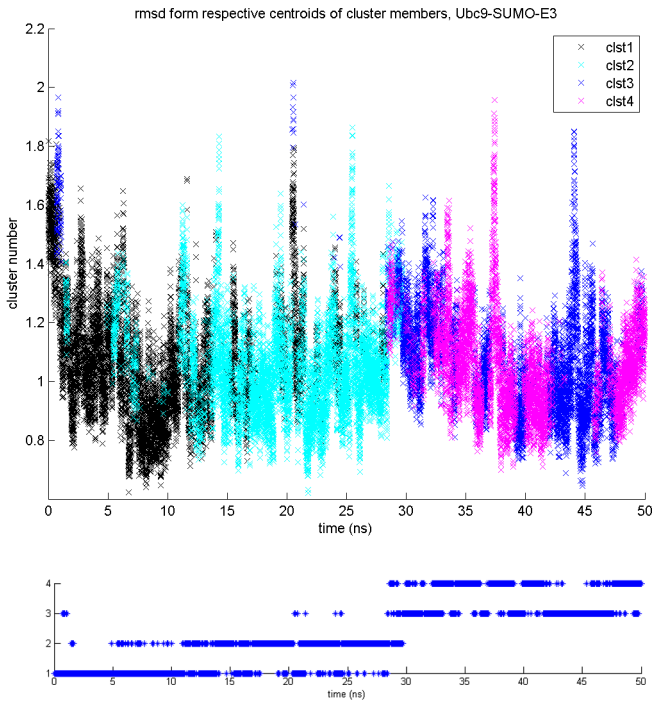

Figure 3.41. Clusters of Ubc9-SUMO-RanBP2 at 2 $\mathring{A}$ cut-off. a) Members of Cluster 1, Cluster 2, Cluster 3 and Cluster 4 are given in black, cyan, blue and magenta respectively. b) Time distribution of members of each cluster.

Figure 3.42. Best member for Clusters of Ubc9-SUMO-RanBP2 at 2 $\mathring{A}$ cut-off. Best member of Cluster 1, Cluster 2, Cluster 3 and Cluster 4 are given in yellow, cyan, blue and magenta, respectively.

Table 3.2. Time Spent at each Cluster for Ubc9 - SUMO - E3 Simulations

Cluster	Percentage of Time Spent
1	29.6
2	28.1
3	22.5
4	19.8

The best members of Cluster 3 and Cluster 4 also show some differences in packing of the β-sheets, the best member of Cluster 4 loosing more of the β-sheet conformation at the N-termini of RanBP2 (Figure 3.42). On the other hand, the equal distribution of the members of Cluster 3 and Cluster 4 in the time interval after 28 ns suggests that the behavior with respect to the packing of β-sheets is under dynamic equilibrium. The representative structures of the clusters do not define a time dependent directed motion of the Ubc9-SUMO-E3 complex structure, such as the orientation change observed in SUMO in Ubc9-SUMO complex. Rather, the structure of Ubc9-SUMO-RanBP2 complex is at equilibrium after unpacking of the initial residues of N-terminal β-sheet of the RanBP2 fragment. The change in the packing of the β-sheets do not lead to significant changes in the interactions or correlations of residues for Ubc9-SUMO-RanBP2 complex (Sections 3.2, 3.4). The change does not seem to be functionally significant in the context of the steps of sumoylation mechanism that are investigated in this study.

3.3.4. Quaternary Structure Change and Ubc9 Mobile Fragments from Clusters of Conformations

The orientation change of SUMO is investigated further by analyzing the representative structures from the clusters of conformations obtained from the simulations of each complex. As implied by the time distribution of clusters, the representative structures for Clusters 4, 5 and 6 of Ubc9-SUMO complex clearly demonstrates the motion of SUMO. Lines represent the deviation of the positions of the best members of Cluster 4 (orange), Cluster 5 (blue) and Cluster 6 (green) from the crystal structure (Figure 3.43 (a)). The deviations are based on α-Carbons. The crystal structure is

given in Figure 3.43 (b).

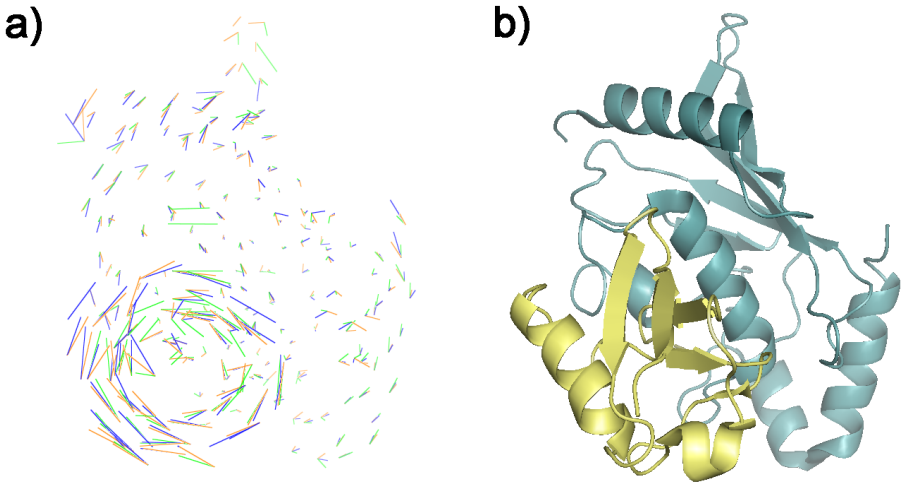

Figure 3.43. SUMO motion in Ubc9-SUMO complex. a) Deviations of best members of Cluster 4 (orange), 5 (blue) and 6 (green) from the crystal structure. b) The crystal structure of Ubc9 (blue)-SUMO (yellow), in the same orientation.

The regions of Ubc9, which gain mobility with SUMO binding and which have reduced mobility with RanBP2 binding as discussed in Section 3.1.2 is represented in Figure 3.44. Deviations of positions in the representative structures of Clusters 1 to 4 of Ubc9-SUMO-RanBP2 complex are given in Figure 3.44 (a). Similarly, the positional deviations for Clusters 4 to 6 of Ubc9-SUMO are given in Figure 3.44 (b). The amplitudes of the motion for the region from Val27 to Met39 of Ubc9 are reduced upon RanBP2 binding, and the direction of motion of the N-terminal helix of Ubc9 is changed.

The restriction of conformational flexibility in terms of the change in the quaternary structure implied to the Ubc9-SUMO complex by RanBP2 binding can also be visualized via the deviations of representative structures from the crystal structure (Figure 3.45). The amplitude and direction changes imply SUMO possesses different dynamics in Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes. The direction of motion is altered towards the target binding sites and mobility of SUMO is restricted in the RanBP2 bound complex.

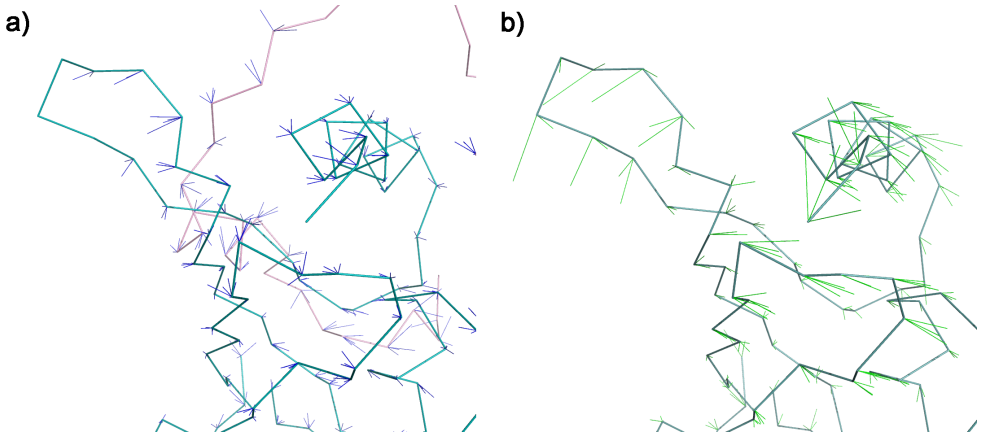

Figure 3.44. Alterations in Ubc9 mobility. The fragment of the crystal structure of Ubc9 (blue)-SUMO-RanBP2 (pink) is given in ribbon representation. a) Deviations in Ubc9-SUMO-RanBP2 simulation. b) The deviations from Ubc9-SUMO simulation.

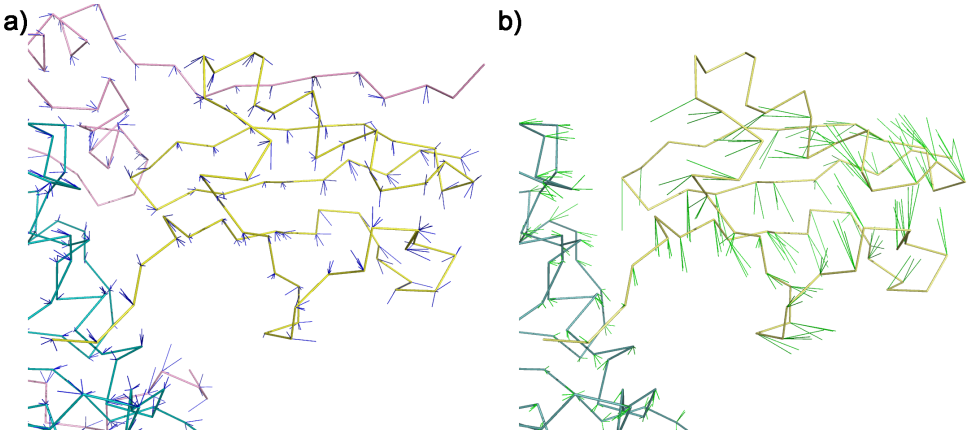

Figure 3.45. Alterations in SUMO mobility. The fragment of the crystal structure of Ubc9(blue)-SUMO(yellow)-RanBP2(pink) is given in ribbon representation. a) Deviations in Ubc9-SUMO-RanBP2 simulation. b) Deviations in Ubc9-SUMO simulation.

3.4. Correlations between the Fluctuations

The cross-correlations between the positional fluctuations are calculated for Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes. The correlations are calculated using the whole trajectory as well as the time intervals defined by the clustering analysis (Section 3.3). For Ubc9-SUMO complex, the time period in which SUMO changes its orientation, from 6 ns to 12 ns, is also investigated separately. The time intervals for which the correlations are calculated for each trajectory is tabulated in Table 3.3.

Table 3.3. Time Intervals for calculated Fluctuation Correlations

Ubc9-SUMO		Ubc9-SUMO-RanBP2	
Limits (ns)	Length (ns)	Limits (ns)	Length (ns)
0.5 - 6	5.5	0.4 - 28	27.6
6 - 12	6	28 - 50	22
12 - 24	12	0.4 - 50	49.6
24 - 31	7		
31 - 42	11		
42 - 58	16		
0.5 - 58	57.5		

The correlation between the positional fluctuations define the regions that display cooperated motions. The changes in correlations in different complexes may have functional implications, especially when combined with the interactions of the regions that is already known of functional importance. The comparatively analysis of the correlations observed for Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes are linked with the motions that may act as markers for RanBP2 binding in Ubc9-SUMO complex, as defined in Sections 3.1.2 and 3.3. The differences are also discussed in the context of the proposed allosteric function of RanBP2 binding (Pichler *et al.*, 2002; Reverter and Lima, 2005).

3.4.1. Correlations involving RanBP2 binding sites

3.4.1.1. Anchor-like behavior of Ubc9, residues Val27 to Glu42. The region between residues Val27 to Glu42 of Ubc9 constitutes two β-sheets and the loop in between. This region is discussed in previous sections with respect to its high mobility, and the association of this mobility with the change in the quaternary structure of Ubc9-SUMO complex (Sections 3.1.2 and 3.3). The correlations of the region Val27 to Glu42 with the remaining residues of the Ubc9-SUMO complex for the whole trajectory of 58 ns, point out this region to have low correlations with remaining Ubc9 residues (Figure 3.46).

When the correlations are calculated for different time windows throughout the trajectory, this region is observed to become negatively correlated with the remaining residues of Ubc9 as simulation time increases (Figures 3.47 to 3.52). This gained anchor-like behavior becomes more evident as SUMO changes its orientation. The negative correlation is maximized in the interval where this region has highest mobility. The increase in the mobility of the region Val27-Glu42, accompanies the orientation change of SUMO to a less favorable conformation for conjunction to targets and the negative direction of the motion to the remaining Ubc9 residues further distinguishes the residues Val27 to Glu42. This motion of Ubc9 may function as a signal for RanBP2 binding to Ubc9-SUMO complex.

3.4.1.2. Existing correlations among RanBP2 binding sites of Ubc9 and SUMO. The residues between Val27 and Glu42 of Ubc9 also form a contact surface for RanBP2. Two regions of SUMO, residues Phe36-Leu47 and residues Asp73-Ile88, are either on the β-sheets packed with RanBP2 or are in the close vicinity of these β-sheets, in Ubc9-SUMO-RanBP2 complex. For Ubc9-SUMO-RanBP2 complex, there are correlated fluctuations between the region Val27-Glu42 of Ubc9, and these two SUMO regions (Figure 3.53). The correlations are stable in both time intervals of 0.4-28 and 28-50 ns along the trajectory, defined by the cluster analysis (Figures 3.54 and 3.55).

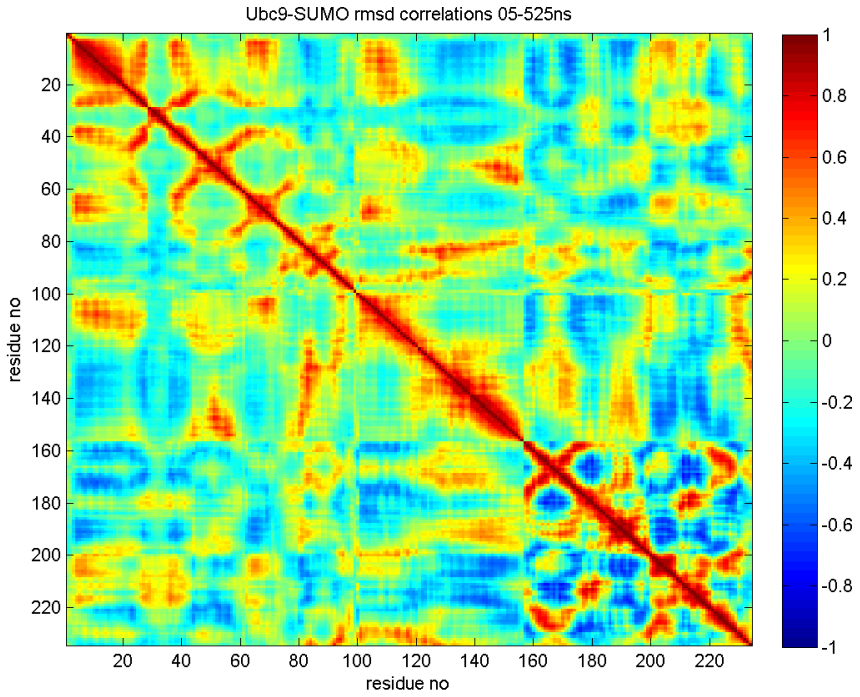

Figure 3.46. Fluctuation Correlations for Ubc9-SUMO complex, from 0.5 to 58 ns

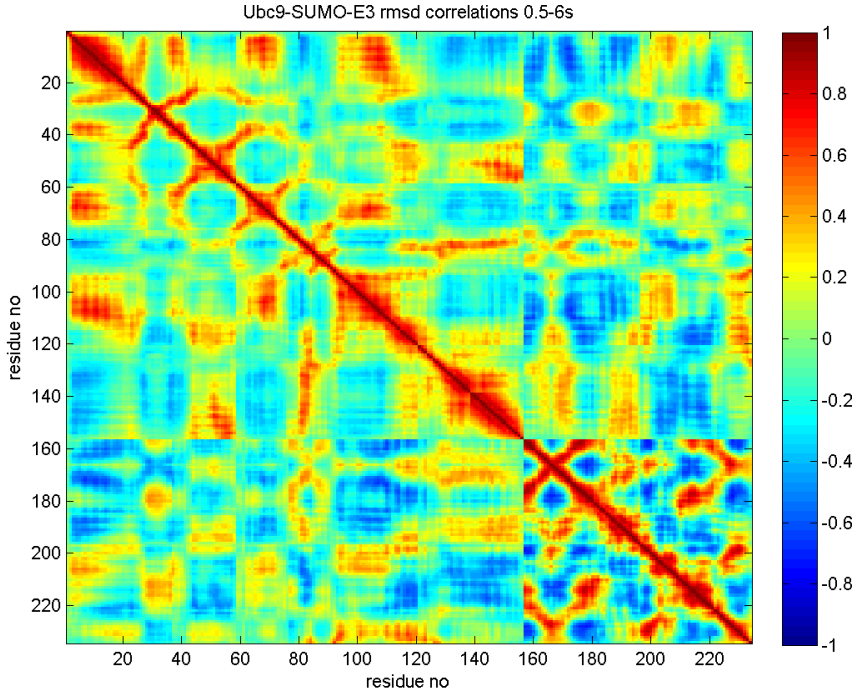

Figure 3.47. Fluctuation Correlations for Ubc9-SUMO complex, from 0.5 to 6 ns

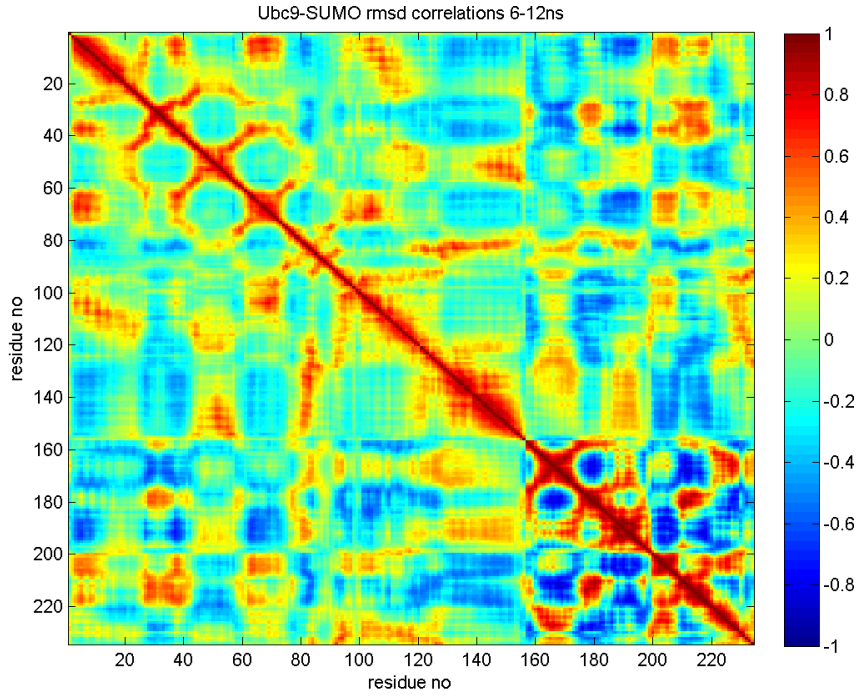

Figure 3.48. Fluctuation Correlations for Ubc9-SUMO complex, from 6 to 12 ns

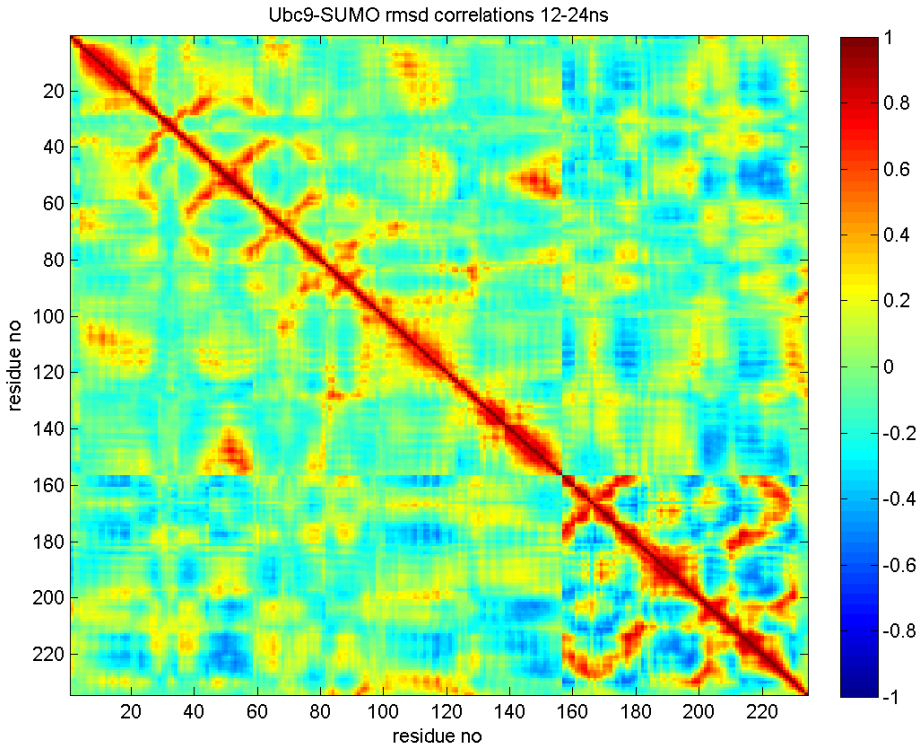

Figure 3.49. Fluctuation Correlations for Ubc9-SUMO complex, from 12 to 24 ns

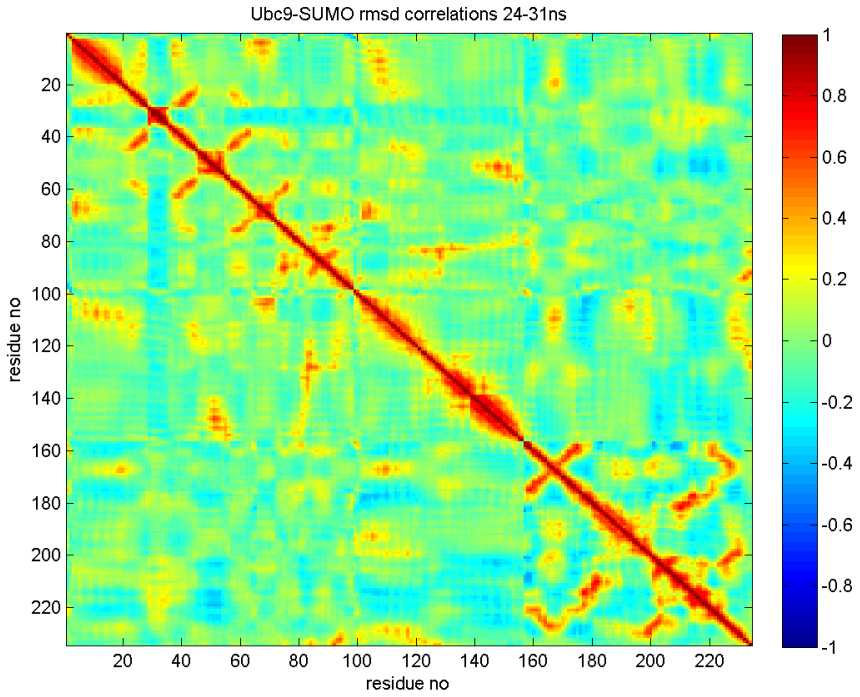

Figure 3.50. Fluctuation Correlations for Ubc9-SUMO complex, from 24 to 31 ns

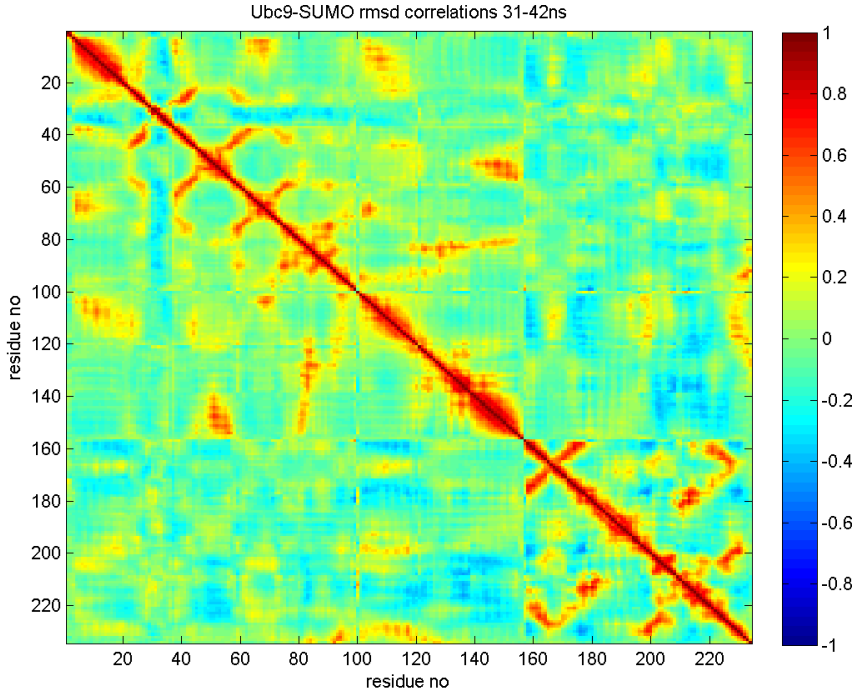

Figure 3.51. Fluctuation Correlations for Ubc9-SUMO complex, from 31 to 42 ns

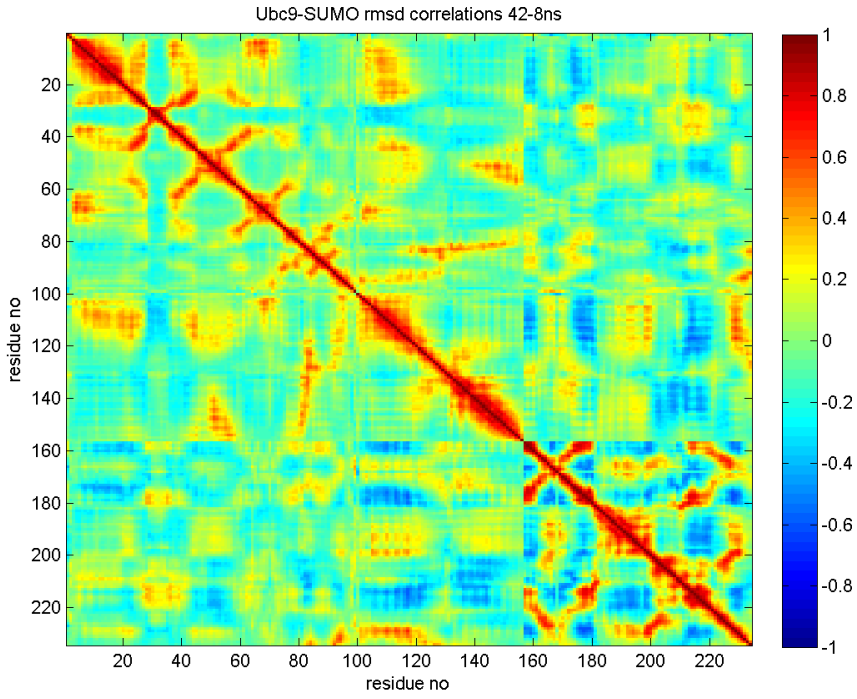

Figure 3.52. Fluctuation Correlations for Ubc9-SUMO complex, from 42 to 58 ns

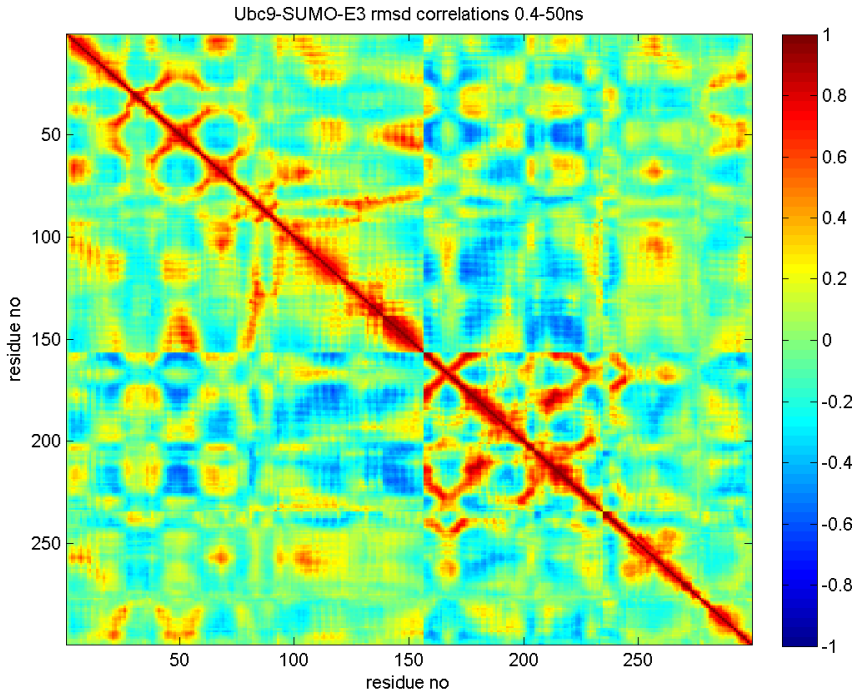

Figure 3.53. Fluctuation Correlations for Ubc9-SUMO-RanBP2 complex, from 0.4 to 50 ns

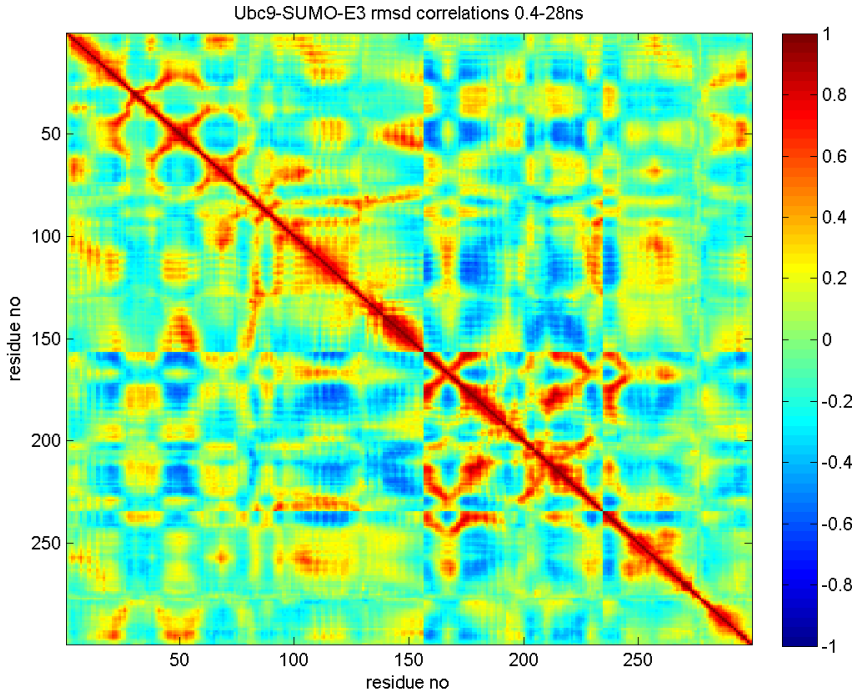

Figure 3.54. Fluctuation Correlations for Ubc9-SUMO-RanBP2 complex, from 0.4 to 28 ns

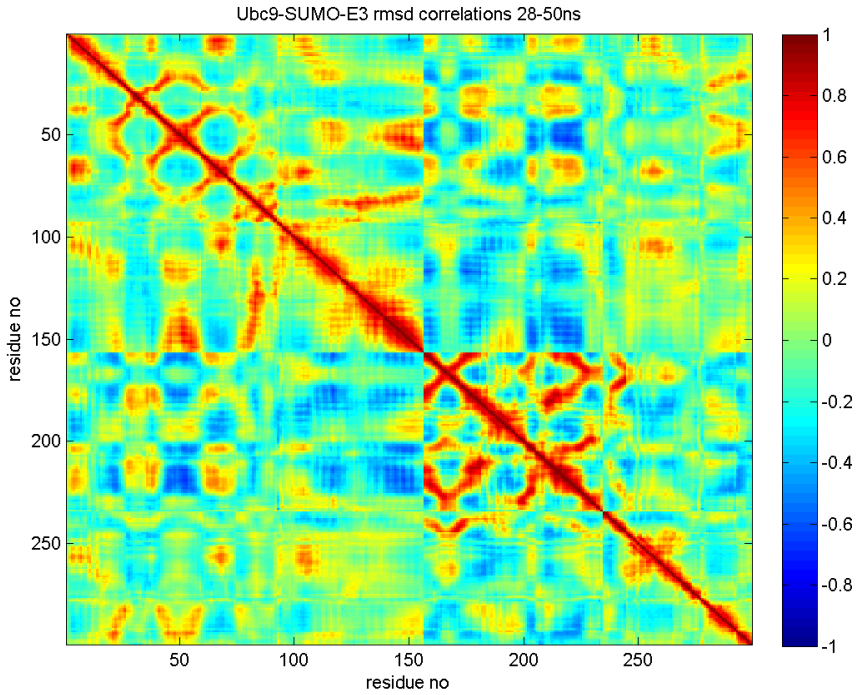

Figure 3.55. Fluctuation Correlations for Ubc9-SUMO-RanBP2 complex, from 28 to 50 ns

The correlations of the same regions are investigated in Ubc9-SUMO complex without the E3 enzyme, RanBP2, in the time windows defined by clustering analysis (Section 3.3). It is seen the correlations between Val27 to Glu42 of Ubc9 and two SUMO regions Phe36-Leu47 and Asp73-Ile88 are fluctuating. The strength of the correlations change in different time intervals, but the correlations are never lost altogether during the 58 ns simulations (Figures 3.47 to 3.52). Although these regions, are not in contact distance in Ubc9-SUMO complex, the RanBP2 binding regions on both Ubc9 and SUMO are correlated in the absence of RanBP2. Nevertheless, RanBP2 binding stabilizes this correlation. This fact suggests a possible functional role for the correlation between Ubc9 residues Val27 to Glu42 and SUMO residues Phe36-Leu47 and Asp73-Ile88

3.4.1.3. New correlations among SUMO and RanBP2 binding sites of Ubc9. Following the change in the orientation of SUMO with respect to Ubc9 in the absence of E3 enzyme, RanBP2, new correlations arise between the fluctuations of Ubc9 residues Ala15-Phe22 and SUMO residues Gly28-Glu33, in Ubc9-SUMO complex (Figure 3.50). The region of residues Ala15-Phe22 on Ubc9 include those that bind to RanBP2, Arg17, Phe22. The residues Gly28-Glu33 of SUMO corresponds to the loop between β-sheets of SUMO that pack with RanBP2 N-terminal β-sheet. In Ubc9-SUMO-RanBP2 complex, both regions correlate with RanBP2 residues. In the absence of RanBP2, the binding sites correlate with each other. It is of importance to note that the change in the orientation of SUMO adjusts these regions, Ala15-Phe22 of Ubc9 and Gly28-Glu33 of SUMO to a closer distance and the correlation may be a result of this close positioning.

3.4.2. Correlations involving functional residues in target recognition

Through the simulation of Ubc9-SUMO-RanBP2 complex, Ubc9 has intra-monomer correlations between residues His83-Ser89 and Asn121-Arg141. Residues His83-Ser89 include the HPN (His83-Pro84-Asn85) motif that has a structural role and Tyr87 that is in contact distance with the consensus sumoylation motif existent in most sumoy-

lation targets. Regarding the residues Asn121-Arg141 of Ubc9: Asn121 to Ala131 includes interactions with consensus sumoylation motif; Glu132 to Arg141 incorporate the residues that are significant in RanGAP1 recognition (Bernier-Villamor *et al.*, 2002; Wu *et al.*, 2003; Yunus and Lima, 2006). The mutations of the residues Glu132 and Tyr134 In the region Glu132-Arg141 are shown to reduce efficient sumoylation of the target RanGAP1, whereas mutations of residues Asn85, Tyr87 in His83-Ser89 reduce sumoylation efficiency for all targets (Bernier-Villamor *et al.*, 2002; Tatham *et al.*, 2003; Yunus and Lima, 2006). For the Ubc9-SUMO complex, the correlations between the region His83-Ser89 and the residues Asn121 to Ala131 are stable throughout the simulation, but the correlations between His83-Ser89 and the additional binding surface, Glu132-Arg141 are fluctuating (Figures 3.47 to 3.52). The correlation of His83-Ser89 and additional binding surface is lost after 12 ns and not observed until 24 ns, where the correlations are significantly weaker than Ubc9-SUMO-RanBP2 complex. The correlation is regained after 31 ns. In the light of the time windows generated by clustering analysis (Section 3.3), the structure occupies conformations in which the correlations are lost, during 13.24 per cent (Cluster 5) of the simulation time. Moreover, correlations between His83-Ser89 and additional binding surface are reduced in the conformations occupying an additional 12 per cent of simulation time (from 24 to 31 ns).

To sum up, the correlations among residues responsible for the structural integrity of the complex's catalytic region (His83-Ser89) and the residues that are functional in selective target recognition (Glu131-Arg141) are not stable in Ubc9-SUMO complex. Same correlations are stabilized in RanBP2 bound complex. Both regions are far from RanBP2 binding surfaces on Ubc9. Thus this stabilized correlation of the proposed allosteric role of RanBP2 in sumoylation mechanism, in addition to the restricted mobility and flexibilities of Asp100 and Lys101 of Ubc9 (Section 3.2.2.1).

3.4.3. Correlations from Gaussian Network Model Analysis

The correlations of motion of the residues are also calculated using Gaussian Network Model (GNM). The best members from each cluster are selected as the rep-

resentative simulation snapshots.

In Ubc9-SUMO-RanBP2 complex, the correlations of motion from GNM analysis do not show significant changes between representative structures from different clusters (Figures 3.56 to 3.59). The extents of correlated residues change in the Ubc9-RanBP2 binding regions at Ubc9 N-terminus and RanBP2 C-terminus. Since the correlations between the binding regions are not lost completely, the correlation changes do not imply further functional roles.

In the Ubc9-SUMO complex, three regions have changes in correlation of motions at different snapshots. Positive correlations arise between residues from Lys18 to Phe22 of Ubc9 and from Gly28 to Ser32 of SUMO, residues from Ile109 to Gly115 of Ubc9 and residue Ser30 of SUMO, residue Ala106 of Ubc9 and residue Gly56 of SUMO. The correlations are seen after the orientation change of SUMO, in the results for Cluster 4, 5 and 6 best members. The motion of SUMO brings the related groups of residues in close distance for mentioned three sets (Figures 3.56 to 3.59).

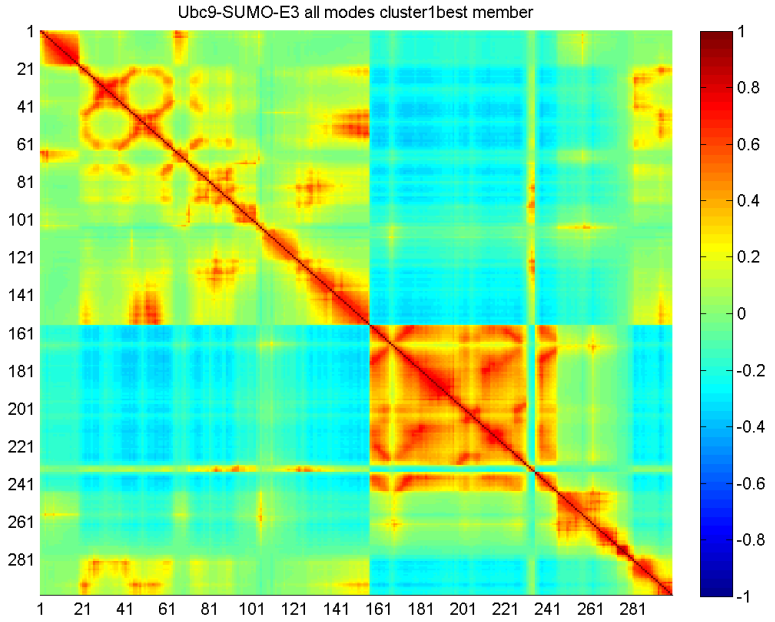

Figure 3.56. Ubc9-SUMO-RanBP2 Correlations of Motion for best member structure of cluster 1. The correlations are calculated from the GNM analysis

The correlations resulting from the model are the outcomes of this close orienta-

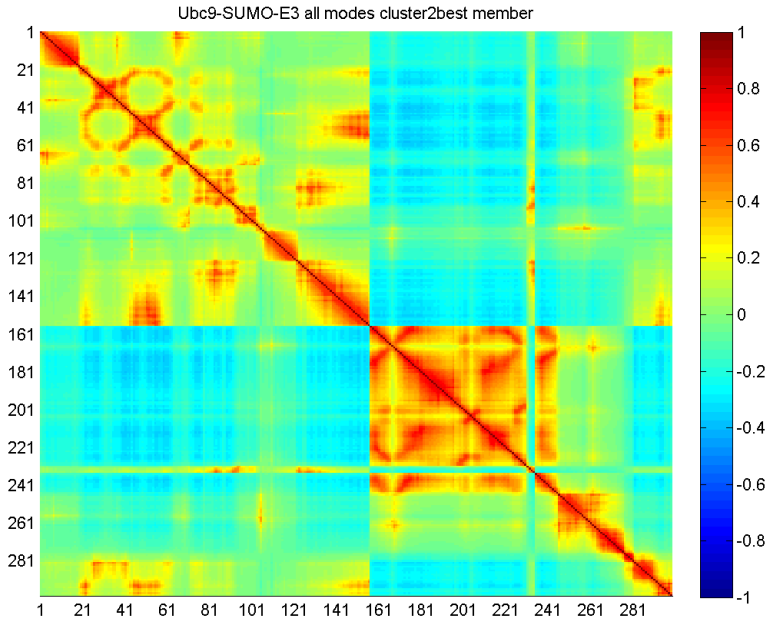

Figure 3.57. Ubc9-SUMO-RanBP2 Correlations of Motion for best member structure of cluster 2. The correlations are calculated from the GNM analysis

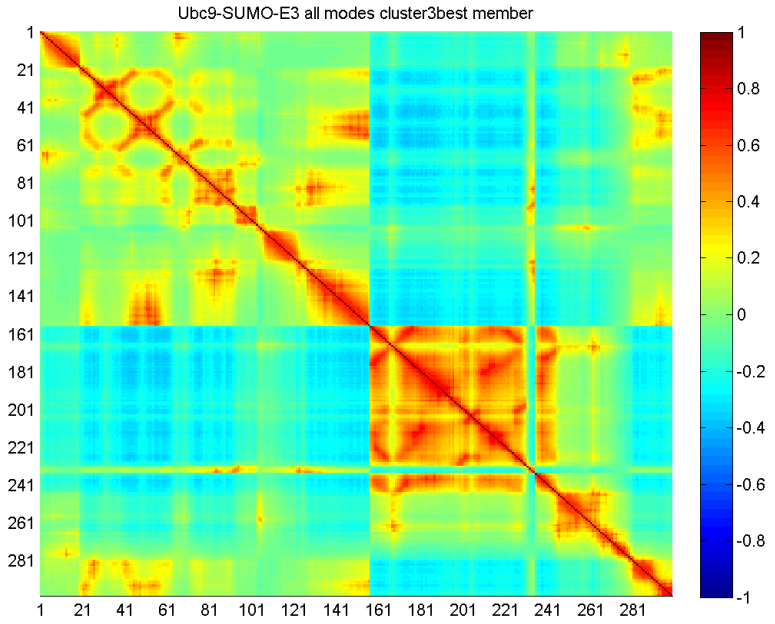

Figure 3.58. Ubc9-SUMO-RanBP2 Correlations of Motion for best member structure of cluster 3. The correlations are calculated from the GNM analysis

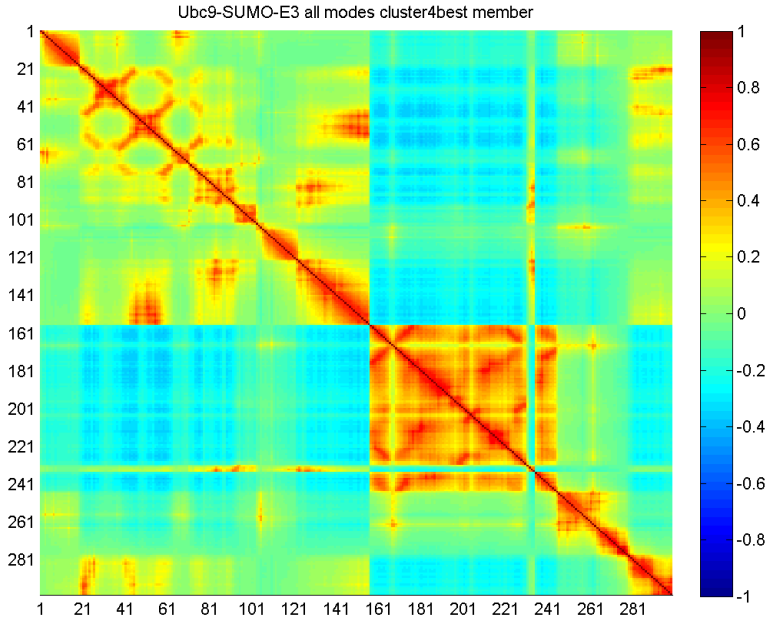

Figure 3.59. Ubc9-SUMO-RanBP2 Correlations of Motion for best member structure of cluster 1. The correlations are calculated from the GNM analysis

tion, but close investigation of the involved residues do not suggest new formed contacts. Among the three regions of interest, the correlations between residues from Lys18 to Phe22 of Ubc9 and from Gly28 to Ser32 of SUMO might be of functional importance. The same region can also be identified from MSF correlations from molecular dynamics simulations, as discussed in Section 3.4.1.3 These residues either include (as for Ubc9) or are in close proximity of (as for SUMO) RanBP2 binding sites. The correlations arising between RanBP2 binding sites may have roles is marking the Ubc9-SUMO complex for RanBP2 binding. Although other interactions can not be identified to support this model, it is important to note that the strength of the correlation is related to the motions of mobile Ubc9 region. In the representative structure of the cluster where the mobile region has highest deviation from the crystal structure (Cluster 5), the correlations between Ubc9 residues Lys18-Phe22 and SUMO residues Gly28-Ser32 are the strongest (Figures 3.60 to 3.65). The mentioned residues on both chains are correlated with RanBP2 in Ubc9-SUMO-RanBP2 complex.

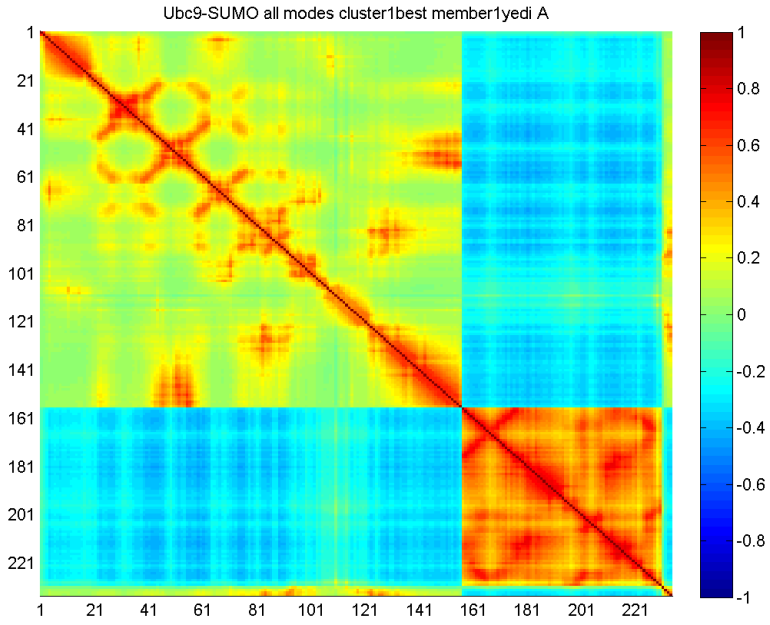

Figure 3.60. Ubc9-SUMO Correlations of Motion for best member structure of cluster 1. The correlations are calculated from the GNM analysis

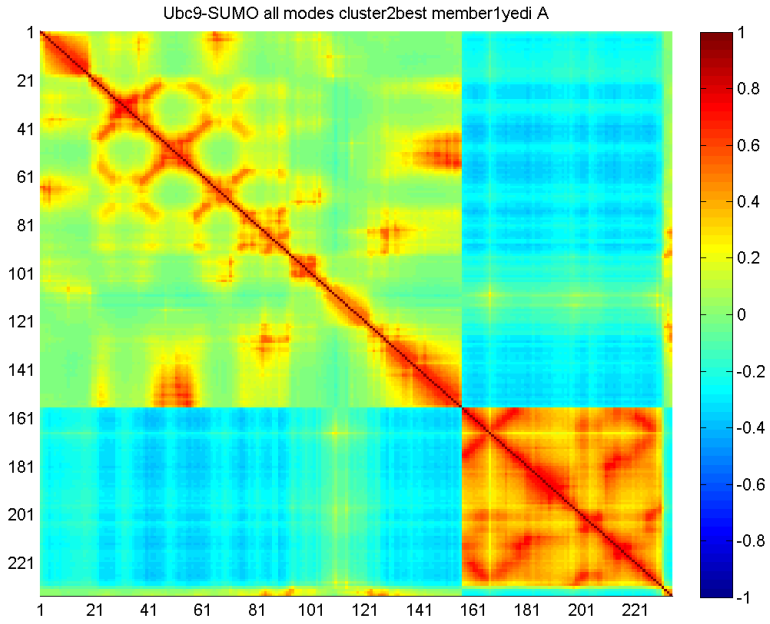

Figure 3.61. Ubc9-SUMO Correlations of Motion for best member structure of cluster 2. The correlations are calculated from the GNM analysis

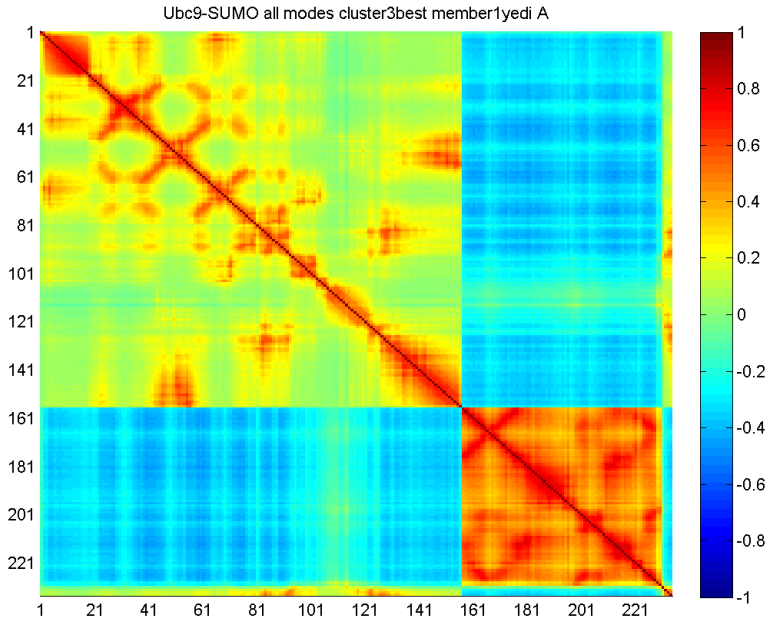

Figure 3.62. Ubc9-SUMO Correlations of Motion for best member structure of cluster 3. The correlations are calculated from the GNM analysis

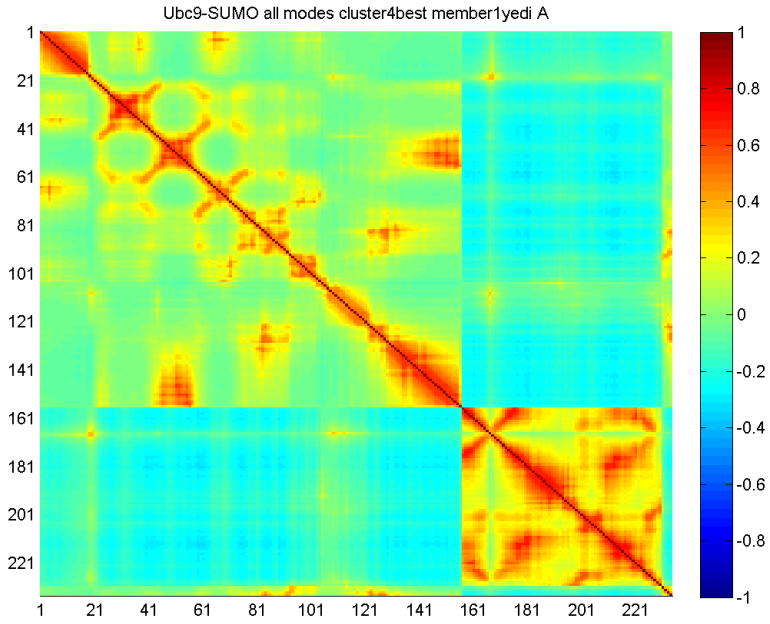

Figure 3.63. Ubc9-SUMO Correlations of Motion for best member structure of cluster 4. The correlations are calculated from the GNM analysis

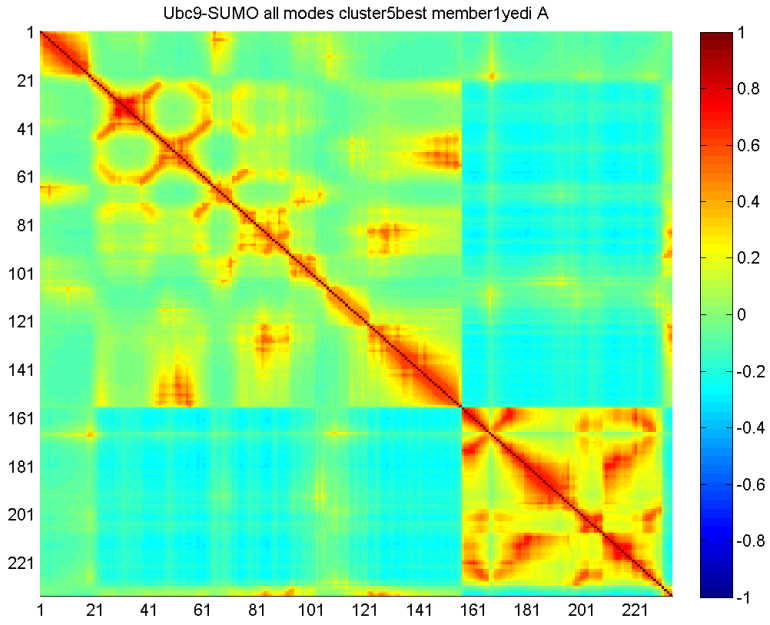

Figure 3.64. Ubc9-SUMO Correlations of Motion for best member structure of cluster 5. The correlations are calculated from the GNM analysis

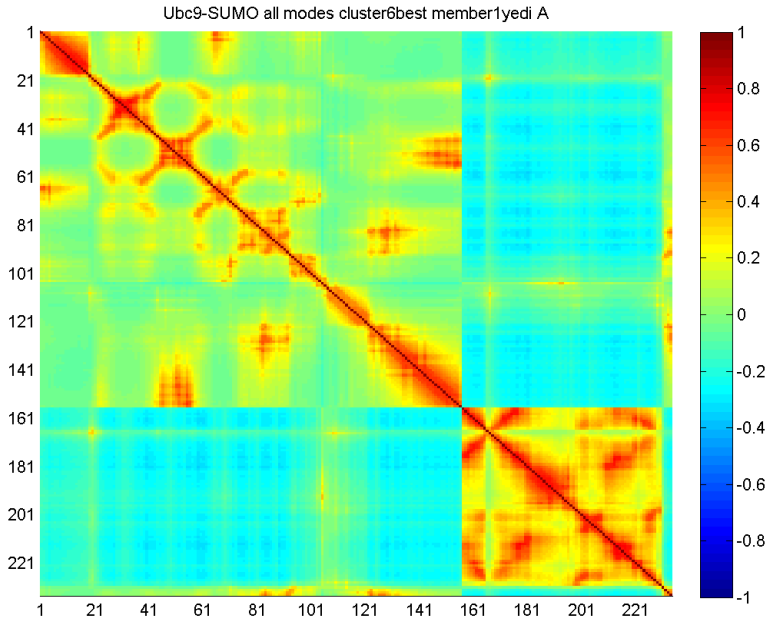

Figure 3.65. Ubc9-SUMO Correlations of Motion for best member structure of cluster 6. The correlations are calculated from the GNM analysis

4. CONCLUSIONS AND FUTURE STUDIES

4.1. Conclusions

It is expected that the comparative analysis of the dynamics of Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes would contribute to the understanding of the functioning of E2 ligase, Ubc9, with and without the E3 ligase, RanBP2, in sumoylation mechanism. Molecular dynamics (MD) simulations run on all the structures, Ubc9, SUMO, Ubc9-SUMO, Ubc9-SUMO-RanBP2, provide the conformations that these structures could sample in time in as much as the present simulation lengths reflect. The detailed analysis of the conformations reveals the following findings:

(1) The RanBP2 binding induces conformational/configurational restrictions on the quaternary structure of Ubc9-SUMO complex. The differences in the interaction networks of residues between Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes, in terms of hydrogen bonding residues, are at least in part responsible for this restriction. (2) In addition to the limitations in the overall orientations of Ubc9 and SUMO chains with respect to each other, RanBP2 binding allosterically reduces mobility and flexibility of the Ubc9 residues, Asp100 and Lys101, which are functional in target recognition. (3) RanBP2 binding is also shown to stabilize intra-correlations of Ubc9, mainly for the residues around the catalytic and additional target binding sites. (4) Furthermore, the motion and the correlated fluctuations in Ubc9-SUMO complex, which may have a role in RanBP2 binding to the complex, are also identified.

The E3 enzyme, RanBP2, is proposed to limit the available conformations/configurations of Ubc9-SUMO complex and prevent non-productive conformations (Pichler *et al.*, 2002; Reverter and Lima, 2005). This is verified with the present simulations by observing the change in the quaternary structure, the relative positions of each chain in Ubc9-SUMO complex. The clustering analysis identifies the different orientations of SUMO in Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes, as well as the relative probabilities of these conformations throughout the simulations. The orienta-

tion change observed in Ubc9-SUMO complex, which is not allowed in Ubc9-SUMO-RanBP2 complex, could be the plausible conformational restriction that E3 enzyme, RanBP2, induces on Ubc9-SUMO complex. The residues forming hydrogen bonds between Ubc9 and SUMO are different in RanBP2 bound and unbound complexes and this difference can be responsible for RanBP2 induced restrictions. The residues identified to form hydrogen bonds between Ubc9 and SUMO in the crystal structure of Ubc9-SUMO-RanGAP1-RanBP2 (Reverter and Lima, 2005), are shown to have less stable interactions in Ubc9-SUMO complex as compared to Ubc9-SUMO-RanBP2 complex. Additionally, potential hydrogen bonds between Glu118 of Ubc9-Gln92 of SUMO and Glu122 of Ubc9-Gln94 of SUMO are identified in Ubc9-SUMO complex that would stabilize SUMO in the acquired orientation.

In addition to the overall conformational restriction of Ubc9 and SUMO relative to each other, it is observed that RanBP2 binding to Ubc9-SUMO complex restricts the mobility and flexibility of Ubc9 residues Asp100 and Lys101, which are functional in target recognition (Tatham *et al.*, 2003; Yunus and Lima, 2006). The mean square fluctuations of these residues are reduced with RanBP2 binding. Also, the timed delayed auto-correlations of virtual bond vectors point out restricted orientational space for Asp100 and Lys101 in the existence of RanBP2. Additionally, these residues are at a closer distance to the active site Cys93 of Ubc9 with the presence of RanBP2. These residues are not close to the RanBP2 binding sites of Ubc9 and SUMO, which implies an allosteric effect of RanBP2 on Ubc9-SUMO complex; this would be of importance in sumoylation efficiency.

The correlations between the fluctuations of residues in Ubc9-SUMO and Ubc9-SUMO-RanBP2 complexes also show some differences. Ubc9 residues Lys74, Tyr87, Ser89, Thr91, Cys93, Asp127, Pro128, Ala129, Gln130 and Ala131, can interact with consensus sumoylation motif existent in most sumoylation targets (Bernier-Villamor *et al.*, 2002). In addition to these interaction sites, Ubc9 has a second contact site with the sumoylation target RanGAP1, via the residues of the helix, Ala131-Asn140, of Ubc9. The strong interactions between Ubc9 and this specific target, RanGAP1, via this additional binding region suggests that RanBP2 may not exert E3 activity for sumoylation

of this target (Bernier-Villamor *et al.*, 2002; Reverter and Lima, 2005). Nevertheless, the present calculations suggest that the correlations among Ubc9 catalytic sites and this additional binding region on Ubc9 are more stabile in Ubc9-SUMO-RanBP2 complex as compared to Ubc9-SUMO complex. This stabilization of correlation suggests RanBP2 indeed exerts E3 activity for the sumoylation target RanGAP1. The correlations in the mean square fluctuations of residues in Ubc9-SUMO complex highlight correlations between the region of Val27 to Glu42 of Ubc9, and the two regions in SUMO: residues Phe36-Leu47 and residues Asp73-Ile88. Although the strength of this coupling is fluctuating in time, it is not completely lost at any instant. These regions on both Ubc9 and SUMO have significance in RanBP2 binding, as seen with the conserved correlated fluctuations of these regions with RanBP2 throughout the 50 ns simulation of Ubc9-SUMO-RanBP2 complex. The pre-existing correlations between RanBP2 binding sites in Ubc9-SUMO complex are suggested to have functional roles. The analysis of the dynamics of Ubc9-SUMO complex also led to identifications of specific motions in the complex, which may mark Ubc9-SUMO complex for RanBP2 binding The residues between Val27-Met39 of Ubc9 have a significantly increased mobility in Ubc9-SUMO complex, as compared to isolated Ubc9 or Ubc9-SUMO-RanBP2 complex. RanBP2 binding reduces the mobility of this region. The change in the mobility of this region is correlated with the orientation change of SUMO. This region becomes more mobile and acquires an anchor like behavior as SUMO continues its motion away from the initial structure. Similarly, new correlations among Ubc9 and SUMO residues, which are of importance in RanBP2 binding, arise with the motion of SUMO. This motion and the correlations are suggested here as of importance for Ubc9-SUMO complex to bind RanBP2.

To sum up, Ubc9-SUMO complex is subjected to a limitation in its available conformations and the orientation of its monomers with respect to the other with RanBP2 binding. Specifically, this restriction in its dynamics helps its functional residues to find their right orientations with respect to the catalytic sites for the function. RanBP2 also stabilizes the correlations among Ubc9 residues that are functional in target recognition and catalytic activity. Additionally, Ubc9-SUMO complex have pre-existing conformational dynamics that are related to its binding to E3. This may imply that the dynamics

of a structure sets the sequence of events in association with other structures, which appears here as:: SUMO binding to Ubc9 is followed by E3 binding.

4.2. Future Studies

The binding of RanBP2 is shown to restrict the conformational/configurational space of Ubc9-SUMO complex. The effects of RanBP2 are limiting the available orientations of Ubc9 and SUMO chains with respect to each other, and limiting the mobility and flexibilities of Ubc9 residues Asp100 and Lys101, that are functional in target recognition. The Ubc9 residues in contact with RanBP2 are determined and there are experimental studies on the RanBP2 binding affinities of the Ubc9 mutants involving these residues (Reverter and Lima, 2005; Tatham *et al.*, 2005). Mutations of the residues Phe22, Val25, Val27, Leu57 and Lys59 results in severe impairment of RanBP2 binding, whereas mutations of residues Glu42, Lys48, Glu54 and Arg61 has moderate effect (Tatham *et al.*, 2005). The effects of RanBP2 binding on Ubc9 can be further investigated by studying the dynamics of mutants with reduced RanBP2 binding affinities. The observed limitations of Ubc9 functional residues, Asp100 and Lys101, in Ubc9-SUMO-RanBP2 complex can be followed in MD simulations of artificially constructed mutants for the residues listed above. The extents of restrictions would provide the information on the specific residues that contribute to the RanBP2 function, and help identification of the residue paths leading to the allosteric effect of RanBP2.

This study also identifies motions and correlations of RanBP2 binding sites in the Ubc9-SUMO complex, that may be of importance for RanBP2 binding to the complex. The finding imply SUMO binding to Ubc9 may enhance RanBP2 binding. The structure of Ubc9-RanBP2 complex only, without the SUMO protein, can also be studied. Investigation of the dynamics of SUMO binding sites on Ubc9 and RanBP2 may complement the finding of this study. The dynamics of complexes involved at different steps in the mechanism, underlying the sequence of events in sumoylation pathway, can be presented.

APPENDIX A: ALIGNMENT EFFECT on POSITIONAL MEAN SQUARE FLUCTUATIONS

The effects of alignment on calculated positional MSF is demonstrated in the MSF plots of each simulated chain. Individual alignment of each chain upon itself corrects misleading results. These results originate from poor alignment of residues of a specific chain, when the whole complex is aligned.

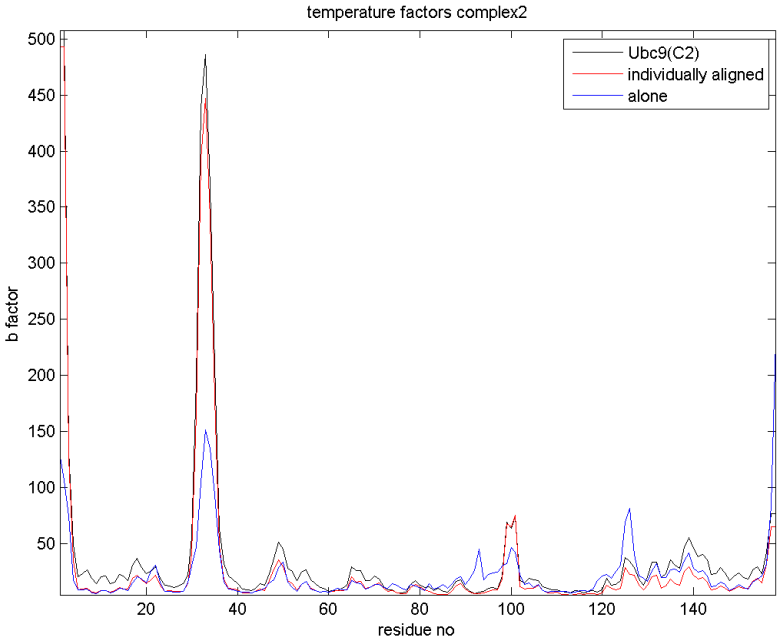

Figure A.1. Alignment effects on Ubc9 positional MSF in Ubc9-SUMO complex. The positional MSF calculated by alignment of the whole complex (black), by alignment of individual the chain upon itself (red) and from isolated Ubc9 simulations (blue).

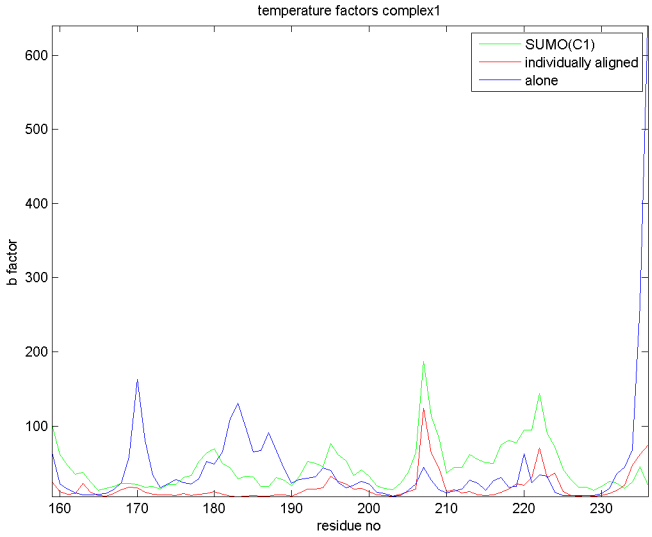

Figure A.2. Alignment effects on SUMO positional MSF in Ubc9-SUMO-RanBP2 complex. The positional MSF calculated by alignment of the whole complex (green), by alignment of individual chain (red) and from the isolated SUMO simulations (blue)

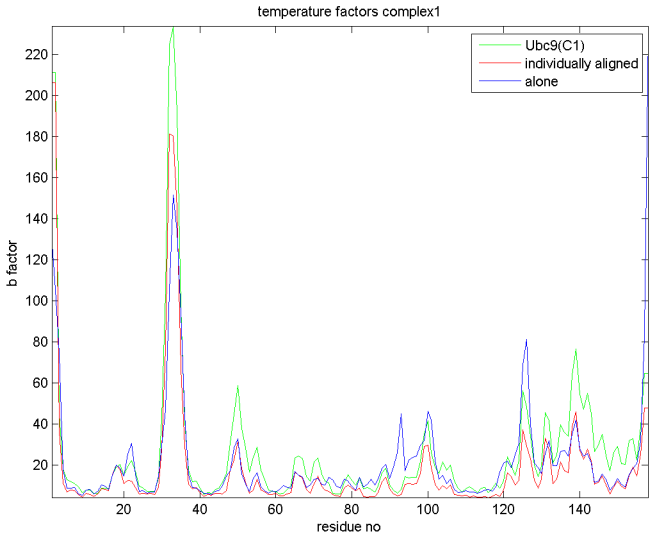

Figure A.3. Alignment effects on Ubc9 positional MSF in Ubc9-SUMO-RanBP2 complex. The positional MSF calculated by alignment of the whole complex (green), by alignment of individual the chain (red) and from the isolated Ubc9 (blue).

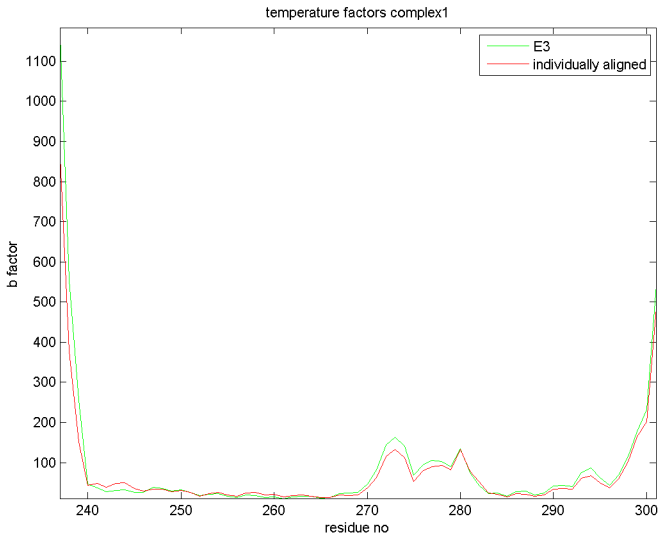

Figure A.4. Alignment effects on RanBP2 positional MSF in Ubc9-SUMO-RanBP2 complex. The positional MSF calculated by alignment of the whole complex (green) and by alignment of individual the chain upon itself (red).

APPENDIX B: CONTACT MAPS

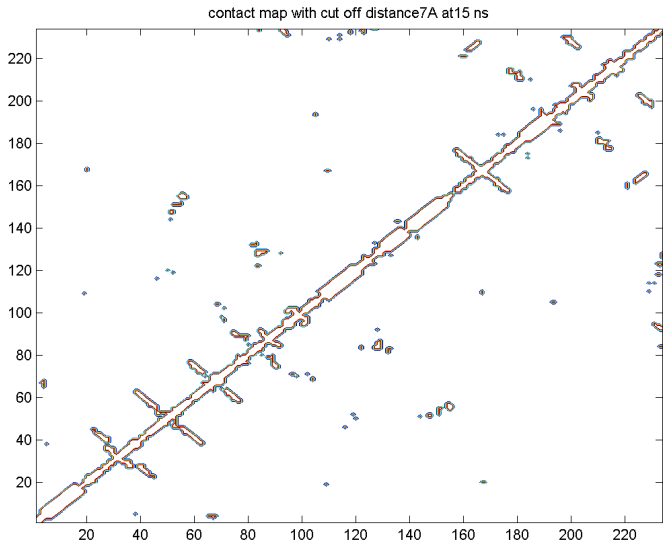

Figure B.1. Ubc9-SUMO Contact Map at 15 ns, with cut-off distance 7$\mathring{A}$

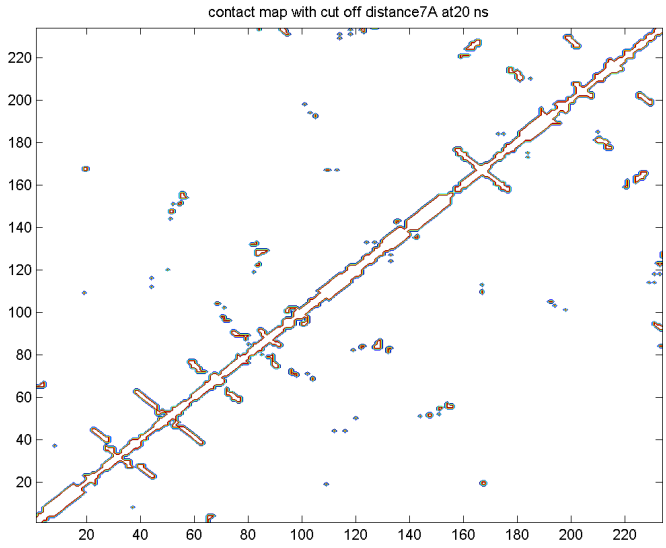

Figure B.2. Ubc9-SUMO Contact Map at 20 ns, with cut-off distance 7$\mathring{A}$

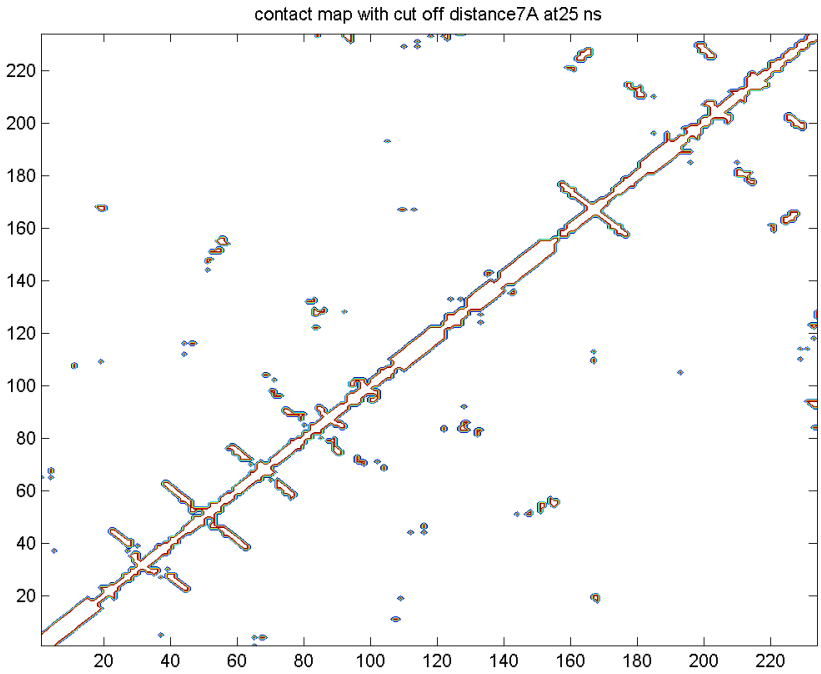

Figure B.3. Ubc9-SUMO Contact Map at 25 ns, with cut-off distance 7$\mathring{A}$

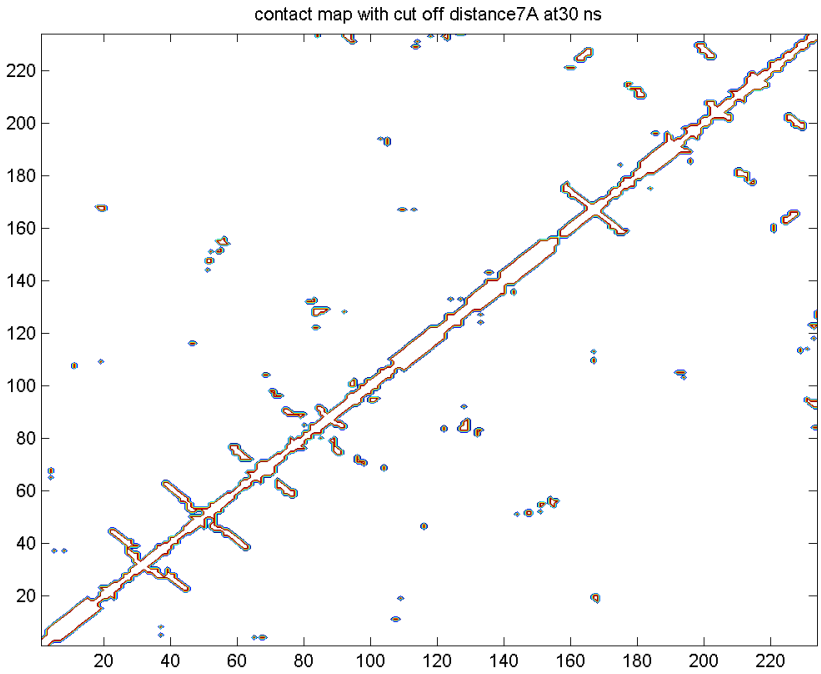

Figure B.4. Ubc9-SUMO Contact Map at 30 ns, with cut-off distance 7$\mathring{A}$

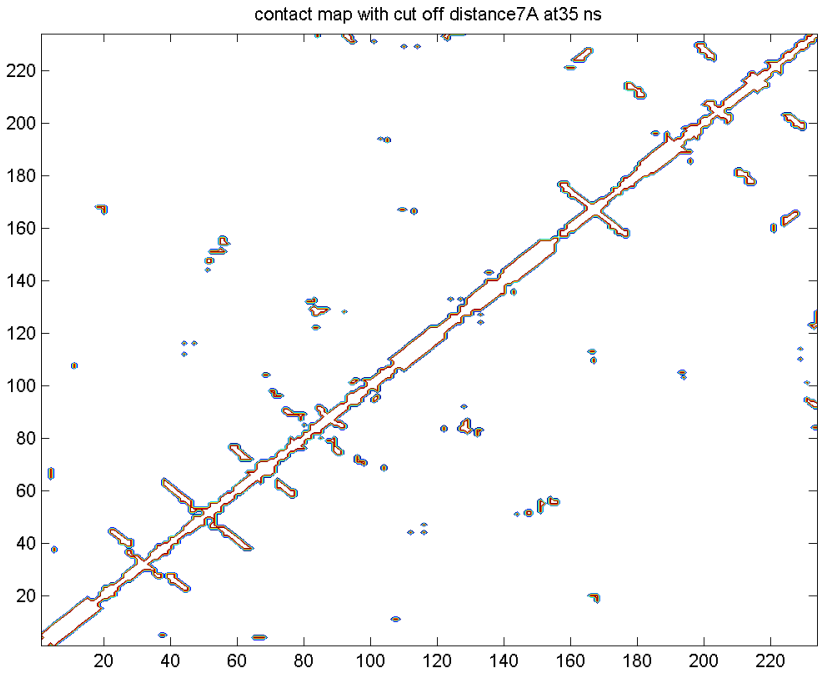

Figure B.5. Ubc9-SUMO Contact Map at 35 ns, with cut-off distance 7$\mathring{A}$

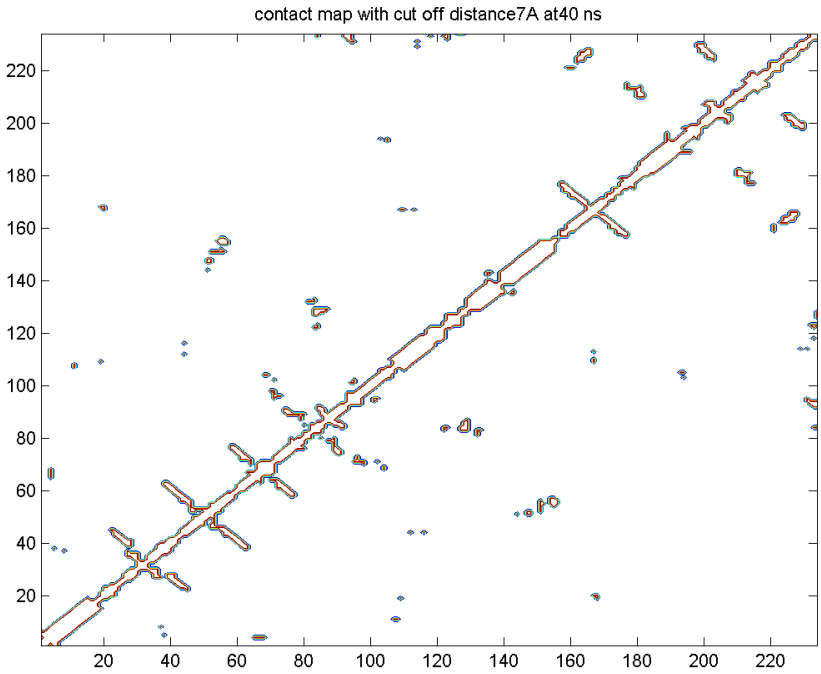

Figure B.6. Ubc9-SUMO Contact Map at 40 ns, with cut-off distance 7$\mathring{A}$

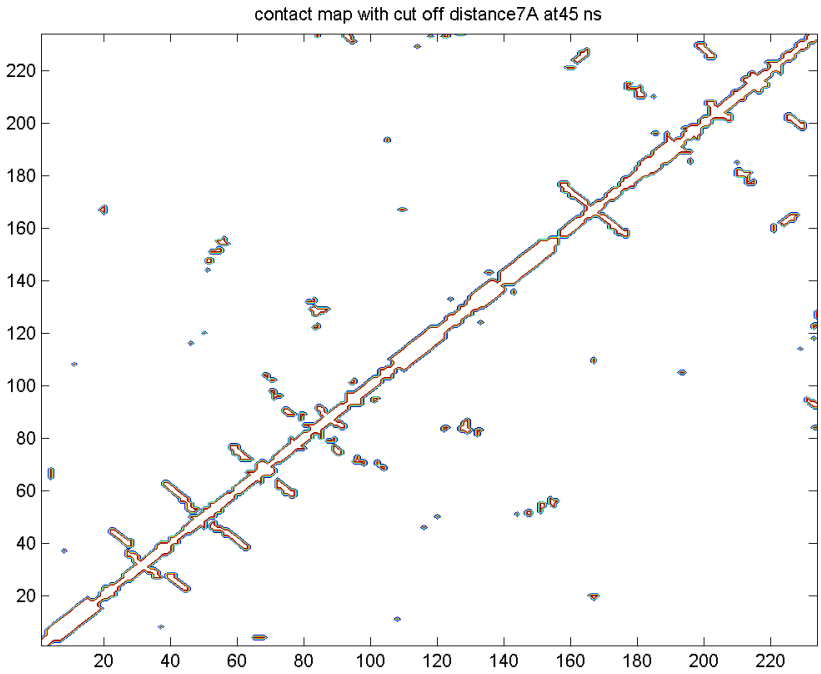

Figure B.7. Ubc9-SUMO Contact Map at 45 ns, with cut-off distance 7$\mathring{A}$

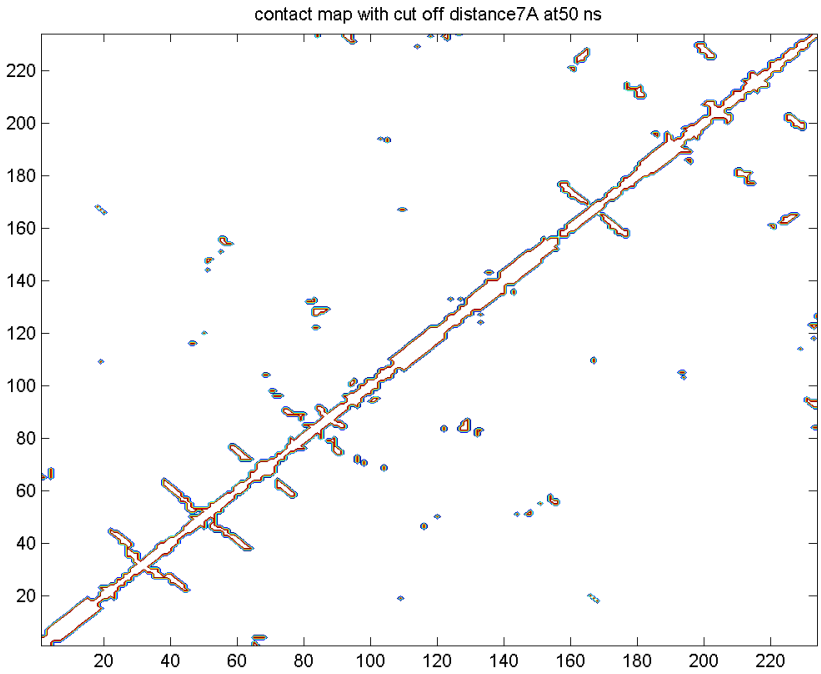

Figure B.8. Ubc9-SUMO Contact Map at 50 ns, with cut-off distance 7$\mathring{A}$

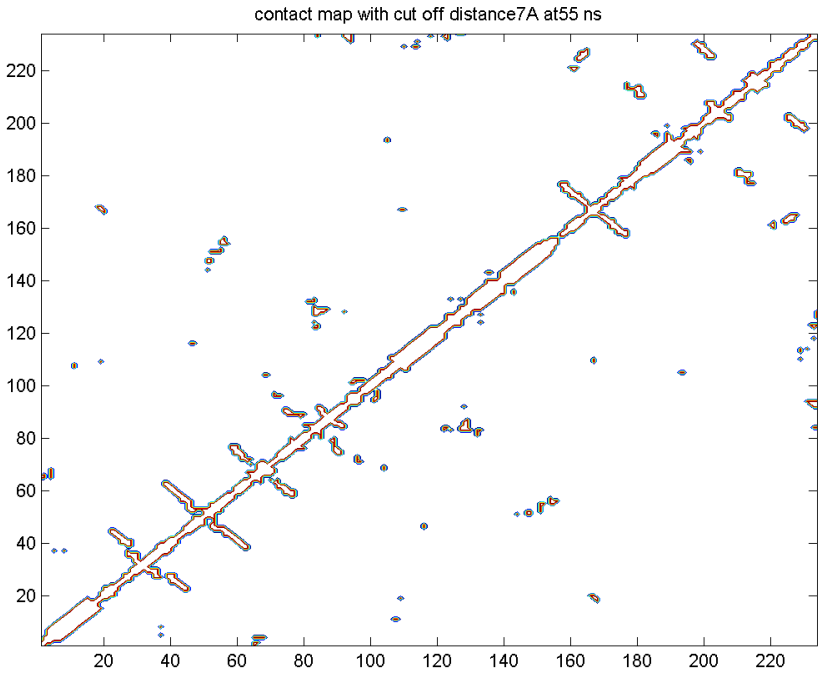

Figure B.9. Ubc9-SUMO Contact Map at 55 ns, with cut-off distance 7$\mathring{A}$

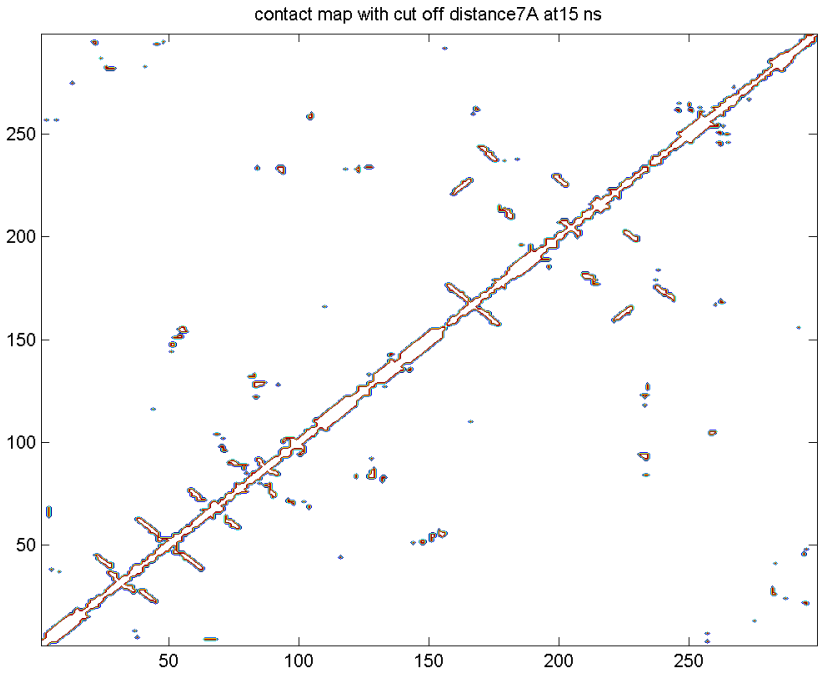

Figure B.10. Ubc9-SUMO-RanBP2 Contact Map at 15 ns, with cut-off distance 7$\mathring{A}$

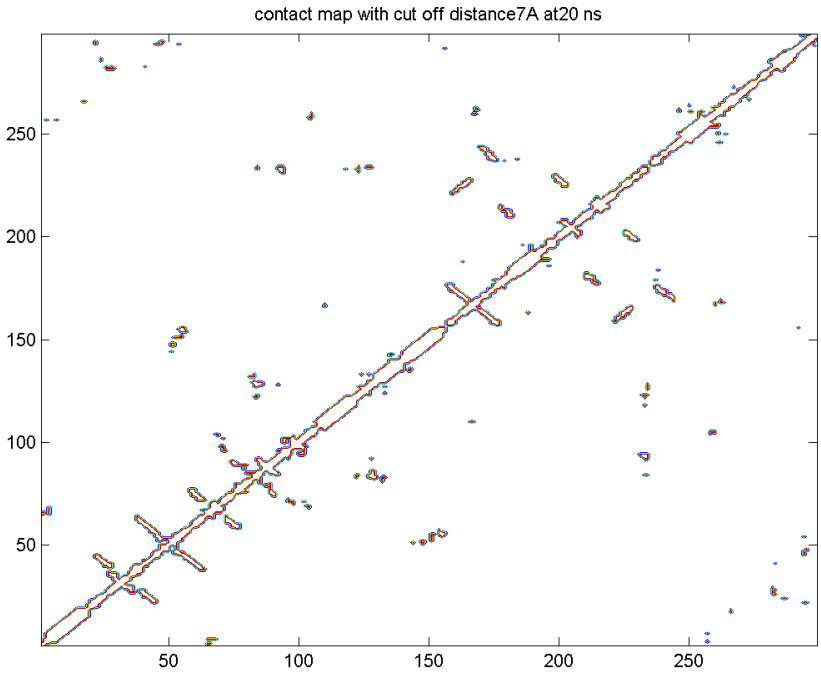

Figure B.11. Ubc9-SUMO-RanBP2 Contact Map at 20 ns, with cut-off distance 7$\mathring{A}$

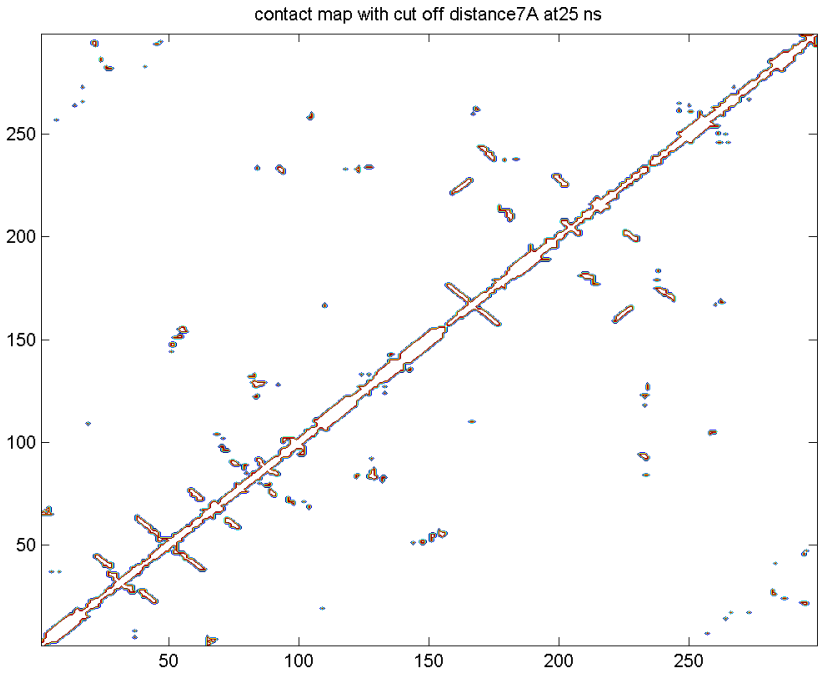

Figure B.12. Ubc9-SUMO-RanBP2 Contact Map at 25 ns, with cut-off distance 7$\mathring{A}$

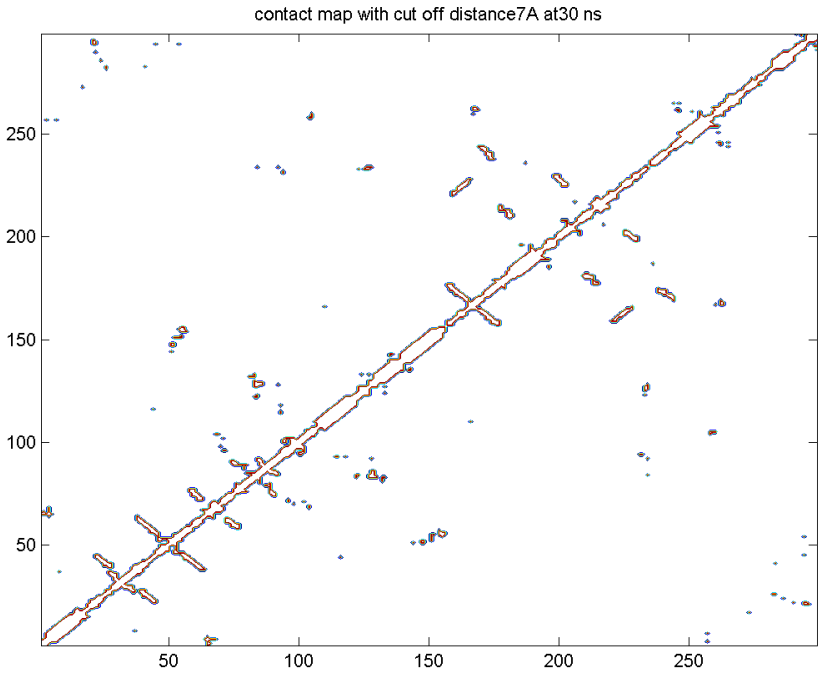

Figure B.13. Ubc9-SUMO-RanBP2 Contact Map at 30 ns, with cut-off distance 7$\mathring{A}$

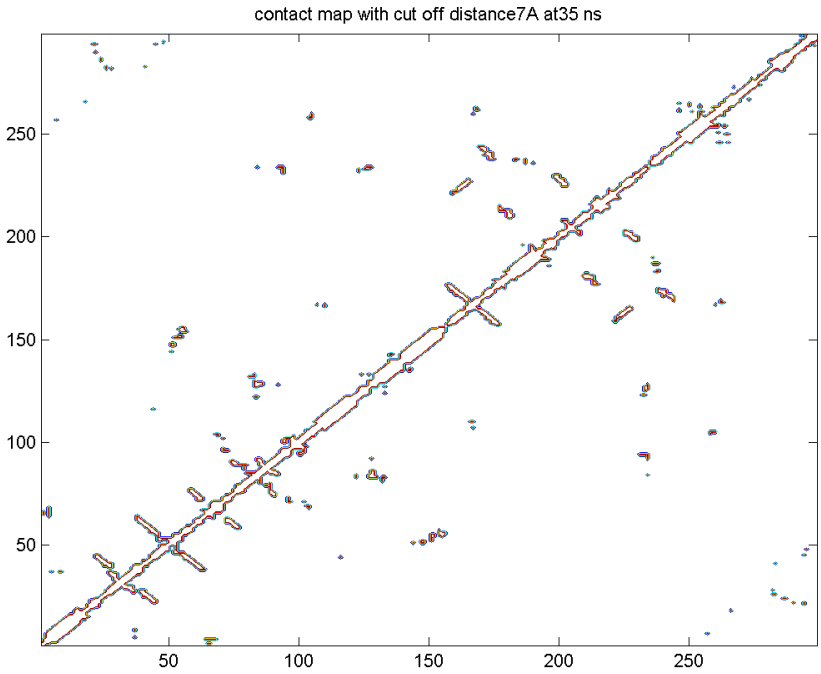

Figure B.14. Ubc9-SUMO-RanBP2 Contact Map at 35 ns, with cut-off distance 7$\mathring{A}$

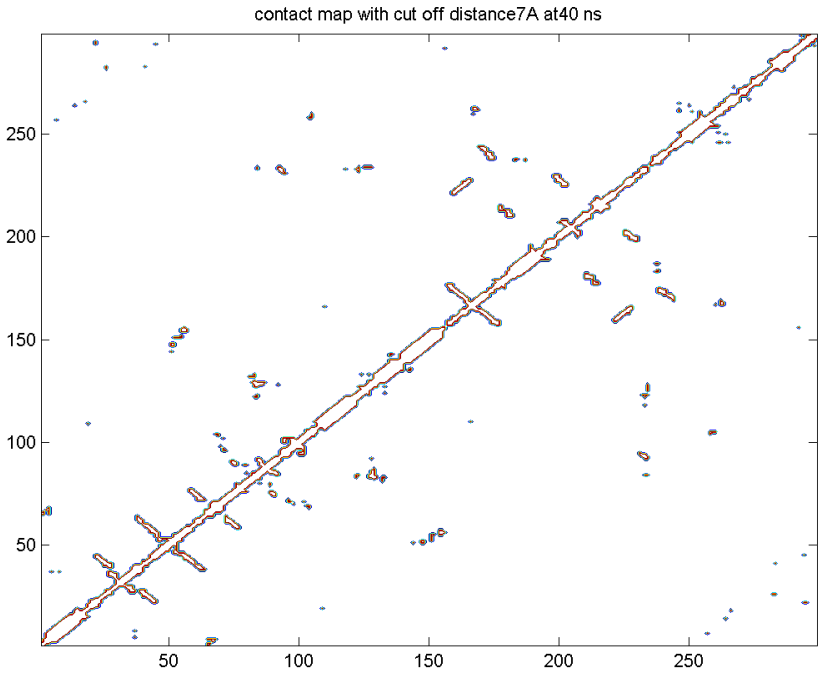

Figure B.15. Ubc9-SUMO-RanBP2 Contact Map at 40 ns, with cut-off distance 7$\mathring{A}$

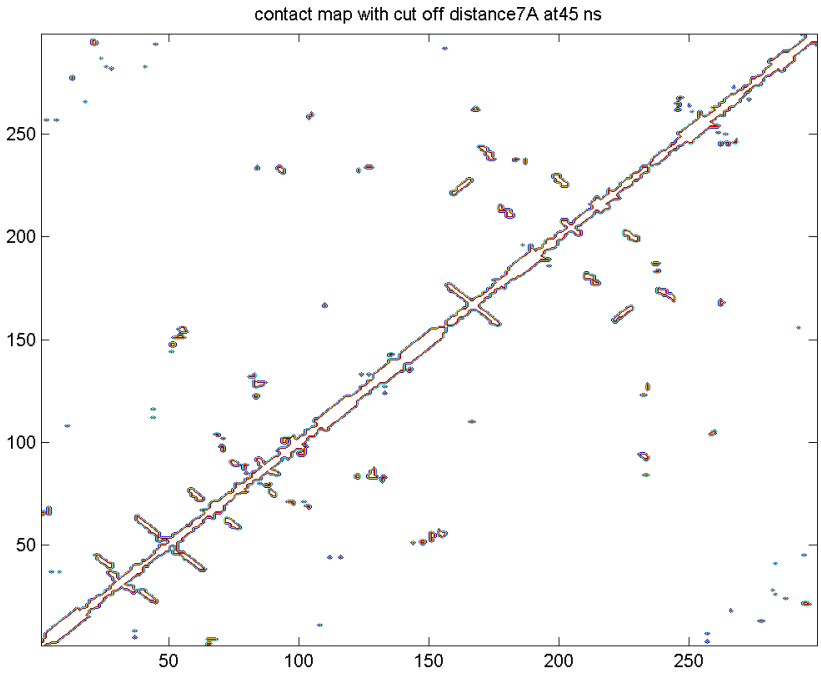

Figure B.16. Ubc9-SUMO-RanBP2 Contact Map at 45 ns, with cut-off distance 7$\mathring{A}$

APPENDIX C: CLUSTERING ANALYSIS

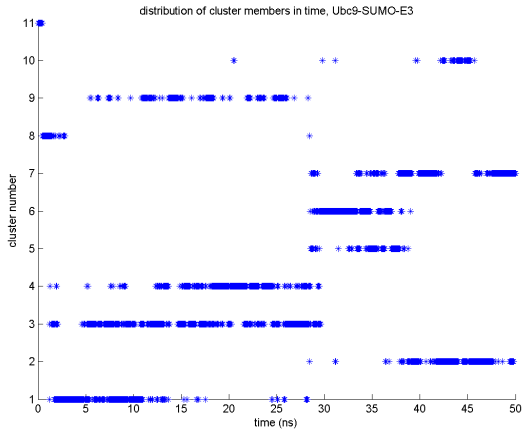

Figure C.1. Clusters of Ubc9-SUMO-RanBP2 at 1.7 $\mathring{A}$ rmsd cut-off. Time distribution of members of each cluster. The cluster number is given in the y - axis and time in nanaoseconds is given in x - axis.

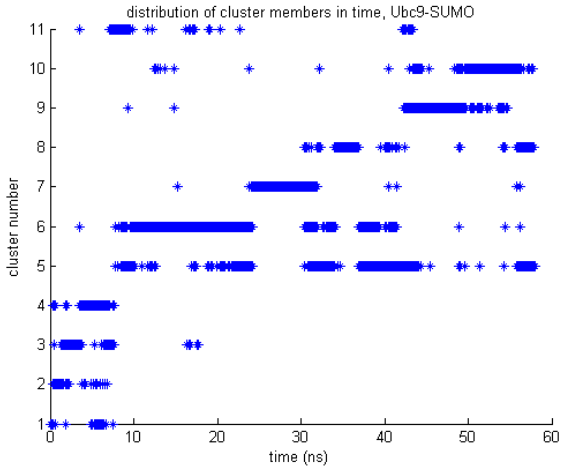

Figure C.2. Clusters of Ubc9-SUMO-RanBP2 at 2 $\mathring{A}$ rmsd cut-off. Time distribution of members of each cluster. The cluster number is given in the y - axis and time in nanaoseconds is given in x - axis.

REFERENCES

Alpaydın, E. 2004, "*Introduction to Machine Learning*", The MIT Press, London.

Bahar I., A. R. Atılgan and B. Erman, 1997, "Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential", *Folding and Design*, Vol. 2, pp. 173-181.

Bayer, P., A. Arndt, S. Metzger, R. Mahajan, F. Melchior, R. Jaenicke and J. Becker, 1998, "Structure determination of the small ubiquitin-related modifier SUMO-1", *Journal of Molecular Biology*, Vol. 280, pp. 275-286.

Becke, A. D., 1988, "Density-functional exchange-energy approximation with correct asymptotic behavior", *Physical Review*, Vol. 38, pp. 30983103.

Becke, A. D., 1993, "A new mixing of HartreeFock and local density-functional theories", *Journal of Chemical Physics*, Vol. 98, pp. 13721377.

Becke, A. D., 1993, "Density-functional thermochemistry. III. The role of exact exchange", *Journal of Chemical Physics*, Vol. 98, pp. 56485652.

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, 2001, "Structural basis for E2 mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1", *Cell*, Vol. 108, pp. 345-356.

Bischoff, F. R., C. Klebe, J. Kretschmer, A. Wittinghofer and H. Ponstingl 1994, "RanGAP1 induces GTPase activity of nuclear Ras related Ran", *The Proceedings of the National Academy of Sciences Online (US)*, Vol. 91, pp. 2587-2591.

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, "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.

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.

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.

Chennubhotla, C., A. J. Rader, L-W. Yang and I. Bahar, 2005, "Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies", *Physical Biology*, Vol. 2, pp. S173-S180.

DeLano, W. L., 2002, *The PyMOL Molecular Graphics System*, http://www.pymol.org.

Essman, U., L. Perera, M.L. Berkowitz, T.A. Darden, H. Lee and L.G. Pedersen, 1995, "A smooth Particle Mesh Ewald method", *Journal of Chemical Physics*, Vol. 103, pp. 8577-8593.

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. 377395.

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.

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

Lee, C., W. Yang and R. G. Parr, 1988, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density", *Physical Review*, Vol.37, pp. 785789.

Liu, Q., C. Jin, X. Liao, Z. Shen, D. J. Chen and Y. Chen, 1999, "The Binding Interface between an E2 (UBC9) and a Ubiquitin Homologue (UBL1)", *Journal of Biological Chemistry*, Vol.274, pp. 16979-16987.

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

Martin, S., A. K. Wilkinson, A. Nishimune and J. M. Henley, 2007, "Emerging extranuclear roles of protein SUMOylation in neuronal function and dysfunction", *Nature Reviews Neuroscienc*, 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.

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

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

Müller, S., H. Cartsen, G. Pyrowolakis, and S. Jentsch, 2001, "SUMO, Ubiquitin's Mysterious Cousin", *Nature Reviews Molecular Cell Biology*, Vol. 2, pp. 202-210.

Okura, T., L. Gong, T. Kamitani, T. Wada, I. Okura, C. F. Wei, H. M. Chang and E. T. Yeh, 1996, "Protection against Fas/APO-1- and tumor necrosis factor-mediated cell death by a novel protein, sentrin", *The Journal of Immunology*, Vol. 157, pp. 4277-81.

Pichler, A., A. Gast, J. S. Seeler, A. Dejean and F. Melchior, 2002, "The Nucleoprotein RanBP2 Has SUMO1 E3 Ligase Activity", *Cell*, Vol. 108, pp. 109-120.

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 and Molecular Biology*, Vol. 11, pp. 984-991.

Panse, V. G., U. Hardeland, T. Werner, B. Kuster and E. Hurt, 2004, "A proteome wide approach identifies sumoylated substrate proteins in yeast", *Journal of Biological Chemistry*, Vol. 279, pp. 41346-41351.

Rader, A. J., C. Chennubhotla, L-W. Yang and I. Bahar, 2006, "The Gaussian Network Model: Theory and Applications" in *"Normal Mode Analysis. Theory and Applications to Biological and Chemical Systems"*, CRC Press, Taylor Francis Group, US.

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

Ryckaert, J. P., G. Ciccotti and H.J.C. Berendsen, 1977, "Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes", *Journal of Computational Physics*, Vol. 23, pp. 327-341.

Shen, Z.,P. E. Pardington-Purtymun, J. C. Comeaux, R. K. Moyzis and D. J. Chen, 1996, "UBL1, a human ubiquitin-like protein associating with human RAD51/RAD52 proteins", *Genomics*, Vol. 36, pp. 271-279.

Tatham, M. H., E. Jaffray, O. A. Vaughan, J. M. 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", *Journal of Biological Chemistry*, Vol. 276, pp. 35368-35374.

Tatham, M. H., K. Suhkmann , S. Kim, B. Yu, E. Jaffray, J. Song, J. Zheng, M. S. Rodriguez, R. T. Hay and Y. Chen,2003, "Role of an N-Terminal Site of Ubc9 in SUMO-1, -2, and -3 Binding and Conjugation", *Biochemistry*, Vol. 42, pp. 9959 -9969.

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. Jaffrey, 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 and Molecular Biology*, Vol. 12, pp. 67-74

van Waardenburg, R. C. A. M., D. M. Duda, C. S. Lancaster, B. A. Schulman and M. A. Bjornsti, 2006, "Distinct Functional Domains of Ubc9 Dictate Cell Survival and Resistance to Genotoxic Stress", *Molecular and Cellular Biology*, Vol. 26, pp. 4958-4969

Wang, J., R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, 2004, "Development and testing of a general amber force field", *Journal of Computational Chemistry*, Vol. 25, pp. 1157-1174

Wohlschlegel, J. A., E. S. Johnson, S. I. Reed and J. R. Yates 3rd, 2004, "Global

analysis of protein sumoylation in *Saccharomyces cerevisiae*", *Journal of Biological Chemistry*, Vol. 279, pp. 45662-45668.

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

Yang, M., C. T. Hsu, C. Y. Ting, L. F. Liu, J. Hwang, 2006, "Assembly of a polymeric chain of SUMO1 on human topoisomerase I in vitro", *Journal of Biological Chemistry*, Vol. 281, pp. 8264-8274.

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

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

Zhao, Y., S. W. Kwon, A. Anselmo, K. Kaur and M. A. White, 2004, "Broad spectrum identification of cellular small ubiquitin-related modifier (SUMO) substrate proteins", *Journal of Biological Chemistry*, Vol. 279, pp. 20999-21002.

Zhou, W., J. J. Ryan and H. Zhou, 2004, "Global analyses of sumoylated proteins in Saccharomyces cerevisiae. Induction of protein sumoylation by cellular stresses", *Journal of Biological Chemistry*, Vol. 279, pp. 32262-32268.