

**Role of RAP80-ABRAXAS-BRCA1-complex in the DNA Damage
Repair: Structural Investigations to Reveal the Molecular Complexity**

By

Vikrant

[LIFE09200804001]

Tata Memorial Centre

Mumbai

A thesis submitted to the

Board of Studies in Life Sciences

In partial fulfilment of requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Of

HOMI BHABHA NATIONAL INSTITUTE




June, 2015

HOMI BHABHA NATIONAL INSTITUTE

Recommendations of the Viva Voce Board

As members of the Viva Voce Board, we certify that we have read the dissertation prepared by Vikrant entitled "**Role of RAP80-ABRAXAS-BRCA1-complex in the DNA Damage Repair: Structural Investigations to Reveal the Molecular Complexity**" and recommend that it may be accepted as fulfilling the dissertation requirements for the Degree of Doctor of Philosophy.

 Date: 30/06/2015

Chairperson: - Dr. Vinay Kumar

 Date: 30/06/2015

Guide/Convenor - Dr. Ashok K Varma

 Date: 30/06/2015

Member 1: - Dr. Rajiv Kalraiya

 Date: 30/06/2015

Member 2: - Dr. Prasanna Venkatraman

 Date: June 30, 2015


External Examiner: Prof. B. Gopal

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

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

Date: 30/06/2015

Place: Kharagpur


Ashok K Varma
(Guide)

STATEMENT BY AUTHOR

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

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

Navi Mumbai

Date: June 30, 2015



(Vikrant)

DECLARATION

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

Navi-Mumbai

Date: June 30, 2015

A handwritten signature in black ink, appearing to read 'Vikrant', written over a horizontal line.

(Vikrant)

List of Publications

Journal

- 1) Role of MERIT40 in stabilization of BRCA1-complex: A protein-protein interaction study; **Vikrant**, *Ulka Sawant, Ashok K Varma* | Biochem Biophys Research Communication | 2014 | Volume 446 | 1139–1144
- 2) Mislocalization of BRCA1-complex due to ABRAXAS Arg361Gln mutation; **Vikrant**, *Rajan Kumar, Quadir, S. Waghmare, Ashok K Varma*; Journal of Biomolecular Structure and Dynamics | 10.1080/07391102.2014.945484
- 3) Structural and Functional Implication of RAP80 Δ 81Glu Mutation; **Vikrant**, *Rajan Kumar, Lumbini R. Yadav, Pallavi Nakhwa, Sanjeev K. Waghmare, Peyush Goyal, Ashok K. Varma*, PloS One | 2013 | Volume 8 | Issue 9 | e72707
- 4) Structural and functional characterization of the MERIT40 to understand its role in DNA repair; **Vikrant**, *Pallavi Nakhwa, Dilip C. Badgujar, Rajan Kumar, Khushboo K.S. Rathore, Ashok K. Varma* | Journal of Biomolecular Structure and Dynamics |2013| <http://dx.doi.org/10.1080/07391102.2013.843473>
- 5) Preliminary Crystallographic Studies of BRCA1 BRCT- ABRAXAS Complex; *Dilip C. Badgujar, Ulka Sawant, **Vikrant**, Lumbini Yadav, M. V. Hosur and Ashok K. Varma* | Acta Crystallography | 2013 | F69 |1401-1404
- 6) Tetrameric ZBRK1 DNA binding domain has affinity towards cognate DNA in absence of Zinc ions; *Lumbini R Yadav, Mahamaya N. Biswal, **Vikrant**, M.V Hosur, Ashok K Varma* | Biochem Biophys Research Communication | 2014 | Volume 450 | 283–288

- 7) Phosphopeptide specific binding of BRCA1-BRCT to KIF1; Vikrant, Ashok Varma
(Manuscript under preparation)

Conferences

- 1) Structural investigation of RAP80; A novel BRCA1 interacting protein involved in the mediation of DNA damage repair function; Vikrant, Ashok K Varma | Acta Crystallography | 2011 | A67 | 346
- 2) Structural and functional basis of MERIT40: an h-BRCA1-RAP80 complex protein involved in DNA damage repair; Vikrant, Ashok K Varma | J Cancer Research Ther | 2012 | 8 | S28

Others

- 1) Conserved Residues at the MAPKs Binding Interfaces that Regulate Transcriptional Machinery; Bhanu P Jagilinki, Nikhil Gadewal, Harshal Mehta, Hafiza Mahadik, Vikrant, Anamika Pandey, Ulka Sawant, Prasad A Wadegaonkar, Peyush Goyal, Satish Kumar, Ashok K Varma | Journal of Biomolecular Structure and Dynamics | 2014 | 10.1080/07391102.2014.915764
- 2) Over-expression, purification and isotopic labeling of a tag-less human glucose dependent insulinotropic polypeptide; Rakesh C. Chandarana, Vikrant, Ashok K. Varma, Anil Saran, Evans C. Coutinho, Jacinta S. D'Souza | 3 Biotech | 2013 | 13205-013-0181

Date June 30, 2015



Signature of Student

ACKNOWLEDGEMENTS

I take this opportunity to offer my sincere thanks to my mentor Dr. Ashok K. Varma for his guidance, incessant encouragement and prudent suggestions. His infectious optimism, bright spirits, never say no attitude and belief in me were invaluable throughout the tenure. Without him it would have been impossible to make this journey successful.

I am grateful to Prof. S.V. Chiplunkar (Director- ACTREC), Dr. R. Sarin (Former Director ACTREC), Dr. S. Zingde (Former Deputy Director ACTREC) for providing research facility and excellent infrastructure. I am fortunate to have Dr. M.V. Hosur, Dr. Vinay Kumar, Dr. V. Prasanna, and Dr. R. Kalraiya as my doctoral committee members. Their critical analysis and helpful suggestions have contributed significantly to the work.

I am extremely thankful to the common instrument facility, Macromolecular Crystallography and X-Ray Diffraction Facility, BTIS, library, IT, administration and accounts department of ACTREC for their constant help and support.

I am thankful to all my lab members Mrs Ulka Sawant, Dr. Dilip Badgujar, Lumbini, Bhanu, Qadir, Rajan, all my batch mates, trainees of varma lab and student's community of ACTREC who helped me in innumerable ways during the entire period.

I owe a lot more than I can express in words to my parents and other family members for all their sacrifices and affectionate blessings. I would reserve my heartiest thanks for my better-half, Nitu, for being my pillar of strength through thick and thin and loving me for who I am.

Last but not least, the Almighty who has provided me strength and blessed me with the presence of many people who have helped me in this research. Though, all cannot be mentioned, but none is forgotten.

CONTENTS

	Page No.
Synopsis	1-14
List of Abbreviations	15-16
List of Table	17
List of Figures	18-21
Chapter-1: <i>Introduction and Review of Literature</i>	22
1.1 Cancer	23
1.2 DNA damage and genome conservation	24
1.3 DNA damage response (DDR) and BRCA1 Associated proteins	32
1.3.1 BRCA1-Associated RING Domain (BARD1)	35
1.3.2 BRCA1/BRCA2-Containing Complex 36 (BRCC36)	37
1.3.3 Receptor Associated Protein 80 (RAP80)	37
1.3.4 Coiled Coil Domain Containing Protein 98 (CCDC98)	39
1.3.5 Mediator of RAP80 Interaction and Targetting 40 (MERIT40)	39
1.3.6 BRCA1/BRCA2-Containing Complex (BRCC45)	40
1.4 Ubiquitination and its role in cell signaling	41
1.5 Conclusion	44
Chapter-2: <i>Materials and Methods</i>	46
2.1 Gene Cloning	47
2.1.1 Preparation of target gene	47
2.1.2 Primer Designing	47
2.1.3 PCR amplification	47

2.1.4	Restriction Digestion	48
2.1.5	Ligation	48
2.1.6	Transformation	49
2.1.7	Competent cells	49
2.2	Protein expression and purification	52
2.2.1	Protein Purification	53
2.2.2	Regeneration of GST resin	57
2.2.3	Regeneration of Nickel resin	58
2.3	Protein characterization	58
2.3.1	Protein estimation	58
2.3.2	Mass Spectrometry	59
2.3.3	Circular Dichroism	61
2.3.4	Fluorescence Spectroscopy	63
2.3.5	Differential Scanning Calorimetry	69
2.3.6	Limited proteolysis	70
2.3.7	Molecular Modeling and Docking	72
2.4	Protein-Protein Interaction	77
2.4.1	Isothermal Titration Calorimetry	78
2.4.2	Surface Plasmon Resonance	79
2.4.3	Pull-down assay	80
2.5	Protein oligomerization and size characterization	81
2.5.1	Glutraldehyde cross-linking	82
2.5.2	Native gel electrophoresis	82

2.5.3	Dynamic light scattering	83
2.6	Crystallization	84
2.6.1	Methods of crystallization	86
2.6.2	X-Ray Diffraction	87
3.0	Domain organization of expression constructs	90
Chapter-3: <i>Structural Basis & Protein-Protein Interactions Studies of RAP80</i>		
3.1	Introduction	92
3.2	Material and Methods	94
3.3	Results and discussion	98
3.3.1	Cloning, expression and purification of RAP80 functional domain	98
3.3.2	Structural insight into RAP80-Ub complex using <i>in silico</i> analysis	98
3.3.3	Structural insights into RAP80-UB complex using <i>in vitro</i> analysis	101
3.3.4	Binding interaction of RAP80 wild-type and Δ E81 with di-Ub	104
3.4	Crystallization of functional domain of RAP80	105
3.5	Conclusion	106
Chapter-4: <i>Biophysical & Structural Characterization of MERIT40</i>		
4.1	Introduction	109
4.2	Material and methods	110
4.3	Results and discussion	113
4.3.1	Cloning, expression and purification of MERIT40	113
4.3.2	Molecular modeling	114
4.3.3	Domain organization of MERIT40	116
4.3.4	Role of MERIT40 in DNA damage repair	117

4.3.5	Oligomeric property	120
4.3.6	Dimer interface of MERIT40	123
4.3.7	Structural analysis	124
4.3.8	Assessment of stability and unfolding	125
4.3.9	Clinical implication of MERIT40 variants	127
4.4	Crystallization	129
4.5	Conclusion	129
Chapter-5: <i>Structural and functional analysis of ABRAXAS</i>		130
5.1	Introduction	131
5.2	Material and methods	133
5.3	Results and discussion	137
5.3.1	Sequence alignment	137
5.3.2	Cloning, expression and purification of functional domains	138
5.3.3`	Oligomeric characterization of wild-type and R361Q mutant	139
5.3.4	Structural organization	139
5.3.5	Thermal stability and folding pathway	142
5.3.6	Binding interaction	144
5.4	Crystallization of ABRAXAS	145
5.5	Conclusion	145
Chapter-6: Structural associates of <i>BRCA1-complex</i>		147
6.1	Introduction	148
6.2	Material and methods	149
6.3	Results and discussion	150

6.3.1	Structural characterization	151
6.3.2	Qualitative interaction analysis	152
6.3.3	Quantitative interaction analysis of MERIT40 and BRCA1	152
6.3.4	Interaction analysis of MERIT40 and ABRAXAS	153
6.3.5	Determining minimal binding region of ABRAXAS for MERIT40	154
6.3.6	Interaction analysis of MERIT40 and RAP80	155
6.4	Co-crystallization of MERIT40 with ABRAXAS & BRCA1-BRCT	155
6.5	Conclusion	156
	Chapter-7: Structure Studies of BRCA1-KIF1b Complex	158
7.1	Introduction	159
7.2	Material and methods	162
7.2.1	Expression and purification of BRCA1 BRCT domain	162
7.2.2	Interaction analysis using ITC	162
7.2.3	Protein crystallization	163
7.2.4	Diffraction data collection and processing	164
7.2.5	Structure solution of BRCA1-KIF1b structure	164
7.3	Results and discussion	166
7.3.1	Purification and characterization of BRCA1 BRCT domain	166
7.3.2	Binding isotherm of BRCA1 BRCT domain and KIF1b peptide	166
7.3.3	Crystallization and X-ray diffraction of BRCA1 BRCT domain	168
7.3.4	Structure solution of the complex	168
7.3.5	Structure analysis and validation	170

7.3.6	Secondary and tertiary structure analysis	171
7.3.7	Structural alignment of different chains of BRCA1 BRCT	171
7.3.8	Structural alignment of BRCA1 BRCT and KIF1b	172
7.4	Conclusion	173
	Chapter-8: <i>Summary and future prospects</i>	175
8.1	RAP80	176
8.2	MERIT40	178
8.3	ABRAXAS	179
8.4	BRCA1-complex	180
8.5	Future perspective	181
	References	182-195
	Publications Reprints	



Homi Bhabha National Institute

Ph. D. PROGRAMME

- | | |
|--|--|
| 1. Name of the Student: | Vikrant |
| 2. Name of the Constituent Institution: | Tata Memorial Centre, Advanced
Centre for Treatment Research and
Education in Cancer |
| 3. Enrolment No. : | LIFE09200804001 |
| 4. Title of the Thesis: | “Role of RAP80-ABRAXAS-BRCA1-
complex in the DNA Damage Repair:
Structural Investigations to Reveal the
Molecular Complexity” |
| 5. Board of study: | Life Science |

SYNOPSIS

Cancer is a complex disease which arises due to abnormal cell division. Instead of undergoing death, cancer cells continue to grow and form new abnormal cells which may leads to metastasis. Breast cancer can be classified into two categories; Ductal carcinoma in situ (non-invasive breast cancer) and Lobular carcinoma in situ (pre-cancer). About 5 - 10% of breast cancer cases are reported to be hereditary in nature, and the most common cause is an inherited defects in the genes, *BRCA1* and *BRCA2* [1] [2]. *BRCA1* is a well established tumor suppressor, and plays an essential roles in various cellular events like apoptosis, cell cycle checkpoint control, transcription and DNA damage repair [3]. BRCA1 protein performs these functions by forming molecular complexes with different cellular binding partners. BRCA1-complex comprises BRCA1, RAP80, ABRAXAS and MERIT40, and is crucial for DNA damage repair mechanism. Several mutations have been identified in BRCA1 associated proteins that disrupt interactions with binding partners, thereby causing various pathogenic implications [4] [5]. The present study aims to structurally characterise the BRCA1 binding proteins to gain functional insights of pathogenic mutation(s) reported in RAP80 and ABRAXAS. This work would also highlights the importance of BRCA1-complex recruitment to the DNA damage site.

The thesis comprises of seven chapters and a summary chapter.

Chapter 1 provides an introduction about the role of BRCA1-complex in the DNA damage repair pathway.

DNA damage repair is a vital process for the survival of living organisms, and mainly occurs through two pathways, homologous recombination (HR) and non homologous end joining (NHEJ). The BRCA1-complex is involved in the homologous recombination

pathway, while DNA-PK and KU70/80 complexes are involved in NHEJ pathway. In response to DNA damage, a series of phosphorylations are triggered leading to the formation of polyubiquitin chains on the histone protein, H2AX. These polyubiquitin chains are recognized by the Ubiquitin Interacting Motif (UIM) present at the C-terminus of RAP80, thereby helping in the recruitment of BRCA1-complex to the DNA damage site. RAP80 forms a stable complex with BRCA1 through the binding partner, ABRAXAS. At the C-terminus, ABRAXAS contains a consensus sequence motif pS-X-X-F (pS – Phosphoserine, X-any amino acids, F-Phenylalanine), through which it recognizes and binds to phospho-specific binding domain BRCA1-BRCT (BRCA1 C-terminal repeats) [6, 7] [2, 8]. MERIT40 (Mediator of RAP80 Interaction and Targeting 40) protein is also a key component of the BRCA1-complex, and is essential for maintaining the complex integrity. However, the exact nature of interactions among components of the BRCA1-complex is not yet established. The two tandem BRCTs are the hotspots for clinical mis-sense mutations, which invariably disrupt phosphorylation dependent interactions with the binding partners [9, 10]. The in-frame deletion mutation of the conserved glutamic acid residue at position 81, c.241-243delGAA (Δ E81), found in a patient diagnosed with breast cancer, is located in the UIM1 motif of RAP80 [5]. Similarly, the familial mutation (R361Q) at the C-terminus of ABRAXAS in Finnish population is associated with cancer. It is suspected that these mutations disassemble the BRCA1-complex [11]. In view of the critical role played by BRCA1-complex in DNA damage repair, we have undertaken structural & biophysical studies on the component proteins of the BRCA1-complex.

Chapter 2 is a brief description of the various methodologies and techniques used to carry out research work reported in the thesis.

The gene of interest was sub-cloned from mammalian c-DNA construct to bacterial expression vector and site- directed - mutagenesis was performed on wild- type template to incorporate the mutation. The protein was purified using different chromatography techniques (affinity and size exclusion chromatography), and highly purified protein was characterized using biophysical tools, such as Circular Dichroism, Fluorescence, Differential Scanning Calorimetry, Dynamic Light Scattering and Mass Spectrometry, etc. Protein-protein interactions were analyzed using Isothermal Titration Calorimetry, Surface Plasmon Resonance and pull down assays. Structural insight was obtained with crystallographic methods, and in-silico tools were used for molecular modeling and docking.

Chapter 3 describes about the structures of RAP80 and its deletion mutant $\Delta E81$. Binding studies of the wild-type and mutant RAP80 with di-ubiquitin molecules have also been mentioned in this chapter.

The wild- type gene was sub-cloned into the bacterial expression system, pGEX-kT vector, and site- directed -mutagenesis has been carried out to produce the RAP80 $\Delta E81$ variant. The protein was purified in large quantities for structural and interactions studies.

The molecular weight of the purified wild-type protein determined using mass spectrometry was 14.95 kDa. CD and fluorescence spectroscopic studies indicate

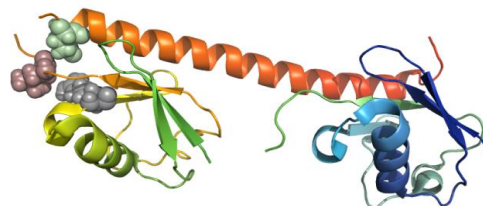


Figure 1: RAP80 wild type and Di-ubiquitin

that the wild-type protein is properly folded and having well defined secondary structure. Crystallization trials were carried out using commercially available Hampton Research screen solutions, in hanging and sitting drop vapour diffusion method. A molecular model of the wild-type protein has been constructed using Swiss-Model server [12]. The modeled structure suggests that RAP80 is predominately a helical protein (**Figure 1**), a result consistent with CD measurements. Further, modeling of the mutant protein indicates disruption in the longest helix, leading to distorted structure pattern. Interestingly, the mutant is found to be relatively more susceptible to digestion by proteases like trypsin and chymotrypsin. The mutant also has lesser affinity for ubiquitin when probed using pull down and Surface Plasma Resonance techniques.

In conclusion, helical to random structure transition in the mutant resulted in loss of several weak intermolecular hydrogen bond and hydrophobic interactions between the UIMs and Di-Ub (K-63 linked), thereby making the binding interactions unfavorable for ubiquitin. Unstable nature of mutant/ di-ubiquitin complex may be responsible for defective recruitment of BRCA1-complex to the DNA damage sites.

Chapter 4 focuses on biophysical, biochemical and structural characterization of MERIT40.

MERIT40 (Mediator of RAP80 Interaction and Targeting 40)

is one of the recently discovered components of the BRCA1-complex, and its function is not yet clearly established. However, it is shown to significantly influence the stability of BRCA1-complex [13]. The wild-type protein has been sub-cloned into bacterial expression pGEX-kT

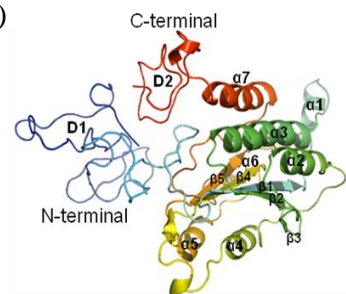


Figure 2: Model structure of MERIT40

vector, and protein was chromatography purified in large quantities for structural and biochemical studies. The full length protein is observed to have a molecular weight of 36.61 kDa, as established through mass spectrometry. Crystallization attempts have not yet yielded good quality single crystals. Ab-initio molecular modeling was performed using the Robetta server, and the molecular model suggests disordered region at both N and C-terminals (**Figure 2**), a finding consistent with results of in-vitro limited proteolysis using trypsin and chymotrypsin enzymes. Structural similarity of MERIT40 model with vWFA like domain suggests proteosome-like activity and involvement in ubiquitination processes for MERIT40 [14]. Gel filtration, native gel electrophoresis and glutaraldehyde cross-linking studies indicated the presence of a concentration independent dimeric species along with the prominent monomeric species. Thermal and chemical denaturation profiles suggest that MERIT40 undergoes unfolding through an intermediate species. To our conclusion, MERIT40 can be classified as an intrinsically disordered protein due to the presence of N-and C-termini disordered regions. Structural homologous of MERIT40 suggests its plausible role in complement activation pathway. MERIT40 undergoes a three state unfolding transition pathway with a dimeric intermediate.

Chapter 5 describes biophysical characterization of ABRAXAS wild-type and R361Q mutant.

ABRAXAS is the third component of the BRCA1-complex [6, 15-18], and is directly associated with RAP80 and BRCA1-BRCT. The *abraxas* gene was sub-cloned into bacterial expression vector pET28a(+) ,and the protein was expressed in *E.coli Rosetta*

2(DE3). Site-directed-mutagenesis was performed using PCR and the restriction enzyme method. The wild-type and mutant proteins were purified using affinity chromatography (Ni-NTA). The molecular weight of the purified protein was estimated around 41 kDa using size exclusion chromatography.

DLS experiment has shown that the mutant had a larger hydrodynamic diameter (wild type $9.0 \pm 0.3 \text{ nm}$ and mutant $10 \pm 0.5 \text{ nm}$). Comparable susceptibility of ABRAXAS mutant towards protease digestion suggests that substitution of R361Q is not destabilizing the domain integrity of ABRAXAS. CD and fluorescence analyses also reveal that the secondary structures of wild-type and mutant proteins are similar, while changes in Trp and Tyr microenvironment suggest a mild alteration in tertiary structure. Thermal stability and folding pathway of wild-type and mutant were found to be similar when probed by varying the temperature (**Figure 3**). However, chemical denaturation using urea as the unfolding agent showed a prominent intermediate in case of wild-type compared to mutant.

In conclusion, in-vitro and in-silico studies indicate that R361Q mutation is causing several localized changes which adversely affect RAP80 binding (**Figure 4**) and recruitment of BRCA1-complex to the DNA damage site.

Chapter 6 describes the molecular assembly of BRCA1-complex.

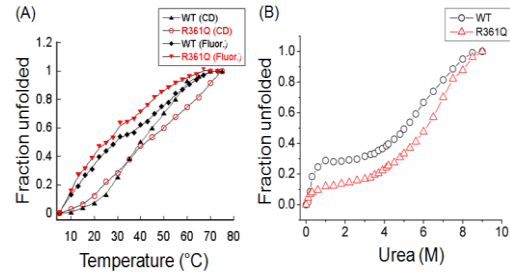


Figure 3: Thermal and urea denaturation of ABRAXAS

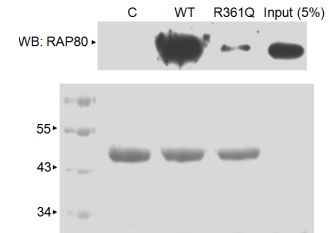


Figure 4: Pull down of ABRAXAS

Though the BRCA1-complex consists of six proteins, RAP80, ABRAXAS, BRCA1, MERIT40, BRCC36 and BRCC45, it is not known how these are organized in the complex [13, 19, 20]. In other words, it is not completely known which components interact directly or indirectly. A direct association between RAP80 and ABRAXAS has been reported earlier[7]. Similarly, pS-X-X-F motif of ABRAXAS directly binds to BRCA1-BRCT[7] in a phosphorylation dependent manner. This study focused on identification of direct binding partners of MERIT40. The ABRAXAS, RAP80 and MERIT40 were sub-cloned into bacterial vector, protein was expressed and purified to carry out interaction analysis using pull down and calorimetry

A quantitative pull down interaction analysis established direct interaction of MERIT40 with BRCA1-BRCT and ABRAXAS. The observed ITC isotherm of MERIT40 and

BRCA1-BRCT interaction depict a low binding affinity ($K=1.02 \pm 0.15 \text{ E}^5$). Interactions profile of MERIT40 and ABRAXAS suggests that ABRAXAS C-

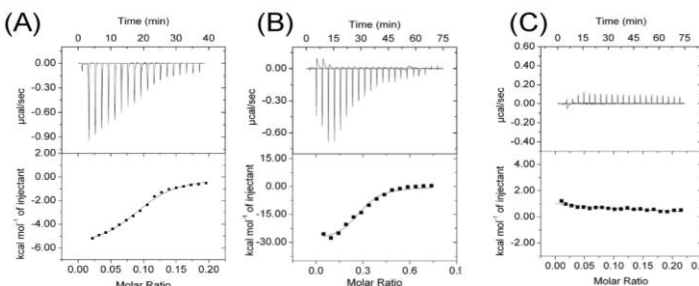


Figure 5: ITC of MERIT40 with (A) BRCA1 BRCT (B) ABRAXAS and (C) RAP80

terminal region is involved in interaction with MERIT40 with a binding affinity ($K= 5.8 \pm 0.35 \text{ E}^5 \text{M}^{-1}$). However, MERIT40 doesn't show any interaction with RAP80 (**Figure 5**). In conclusion, interactive relationship of MERIT40 with ABRAXAS as well as BRCA1-BRCT mediates the formation of stable BRCA1-complex. Thus, MERIT40 is the core mediator molecule, and is obligatory for organization of stable BRCA1-complex.

Chapter 7 of the thesis describes the crystallographic studies of BRCA1 interacting partners.

One of the basic requirements for three-dimensional structure determination of macromolecules using X-ray crystallography is the highly purified protein in large amount and well-ordered crystals. Protein crystallization experiments were carried out by the vapour diffusion method separately in hanging and sitting drops. The concentration of protein used to prepare the drop, which equilibrated with a precipitating solution, ranged from 15 – 25 mg/ml. Commercially available solutions, Hampton Crystal Screen 1, 2 , SaltX, Sigma crystal screen and Qiagen screens were used as the precipitating solutions. The protein and precipitating solutions were mixed in 1:1 volume ratio, and the mixture was allowed to equilibrate at 22°C or at 4°C with 500-1000 µl mother liquor in a closed system. We observed a clear drop or precipitation in most of the drops in cases of MERIT40, RAP80 and ABRAXAS. However, the crystallization trial of BRCA1-BRCT and KIF1 peptide complex resulted in diffraction quality crystals. These crystals were used for x-ray diffraction data collection, and the structure of the complex has been determined to low resolution.

kif1b gene encodes a motor protein that transports mitochondria and synaptic vesicle precursors, and mutations in this gene cause the formation of Charcot-Marie-Tooth disease. Kinesin-like KIF1B can be a potential interacting partner of the BRCA1-BRCT domain [21] and is suspected to be phosphorylated in case of cancer. In order to understand the molecular associations with BRCA1-BRCT and KIF1B, the synthetic construct of KIF1B (NH₂- DRTP(**pS**)PTFST-COOH) (pS- phosphoserine) peptide was analyzed for its binding with BRCA1-BRCT. The binding isotherm suggests a moderate binding with affinity constant, $K = 3.53 \times 10^5 \text{ M}^{-1}$. Crystals of the complex were obtained by co-crystallization under following conditions: 0.2 M ammonium sulphate, 30% PEG

5000 MME, 0.1 M MES pH 7.5. Diffraction data were collected at 100 K by the oscillation method using the MAR DTB mounted on a Microstar rotating anode X-ray generator (Bruker). The diffraction data processed using Mosflm software, indicates that the crystals belong to the space group $P2_12_12_1$ with unit cell parameters: $a=84.01\text{ \AA}$, $b=180.47\text{ \AA}$, $c=194.22\text{ \AA}$. The useful data extended to 3.5 \AA with a completeness of 86%. The structure was solved by molecular replacement method as implemented in the software package *Phaser*. There are six molecules in the asymmetric unit leading to a solvent content of 60%. The structure was further refined using *Refmac-5*. In the crystal structure, there is clear electron density for the KIF1 oligopeptides bound to all the six subunits. The major interactions between the KIF1 peptide and the BRCA1 BRCT domain is through hydrogen bonds between phosphoserine 909 from the peptide and residues Ser1655 and Gly1656 from the BRCT domain. Efforts are being made to improve the quality of crystals for better diffraction.

Chapter 8 describes about important findings and conclusions of the work

RAP80, ABRAXAS, MERIT40 and BRCA1 play crucial role in DNA damage repair pathway, and mutations such as $\Delta E81$ in RAP80 and R361Q in ABRAXAS cause significant impairment of BRCA1-complex function. Limited proteolysis of purified RAP80 and its $\Delta E81$ mutant shows that the mutant is less compact and more susceptible to proteolysis. Molecular modeling studies indicate that RAP80 $\Delta E81$ mutation leads to destabilization of an alpha-helix causing loss of binding with ubiquitin. This, in turn, leads to defective recruitment of BRCA1 to the DNA repair site. Similarly, biophysical and in-silico experiments indicate that R361Q mutation brings some localized changes in the structure of ABRAXAS that is likely to affect transport of BRCA1-complex from the

cytoplasm to the nucleus. MERIT40 is a recently identified protein that acts as a stabilizer of the BRCA1-complex. It is found to bridge RAP80 and BRCA1 through direct interactions with adapter protein ABRAXAS, and BRCA1, thus maintaining the integrity of the whole complex. The sequence of MERIT40 is not similar to any proteins contained in the PDB to enable homology modeling. Ab-initio modeling using Robetta server revealed that MERIT40 is structurally similar to another protein involved in complement activation. These studies presented in this thesis unravel the interactions involved among different BRCA1-complex members which are essential for DNA repair function of complex. It will further explore the possibility of structure based inhibitor design for therapeutic application that can compensate the effect of mutations.

References:

1. Alberg, A.J., A.P. Lam, and K.J. Helzlsouer, *Epidemiology, prevention, and early detection of breast cancer*. Curr Opin Oncol, 1999. **11**(6): p. 435-41.
2. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1*. Science, 1994. **266**(5182): p. 66-71.
3. Venkitaraman, A.R., *Cancer susceptibility and the functions of BRCA1 and BRCA2*. Cell, 2002. **108**(2): p. 171-82.
4. Pylkas, K., et al., *Analysis of large deletions in BRCA1, BRCA2 and PALB2 genes in Finnish breast and ovarian cancer families*. BMC Cancer, 2008. **8**: p. 146.
5. Nikkila, J., et al., *Familial breast cancer screening reveals an alteration in the RAP80 UIM domain that impairs DNA damage response function*. Oncogene, 2009. **28**(16): p. 1843-52.
6. Wang, B., et al., *Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response*. Science, 2007. **316**(5828): p. 1194-8.
7. Kim, H., J. Huang, and J. Chen, *CCDC98 is a BRCA1-BRCT domain-binding protein involved in the DNA damage response*. Nat Struct Mol Biol, 2007. **14**(8): p. 710-5.
8. Monteiro, A.N., A. August, and H. Hanafusa, *Evidence for a transcriptional activation function of BRCA1 C-terminal region*. Proc Natl Acad Sci U S A, 1996. **93**(24): p. 13595-9.
9. Yu, X., et al., *The BRCT domain is a phospho-protein binding domain*. Science, 2003. **302**(5645): p. 639-42.
10. Manke, I.A., et al., *BRCT repeats as phosphopeptide-binding modules involved in protein targeting*. Science, 2003. **302**(5645): p. 636-9.
11. Solyom, S., et al., *Breast cancer-associated Abraxas mutation disrupts nuclear localization and DNA damage response functions*. Sci Transl Med, 2012. **4**(122): p. 122ra23.
12. Huen, M.S., et al., *RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly*. Cell, 2007. **131**(5): p. 901-14.
13. Kolas, N.K., et al., *Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase*. Science, 2007. **318**(5856): p. 1637-40.
14. Mailand, N., et al., *RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins*. Cell, 2007. **131**(5): p. 887-900.
15. Wang, B. and S.J. Elledge, *Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brcal/Brcc36 complex in response to DNA damage*. Proc Natl Acad Sci U S A, 2007. **104**(52): p. 20759-63.
16. Arnold, K., et al., *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling*. Bioinformatics, 2006. **22**(2): p. 195-201.
17. Shao, G., et al., *MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks*. Genes Dev, 2009. **23**(6): p. 740-54.
18. Holm, L. and P. Rosenstrom, *Dali server: conservation mapping in 3D*. Nucleic Acids Res, 2010. **38**(Web Server issue): p. W545-9.
19. Celeste, A., et al., *Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks*. Nat Cell Biol, 2003. **5**(7): p. 675-9.
20. Sobhian, B., et al., *RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites*. Science, 2007. **316**(5828): p. 1198-202.
21. Wang, Y., et al., *BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures*. Genes Dev, 2000. **14**(8): p. 927-39.
22. Feng, L., J. Huang, and J. Chen, *MERIT40 facilitates BRCA1 localization and DNA damage repair*. Genes Dev, 2009. **23**(6): p. 719-28.
23. Wang, B., et al., *NBA1, a new player in the Brcal A complex, is required for DNA damage resistance and checkpoint control*. Genes Dev, 2009. **23**(6): p. 729-39.
24. Rodriguez, M., et al., *Phosphopeptide binding specificities of BRCA1 COOH-terminal (BRCT) domains*. J Biol Chem, 2003. **278**(52): p. 52914-8.

List of publications:

(a) Published

- 1) **Vikrant**, Ulka Sawant, Ashok K Varma, (2014). Role of MERIT40 in stabilization of BRCA1-complex: A protein-protein interaction study. Biochem Biophys Research Communication. (<http://dx.doi.org/10.1016/j.bbrc.2014.03.073>)
- 2) Dilip C. Badgujar, Ulka Sawant, **Vikrant**, Lumbini Yadav, M. V. Hosur and Ashok K. Varma, (2013). Preliminary Crystallographic Studies of BRCA1 BRCT- ABRAXAS Complex. Acta Crystallography. F69. 1401-1404
- 3) **Vikrant**, Pallavi Nakhwa, Dilip C. Badgujar, Rajan Kumar, Khushboo K.S. Rathore, Ashok K. Varma, (2013). Structural and functional characterization of the MERIT40 to understand its role in DNA repair. Journal of Biomolecular Structure and Dynamics. (<http://dx.doi.org/10.1080/07391102.2013.843473>)
- 4) **Vikrant**, Rajan Kumar, Lumbini R. Yadav, Pallavi Nakhwa, Sanjeev K. Waghmare, Peyush Goyal, Ashok K. Varma, (2013). Structural and Functional Implication of RAP80 Δ 81Glu Mutation. PloS One. 8. e72707
- 5) **Vikrant**, Ashok K Varma, (2011). Structural investigation of RAP80; A novel BRCA1 interacting protein involved in the mediation of DNA damage repair function. Acta Crystallography. A67. 346
- 6) **Vikrant**, Rajan Kumar, Quadir, S. Waghmare, Ashok K Varma. Mechanism of BRCA1-complex due to ABRAXAS Arg361Gln mutation. JBSD, 10.1080/07391102.2014.945484

(C) Other Publications:

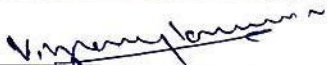
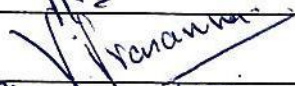
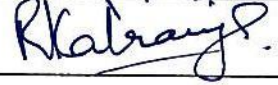
- 1) Lumbini, Mahamaya, **Vikrant**, M V Hosur, Ashok Varma. Tetrameric ZBRK1 DNA binding domain has affinity towards cognate DNA in absence of Zinc ions- **BBRC** | Volume 450 | Pages 283-288
- 2) Rakesh C. Chandarana, **Vikrant**, Ashok K. Varma, Anil Saran, Evans C. Coutinho, Jacinta S. D'Souza (2013). Over-expression, purification and isotopic labeling of a tag-less human glucose dependent insulinotropic polypeptide. 3 Biotech. 13205-013-0181
- 3) Bhanu P Jagilinki, Nikhil Gadewal, Harshal Mehta, Hafiza Mahadik, **Vikrant**, Anamika Pandey, Ulka Sawant, Prasad A Wadegaonkar, Peyush Goyal, Satish Kumar, Ashok K Varma (2014) Conserved residues at the MAPKs binding interfaces that regulate transcriptional machinery. **JBSD** | DOI: <http://dx.doi.org/10.1016/j.bbrc.2014.05.104>

Date: 16/4/2014



Signature of Student

Doctoral Committee:

S. No.	Name	Designation	Signature	Date
1	Dr. Vinay Kumar	Chairperson		17/4/2014
2	Dr. Ashok Varma	Convener		16/4/14
4	Dr. V. Prasanna	Member		16/4/14
5	Dr. Rajiv Kalraiya	Member		16/4/14

Forwarded Through:



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



Prof. K. Sharma
Director, Academics
Tata Memorial Centre

Dr. S. V. Chiplunkar
Director
Advanced Centre for Treatment, Research &
Education in Cancer (ACTREC)
Tata Memorial Centre
Kharghar, Navi Mumbai 410210.

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

List of Abbreviations

BRCA1	: BReast CAncer susceptibility gene1
CtIP	: CtBP Interacting Protein
RING	: Really Interesting New Gene
BARD1	: BRCA1 Associated Ring Domain-1
HR	: Homologous Recombination
NHEJ	: Non-Homologous End Joining
ATM	: Ataxia Talangiectasia Mutated
ATR	: ATM and RAD3 related
MDC1	: Mediator of DNA Damage Checkpoint -1
RNF	: RING finger containing protein
UBC	: Ubiquitin Conjugating Enzyme
MRE11	: Meiotic Recombination 11 homolog
MRN	: Mre11-Rad50-Nbs1 complex
RAP80	: Receptor Associated Protein-80
Chk1	: Checkpoint kinase 1
BACH1	: BTB and CNC homology 1
ACC1	: Acetyl Coenzyme A Carboxylase
CD	: Circular Dichroism
MALDI-TOF	: Matrix-Assisted Laser Desorption and Ionisation-Time Of Flight
ITC	: Isothermal Titration Calorimetry
MR	: Molecular Replacement

COOT	: Crystallographic Object-Oriented Toolkit
NCoA2	: Nuclear receptor CoActivator 2
TEV	: Tobacco Etch Virus protease
CCP4	: Collaborative Computational Project No.4
FPLC	: Fast Protein Liquid Chromatography
NMR	: Nuclear Magnetic Resonance
RMSD	: Root Mean Square Deviation
Top BP1	: Topoisomerase (DNA) II Binding Protein 1
53BP1	: p53 Binding Protein
PARP1	: Poly (ADP-ribose) Polymerase 1
RFC	: Replication Factor C
ARD	: Ankyrin Repeat Domain
SPR	: Surface Plasmon Resonance

List of Tables

Table 4.1	Protein geometry evaluation using Mol Probity web server
Table 4.2	Structural comparison of MERIT40 with structurally similar protein using DALI
Table 6.1	Mass spectrometry profiles of MERIT40, BRCA1, ABRAXAS and RAP80 using Mascot analysis
Table 7.1	ITC reaction parameters
Table 7.2	Summary of data collection, processing and refinement statistics
Table 7.3	Binding comparison of various BRCA1 BRCT binding peptides

List of Figures

Figure 1.1	BRCA1 structural organization and its cellular binding proteins
Figure 1.2	Model depicts the BRCA1 mediated checkpoint regulation
Figure 1.3	Mechanism of BRCA1 localization and nuclear import/Export
Figure 1.4	BRCA1-mediated transcription regulation in response to DNA damage
Figure 1.5	Transcription regulatory functions of BRCA1
Figure 1.6	Mechanism of BRCA1 regulation of responses to microtubule damage.
Figure 1.7	BRCA1 complex and its recruitment at the site of DNA damage in response to IR
Figure 1.8	Functional domains of BARD1 and its associated proteins
Figure 1.9	E2 Catalysis mechanism for addition of ubiquitin
Figure 1.10	Various type of Lys residue specific linkage during (poly) ubiquitination and their fate in various cellular events
Figure 1.11	Importance of E2 in ubiquitylation
Figure 2.1	The phase diagram
Figure 3.1	DNA damage response in repair pathway
Figure 3.2	Graphical representation of RAP80
Figure 3.3	Basic scheme of sub-cloning of different functional domain of RAP80 in pGEX-kT
Figure 3.4	Agarose gel showing the insert release from potential clones plasmids
Figure 3.5	Expression and purification profile of RAP80 (1-130)
Figure 3.6	Purification of RAP80 (1-405)
Figure 3.7	Binding interaction of RAP80 UIMs and Δ E81 with Di-Ub (K-63 linked)
Figure 3.8	Resistivity profile of RAP80 wild-type and Δ E81 towards protease digestion
Figure 3.9	Comparison of secondary structural components of RAP80 (1-130)

Figure 3.10	Thermal stability of RAP80 (1-130)
Figure 3.11	Differential Scanning Calorimetry profile of RAP80 wild-type
Figure 3.12	Binding analysis of RAP80 wild-type and Δ E81 with Di-Ub (K-63 linked)
Figure 3.13	Mechanism of consequence due to RAP80 Δ E81
Figure 4.1	Graphical representation of structure of MERIT40
Figure 4.2	Primer used during PCR amplification of MERIT40
Figure 4.3	Agarose gel showing the PCR amplified product
Figure 4.4	Expression and purification of MERIT40
Figure 4.5	Modeled structure of MERIT40 and its validation
Figure 4.6	Limited proteolysis followed by mass spectrometry of MERIT40
Figure 4.7	Identity determination of MERIT40
Figure 4.8	Multiple structural alignment of MERIT40 model with DALI searched templates
Figure 4.10	Oligomeric characterization of MERIT40
Figure 4.11	MALDI-TOF mass spectrometry profile of full length MERIT40
Figure 4.12	Dimeric structure interface of MERIT40
Figure 4.13	Far-UV CD spectra, and urea unfolding using fluorescence, of MERIT40
Figure 4.14	Denaturation profile of MERIT40 using CD, DSC and fluorescence spectroscopy
Figure 4.15	Mutational analysis of MERIT40. (A) ligplot of Lys274. (B) ligplot of Arg274
Figure 5.1	Graphical representation of ABRAXAS
Figure 5.2	Basic scheme of sub-cloning of different functional domain of ABRAXAS in pET28
Figure 5.3	Multiple sequence alignment of ABRAXAS showing highly conserved nature of Arg361
Figure 5.4	Agarose gel showing the PCR amplified product
Figure 5.5	Purification of ABRAXAS. SDS-PAGE showing purified protein

Figure 5.6	Dynamic Light Scattering profile of ABRAXAS (6-373) wild-type and R361Q
Figure 5.7	Resistivity profile of ABRAXAS wild-type and mutant towards Protease digestion
Figure 5.8	Secondary and tertiary structure evaluation of ABRAXAS wild-type and mutant
Figure 5.9	Thermal and chemical stability of ABRAXAS (6-373) wild-type and mutant
Figure 5.10	Binding analysis of ABRAXAS wild-type and mutant with RAP80
Figure 5.11	The figure illustrates the plausible mechanism of defective homologous recombination repair and other consequences due to ABRAXAS R361Q mutation
Figure 6.1	Schematic representation of members of BRCA1 complex
Figure 6.2	Protein purification and structural characterization of ABRAXAS, BRCA1-BRCT, MERIT40 and RAP80
Figure 6.3	Pull-down assay to determined direct binding partner(s) of MERIT40
Figure 6.4	ITC thermogram of interaction analysis between MERIT40 and BRCA1-BRCT
Figure 6.5	Interaction analysis of MERIT40 with ABRAXAS using ITC
Figure 6.6	Isothermal titration calorimetry of MERIT40
Figure 6.7	Isothermal titration calorimetry of MERIT40 with RAP80 (1-405)
Figure 6.8	Anticipated model of MERIT40 mediated BRCA1-complex stabilization
Figure 7.1	Schematic representation of the structure of KIF1b
Figure 7.2	Schematic representation of Kif1b peptide used in the binding and structure determination study
Figure 7.3	SDS-PAGE showing the purified protein BRCA1 BRCT
Figure 7.4	Binding isotherm of BRCA1 BRCT-KIF1b
Figure 7.5	Structure solution of BRCA1 BRCT-KIF1b complex

Figure 7.6	Peptide showing electron density map
Figure 7.7	Structural evaluation of BRCA1 BRCT-KIF1b using Ramachandaran plot
Figure 7.8	Schematics representation of secondary structural component of BRCA1 BRCT-KIF1b
Figure 7.9	Structural superposition of different chains of BRCA1-BRCT in asymmetric unit, without peptides
Figure 7.10	Structural overlay of BRCA1-BRCT of the various reported complex with solve structure of BRCA1-BRCT, without peptides

Chapter 1

Introduction and Review of Literature

1.1 Cancer

Cancer is a deadly disease which arises due to abnormal cell division in the body. Instead of undergoing death, cancerous cells continue to grow and form new abnormal cells, which may undergo metastasis. Breast cancer is a malignant tumor that can invade the surrounding tissues or metastasize to other organ(s). Breast cancer can be classified into four categories, ductal carcinoma *in situ* (DCIS; also known as *intraductal carcinoma*), lobular carcinoma *in situ*, invasive (infiltrating) ductal carcinoma, and invasive (infiltrating) lobular carcinoma [22].

DCIS is the most common type of non-invasive breast cancer in which cancer cells remain in the milk ducts and do not spread through the walls into the surrounding breast tissues. Lobular carcinoma *in situ* is not considered as a true cancer. Invasive (or infiltrating) ductal carcinoma is the most common type of invasive breast cancer. It starts from a milk duct of the breast tissue and migrates through the walls of the duct, and further grows into the fatty tissue and metastasizes to other organs. Invasive lobular carcinoma (ILC) originates in the milk-producing glands (lobules) and undergoes metastasis.

It has been observed that ~ 5- 10 % of breast-cancer cases are hereditary in nature. The most common cause of the hereditary breast cancer is an inherited genetic alteration in the tumor suppressor genes *BRCA1* and *BRCA2* [23]. Both these genes appear to be equally responsible in the early onset of breast cancer [24-26]. The reported risk for breast and ovarian cancer is as high as 80% for members of families with *BRCA* mutations [1, 2]. *BRCA1* mutation(s) can be considered as autosomal dominant with high penetrance. However, such mutations behave recessively as per the concept of cancer

etiology [27]. The wild- type copy of *brca1* gene is found to be lost in breast and ovarian cancer patients having chromosomal aberration [28, 29]. All these observations support the hypothesis that *BRCA1* acts as a tumor suppressor gene.

1.2 DNA damage and genome conservation

Naturally, DNA damages happen about 10,000 to 100,000 times per day per mammalian cell [30]. DNA damage manifests as alterations in the molecular structure of DNA, such as a strand breakage or nucleotide base change. Several factors are responsible for DNA damage, including metabolic processes of the body and environmental radiations. The metabolic products such as reactive oxygen, nitrogen and carbonyl species, and lipid peroxidation products or alkylating agents can damage the DNA [31].

DNA damage can be repaired by various processes, out of which homologous recombination (HR) and non homologous end joining (NHEJ) assign a substantial role for *BRCA1* [32]. Homologous recombination utilizes the coding information present on sister chromatids for accurate repair. However, NHEJ is an error-prone process because it does not consider sister chromatids to rejoin broken ends. Still, NHEJ is the major pathway of double strand break (DSB) repair in mammalian cells and happens throughout the cell- cycle. *BRCA1* favours HR over NHEJ, and contributes in the NHEJ by assisting alignment of short regions of homology on either side of the DNA strand break. This process of damage repair is called as microhomology NHEJ [33, 34]. Cells prefer to take NHEJ pathway to repair their damaged DNA in the absence of *BRCA1* gene, thus making the resultant repair product highly error-prone.

The task of conservation of genome is achieved by the recital action of cellular events, such as DNA replication, DNA repair, apoptosis and senescence. Most of these

processes are coordinated by the DNA-damage checkpoint, which is a complex signaling network triggered by genomic instability or DNA damage [35]. Two phospho-inositide 3-kinase-like protein kinases (PIKKs), ataxia telangiectasia mutated (ATM), and ATM- and Rad3-related (ATR), are the prime sensors of DNA damage, and act as master regulators of cell-cycle checkpoint [35]. ATM primarily responds to DNA double-strand breaks (DSBs), whereas ATR becomes activated to a much broader range of DNA damage, including DSBs [36]. The initial step in ATR activation is the recognition of single stranded DNA (ssDNA), and junctions between ssDNA and double-stranded DNA (dsDNA), which are induced by DNA damage. Although, the ATR activation at DSBs is highly dependent on ATM, it can also respond directly at interference of DNA replication [37, 38]. When ATM and ATR are activated, DSBs are recognized by MRN complex, which consists of meiotic recombination protein 11 (MRE11), RAD50 and Nijmegen breakage syndrome protein 1 (NBS1) [39]. The MRN11-RAD50-NBS1 (MRN) complex is crucial for the recruitment and stimulation of ATM at DSBs [40-42] [39]. ATM remains active throughout the cell-cycle, whereas, ATR acts often slowly and predominantly only during the S and G2 phases of the cell cycle. ATM generates long stretches of ssDNA adjacent to the breaks due to resection of DSBs by exo- and endonucleases [37]. The exo- and endonucleases involved in DSB resection include the EXO1 (an exonuclease), MRN complex (an exo- and endonuclease) and CtIP (an endonuclease and an activator of the exonuclease activity of MRN) [43] [44]. Proteins at ATR signaling cascade, including RAD17, RPA, ATRIP, 9-1-1 complex and TopBP1, get phosphorylated by ATR during checkpoint activation. Phosphorylation of ATR

substrates has a clear role in downstream processes such as cell-cycle arrest and DNA repair.

The well-known substrate of ATR is the checkpoint kinase 1 (Chk1); an effector kinase involved in the ATR mediated checkpoint pathway [45]. Chk1 is not only crucial for the stability of DNA replication forks but is also important for the checkpoint response during S phase. In response to DNA damage, the Chk1 is phosphorylated by ATR at multiple sites, which stimulates the kinase activity and releases Chk1 from chromatin. Furthermore, Chk1 phosphorylates substrates such as Cdc25A and Cdc25C which are involved in cell-cycle arrest at the G1/S and G2/M transitions, respectively. The phosphorylation of Chk1 by ATR is mediated by claspin which is a component of the DNA replication fork complex [46]. Claspin and Chk1 stabilize each other even in the absence of DNA damage [47, 48]. Claspin is phosphorylated by ATR in response to DNA damage, and endorses the phosphorylation and further activation of Chk1 [49, 50]. Claspin is degraded in the G2 and M phases of the cell cycle by the Skp1–Cullin1–F-box (SCF) ubiquitin ligase [51, 52]. It is degraded in G1 by the anaphase-promoting complex (APC) while under DNA damage condition, the degradation of claspin in G2 is disrupted [53]. In response to DNA damage, claspin can accumulate itself during the S and G2 phases to ensure efficient Chk1 activation during cell-cycle. The claspin-mediated phosphorylation of Chk1 describes a good example of phosphorylation regulation of ATR substrates at multiple levels.

BRCA1 is a substrate of ATM, ATR and CHK2 protein kinases, and plays an essential role in S/G2-M checkpoint control [37]. BRCA1 has RING domain at its N-terminal, DNA binding region in the centre and two BRCT repeats at the C-terminal (**Figure 1.1**).

Central domain(s) of BRCA1 exhibits DNA-binding activity; the region from (452-1079) amino acids bind to branched or cruciform DNA, whereas a region from (1293–1560) amino acids independently acts as a transcription regulator [18, 54]. The BRCT domain has been found in different proteins, including RAD9, AD4, 53BP1, *Crb2*, RAP1, BARD1 and MDC1. It possesses phosphopeptide-binding ability through which it functions in cell -cycle regulation, DNA-damage checkpoint control and DNA repair [10, 21]. A superficial pocket at the N-terminal of BRCT motif interacts with the phosphoserine (pSer) at (0) position of the ligand, whereas a hydrophobic pocket at the interface of the two BRCT repeats accommodates (+3) position phenylalanine residue. The three dimensional structural folding at the N- and C-terminal of the BRCT pocket is similar in the all phosphopeptides-BRCTs complexes, consistent with its unbiased binding to pSer (0) [55, 56]. However, the linker region at the interface is significantly different, which determines BRCTs binding specificity.

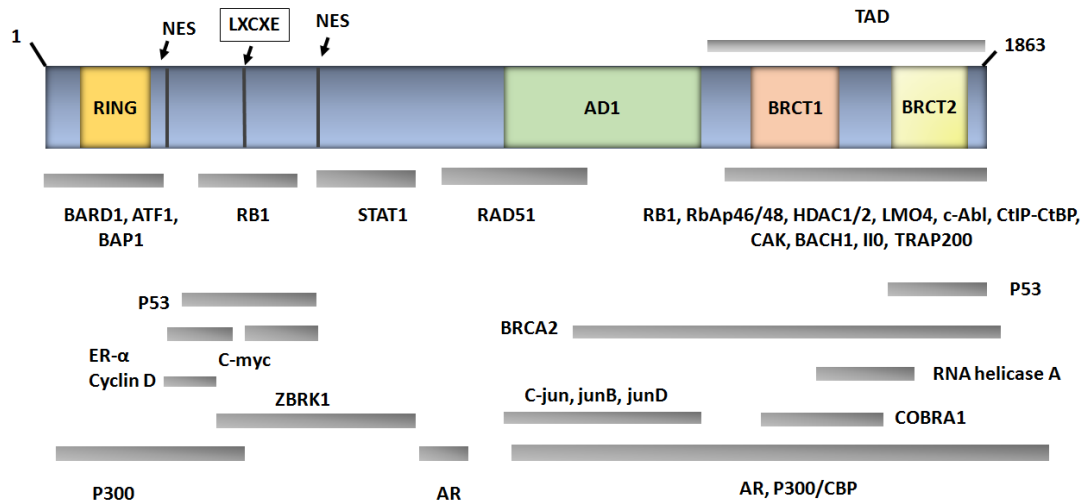


Figure 1.1: BRCA1 domain organization and its cellular binding proteins.

ATM/ATR kinases hyper-phosphorylate BRCA1 at different sites, and this leads to dissociation of BRCA1 from processive RNAPII complexes [57]. Phosphorylation of BRCA1 occurs at Ser 1387, Ser 1423 and Ser 1524 positions in the central domain. Mutation on BRCA1 Ser 1387 disrupts the S-phase checkpoint, whereas mutations at Ser1423 and Ser1524 block BRCA1-mediated cell cycle G2/M checkpoints [58]. Furthermore, BRCA1 phosphorylation by CHK2 at Ser 988 position is essential for HR activity [59], G2/M cell-cycle checkpoint activation and also for suppression of genotoxicity induced mammary and uterine cancers [60].

BRCA1 is a part of BASC (BRCA1-Associated genome Surveillance Complex) complex, which when phosphorylated by ATM/ATR kinases generate binding sites for BRCA1 associated factors, such as MRE11, RAD50 and NBS1 (MRN complex) [18]. MRN complex intervenes the G2/M cell-cycle checkpoint signal, and also plays a role in DSB repair by holding two DNA ends close to each other, which is required for non-homologous end-joining function (**Figure 1.2**) [61]. BRCA1 specifically interacts with processive polymerase II (IIO) and acts as a part of genome surveillance complex [57] [62].

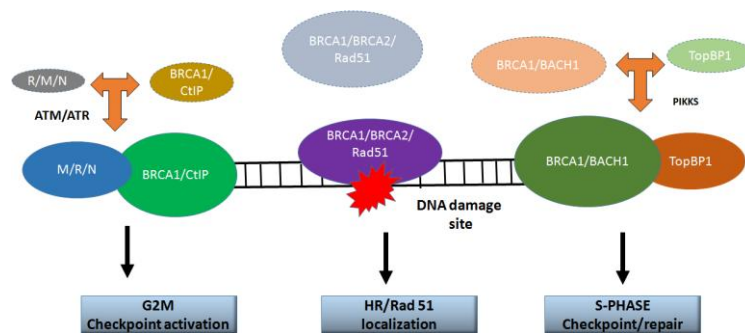


Figure 1.2: Model depicts the BRCA1 mediated checkpoint regulation. BRCA1-containing complexes interact efficiently after DNA damage with TopBP1 and the M/R/N complex. Upon DNA damage, specific kinase signaling cascade promotes BRCA1 super-complex assembly (colored circles).

BRCA1 BRCT interacts with the C-terminal of p53 and stabilizes the complex. This stable complex is essential for stimulation of transcriptional activity, as demonstrated by an increase in p53-mediated activation of responsive promoters such as p21 waf1/ cip1 and bax. Interestingly, stabilization of p53 induces a subset of p53-regulated genes involved in cell cycle arrest and DNA repair but not in apoptosis [63, 64]. The DNA damaging agents were found to stabilize p53 and promote transcription of the number of genes, including apoptosis mediators. This demonstrates that BRCA1 is involved in p53-mediated growth arrest rather than programmed cell death. The p53 acts to down-regulate BRCA1 level in a negative feedback loop thereby controlling transcription activation and cell cycle [65].

BRCA1 is a nuclear protein which contains two SV40-like nuclear localization signals (NLS; 503-KRKRRP-508 and 606-PKKNRLRRKS-615) that facilitate its nuclear import [66, 67]. Translocation of BRCA1 to the nucleus is dependent on its association with the RING structure of BARD1. Binding of BARD1 to BRCA1 masks the nuclear export signal of BRCA1, thereby retaining it into the nucleus. BRCA1 and BARD1 heterodimeric complex are also required for localization of each other to DNA damaged induced foci [68] (**Figure 1.3**). Ubiquitylation is a regulatory mechanism of such a shuttling of proteins from nucleus to cytoplasm and vice versa, e.g. the nuclear export of p53 [69]. Ubiquitylation of BARD1-BRCA1 complex sends a signal for nuclear export rather than for degradation [70]. BRCA1-BARD1 complex also has the unique potential to catalyze formation of Lys6-linked polyubiquitin chain, which transduces signals other than those for protein degradation [71].

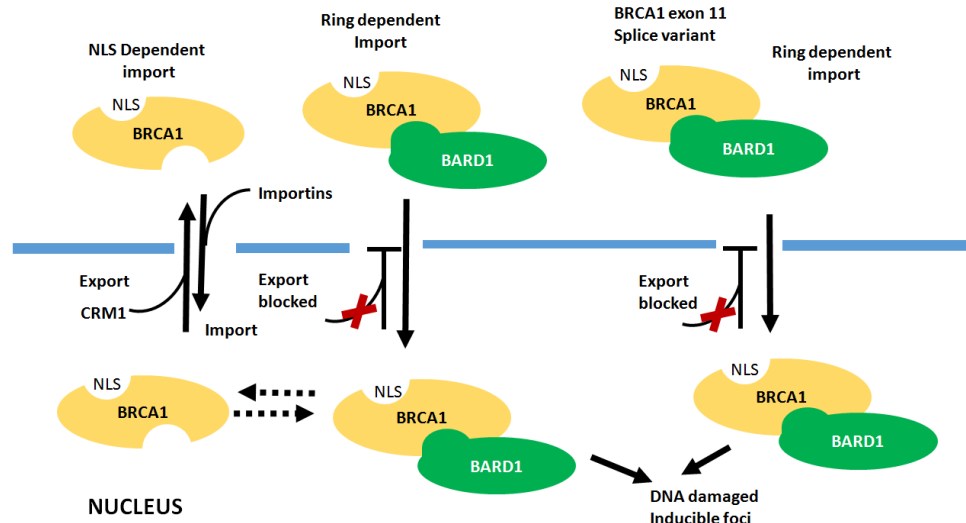


Figure 1.3: Mechanism of BRCA1 Nuclear Import/Export localization. BRCA1 can enter the nucleus with two different mechanisms: 1) via the importin receptor pathway mediated by the two NLSs, and 2) through its association with BARD1 RING domain.

BRCA1 regulates a large number of chromatin remodeling activities through its activation domain 1 (AD1) and the two C-terminal BRCA1 repeats (BRCTs). These remodeling activities do not require histone acetylation and intercede through the BRCA1-dependent recruitment of a cofactor of BRCA1 (COBRA1) [72]. The COBRA1 complex binds and down regulate the transcription activity of estrogen receptor-alpha (ER- α), thus contributing to BRCA1-mediated repression of ER- α activity [73].

ZBRK1 (zinc finger and KRAB domain protein) represses the Gadd45a gene transcription through interaction with a specific DNA sequence[74]. This repression function gets inactivated in response to DNA damage due to ubiquitin-proteasome dependent degradation of ZBRK.. A BRCA1-dependent C-terminal repression domain is present in ZBRK1 that mediates the oligomerization of ZBRK1, leading BRCA1 binding to DNA (**Figure 1.4, Figure 1.5**) [75, 76]. BRCA1 central domain interacts with ZBRK1 and unmask BRCA1-mediated transcriptional activation of Gadd45a and other DNA

damage inducible genes [77]. BRCA1 over-expression favors the BRCA1-dependent promoter activation of DNA damage players over the ZBRK1-mediated repression of downstream proteins [74]. BRCA1 interacts and regulates transcriptional and transforming activity of c-Myc, an oncoprotein found to be over-expressed in many cancer types. BRCA1 inhibits c-Myc mediated transactivation of telomerase reverse transcriptase (TERT) and telomerase enzymatic activity [78] [79].

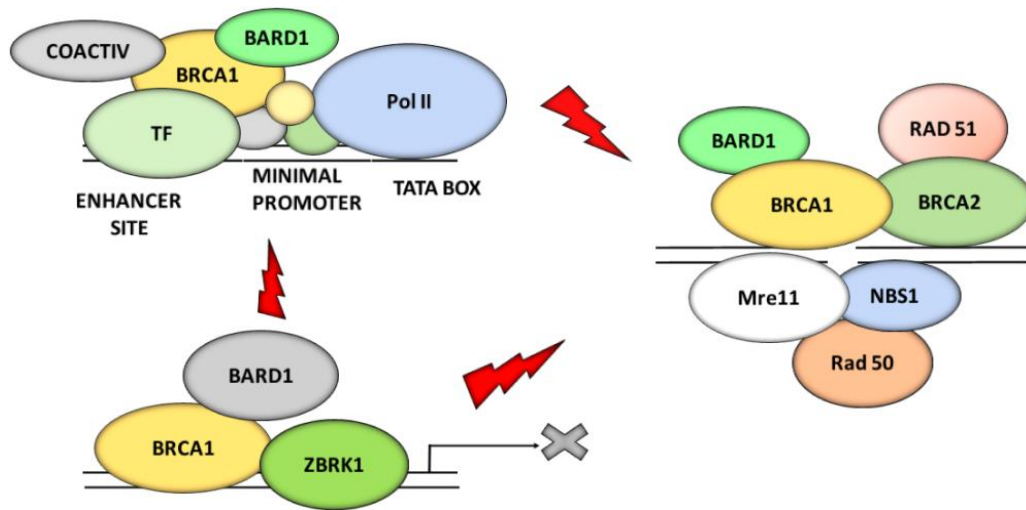


Figure 1.4: BRCA1-mediated transcription regulation in response to DNA damage.

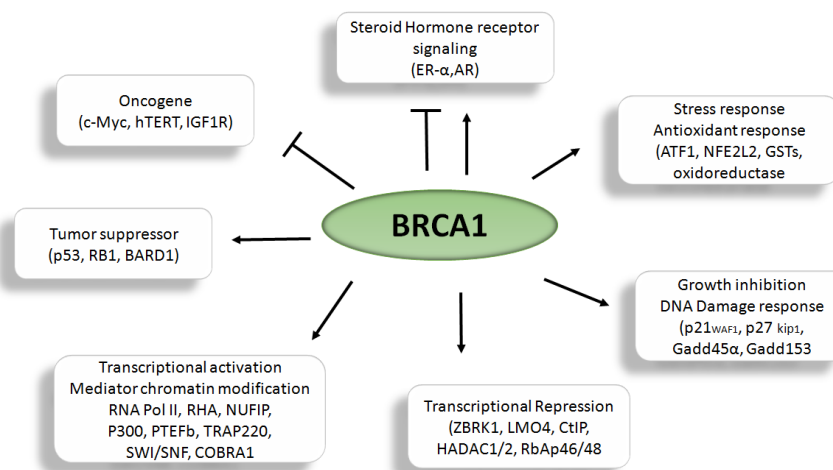


Figure 1.5: Transcription regulatory functions of BRCA1.

On DNA damage, cell sequesters the Cdc25C protein in the cytoplasm to prevent the activation of cyclinB–cdc2 kinase complex. Cdc25C is one of the targets of 14-3-3 σ chaperone protein, and is transcriptionally regulated by BRCA1 [80]. Similar to 14-3-3 σ , BRCA1 depleted cells could not elicit G2/M arrest in response to ionizing radiation (IR) [81]. BRCA1 also inhibits phosphorylation of cdc2 by upregulating the transcription of the wee-1 kinase leading to subsequent inhibition of cyclinB–cdc2 kinase [80].

BRCA1-BARD1 complex targets γ -tubulin for ubiquitination thereby regulating centrosome during cell division [55, 82]. Absence of ubiquitination due to mutation in γ -tubulin at the residues Lys-48 or Lys-344 leads to centrosome hyperactivity. BRCA1 restrains the association of γ -tubulin with centrosomes and inhibits microtubule nucleation function, thus making centrosomes quiescent [83]. BRCA1 prevents the hypertrophy and aneuploidy by acting as a member of the surveillance complex to maintain centrosome fidelity. In this case, BRCA1 acts as a sensor of microtubule disorganization caused by drugs, and promotes apoptosis through JNK activation (**Figure 1.6**). BRCA1 interacts with cdc20/Anaphase Promoting Complex which controls the transcription of MAD2 and Spindle Assembly Checkpoint [84].

1.3 DNA damage response (DDR) and BRCA1 Associated proteins

The moment DSB takes place, phosphorylation on the histone H2A variant, H2AX, occurs which results in its accumulation within chromatin at the sites of DSBs to form γ -H2AX foci [85, 86]. ATM and ATR mediated phosphorylation of histone H2AX at S139 position promotes the direct recruitment of MDC1 through MDC1-BRCT domains. This leads to the recruitment of a ubiquitin ligase RNF8/UBC13 to the damage sites [17, 87].

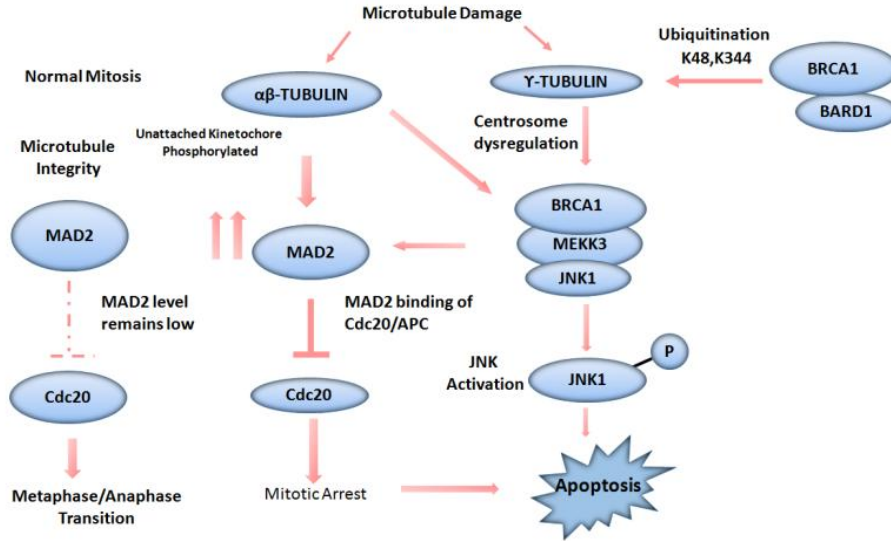


Figure 1.6: Mechanism of BRCA1 regulation of responses to microtubule damage.

MDC1 then further interacts with γ -H2AX and recruits additional activated ATM to the sites of DSBs [88], resulting in polyubiquitin chain formation at DSBs. This positive feedback loop promotes further modification of H2AX and allows expansion of regions surrounding DSBs, thus providing binding sites for DNA damage and repair factors, including the MRN complex and BRCA1. The polyubiquitin chains are subsequently recognized by RAP80 in the BRCA1 complex, thereby facilitating the formation of BRCA1 foci at DSB [35]. BRCA1 further recruits downstream proteins and promotes homologous recombination repair.

BRCA1 interacts with a variety of molecules through its various functional domains and is potentially responsible for multiple functions in cell-cycle regulation, DNA damage repair and transcription [89]. For example, BRCA1 interacts with cellular proteins via their pSXXF motifs and forms three distinct complexes-BRCA1 A, BRCA1 B and BRCA1 C [90]. The unique proteins present in BRCA1 A, BRCA1 B and BRCA1 C complexes are ABRAXAS, BACH1 and CTIP, respectively. However, the binding motif

of each molecule occupies similar conformation at the phosphopeptide binding site of BRCA1 BRCT [16] [91]. Interestingly, it has been reported that doubly phosphorylated Abraxas binds to BRCA1 BRCT in a mutually exclusive manner of BACH1 and CTIP [55]. BRCA1-BRCA2-containing complex (BRCC36) is a member of the four-subunit of BRISC (BRCC36-isopeptidase complex) that contains ABRAXAS paralog, KIAA0157/Abro1 [92]. The BRISC complex contains three additional protein components RAP80, BRCC45 and MERIT40. In addition, the BRCA1-complex contains a five-member stoichiometric complex consisting of RAP80, ABRAXAS, BRCC36, BRCC45 and MERIT40 (**Figure 1.7**) [13].

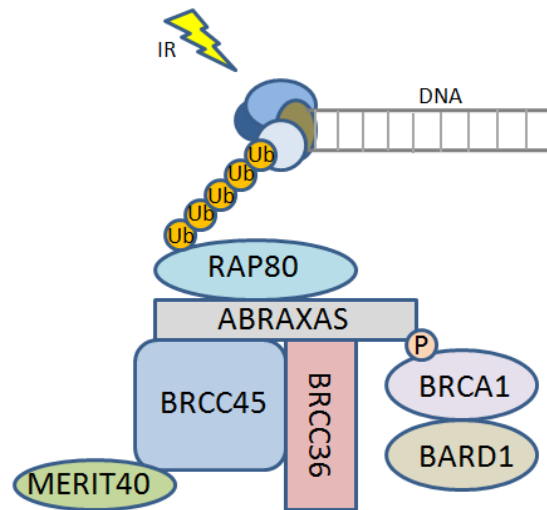


Figure 1.7: BRCA1 complex and its recruitment at the site of DNA damage in response to IR.

The RING domain at the N-terminal of BRCA1 binds to the RING domain of BARD1, and together they form a heterodimeric E3 ubiquitin ligase complex. The C-terminal of BRCA1 binds to RAP80, ABRAXAS, MERIT40 and other DNA damage response proteins to accomplish an effective damage repair process at the site of DNA damage. Mutations which delete or disrupt the C-terminal BRCT domain have been shown to cause significant relocalization of the BRCA1 from nucleus to cytoplasm [93]. Most of

the cancer-causing missense mutations which destabilize the structure of BRCTs have been identified at the interface between the two BRCT repeats of BRCA1 [94].

1.3.1 BRCA1-Associated RING Domain (BARD1): The BRCA1-associated RING domain protein 1 (BARD1) is the obligatory binding partner of the tumor suppressor BRCA1 and plays a central role in the regulation of its cellular localization, stability and function [95, 96]. It has been reported that Bard1-null mice have shown a characteristic phenotype which is identical to *Brcal*-null animals, and cellular levels of BARD1 and BRCA1 are reciprocally regulated. Furthermore, BARD1 and BRCA1 are always found in close association during various processes such as DNA damage repair [97]. BARD1 is a putative tumor suppressor, and is found to be mutated in a subset of breast and ovarian cancer patients [98]. Human BARD1 comprises different functional domains, including RING finger (46-90 residues) at the N-terminal, three ankyrin repeats (420-525 residues) at the central domain, and two tandem BRCT domains (568-777 residues) at the C-terminal (**Figure 1.7**). The RING domain of BARD1 belongs to the class of zinc finger domains having ubiquitin ligase activity[99, 100]. BARD1 exists as a heterodimer with BRCA1, which is important for cell viability, and their absence results in accumulation of chromosomal aberration and early embryonic death (**Figure 1.7, 1.8**) [97]. BRCA1-BARD1 heterodimeric complex is stabilized by four-helix bundle of α -helices found at the edge of RING domains of both polypeptides [101]. Ankyrin repeats comprise of a helix-turn-helix and a *beta*-hair-pin which together staking in multiple numbers from loops that protrude from one face of the structure to constitute the protein interaction domain.

BARD1 translocates between nucleus and cytoplasm; however, an optional localization occurs during apoptosis. BRCA1/BARD1 complex can undergo auto ubiquitination, as well as catalyze Lys6-linked polyubiquitin chain formations, which send other than degradation signal to its substrate [71]. However, BRCA1-BARD1 ubiquitinate RNAP II which targets it for proteasome-intervene degradation, and further inhibition of transcription and RNA processing in response to DNA damage [102].

BARD1 coordinates the proapoptotic stress signals and p53- dependent apoptosis, and its depletion annihilates p53 upregulation and apoptotic response to stress, further leads to genomic instability and a premalignant tumor [103].

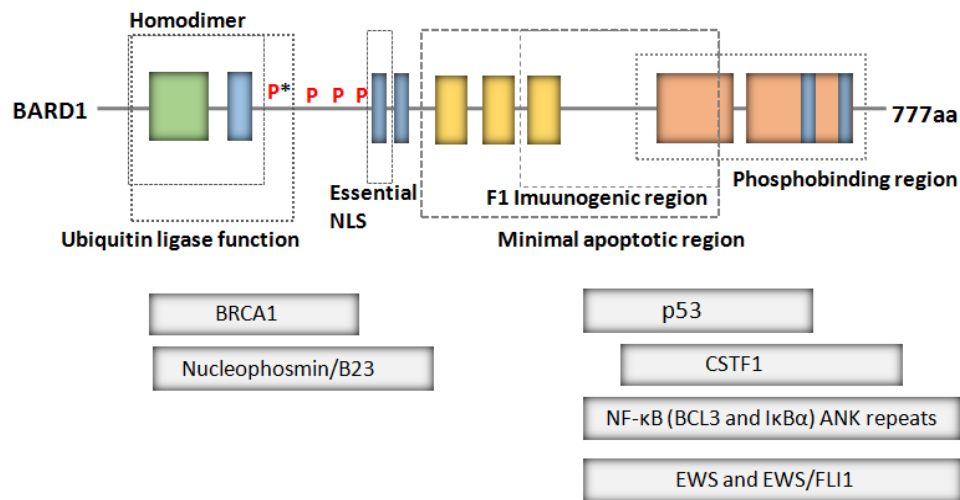


Figure 1.8: Functional domains of BARD1 and its associated binding proteins.

BARD1 performs few cellular functions without *BRCA1* coordination. In the testis, *BARD1* was found in premeiotic cells to regulate apoptosis in the absence of *BRCA1* [104]. *BARD1* and *BRCA1* expressions have been differentially regulated in an ovulatory cycle in case of breast, ovarian and uterine tissue [95]. *BARD1* expression was found to be upregulated from dioestrus to postoeustrus, whereas reverse expressions were found in *BRCA1* from oestrus to postoeustrus [105]. *BARD1* expression is also associated with

apoptosis as seen in hypoxia induced by artery blockage, which suggests differential expression of BARD1 as compared to BRCA1 in the tissues undergoing apoptosis. Furthermore, overexpression of exogenous *BARD1* leads to apoptosis that is associated with p53 stabilization and activation of caspase 3. BARD1 performs p53 dependent apoptosis function, which is inhibited by BRCA1 in p53 or BRCA1-deficient cell lines [103].

1.3.2 BRCA1/BRCA2-Containing Complex 36 (BRCC36): BRCC36 is a member of BRCC multiprotein complex, and is found to be associated with BRCA1 and BRCA2 (**Figure 1.7**). It plays an important role in the regulation of the ubiquitin E3 ligase activity of BRCC complex. The BRCC36 gene is located at the Xq28 locus, a chromosomal break point in patients with prolymphocytic T cell leukemia. BRCC is considered as E3 ligase complex and performs E2 mediated ubiquitination of the tumor suppressor p53 [106].

BRCC36 shares sequence homology with the human Poh1/Pad1 and Jab1 which are the subunits of 26S proteasome and COP9 signalosome [107]. Cancer-associated mutations of BRCA1 disrupt the interactions with BRCC36 [107]. Moreover, the four-subunit of BRCC complex (BRCA1/BARD1/ BRCC45/BRCC36) has enhanced E3 ligase activity compared to the two subunit BRCA1/BARD1 complex. Furthermore, anomalous expression of BRCC36 was observed in the majority of breast cancer cell lines and invasive ductal carcinomas [107].

1.3.3 Receptor Associated Protein 80 (RAP80): RAP80 was originally identified as a protein associated with retinoid receptor (RTR) which plays an important role in embryonic development as well as in the adult [108]. RTR expresses in embryonic stem

cells and is regulated during differentiation of embryonal carcinoma and embryonic stem cells [109]. RAP80 mRNA is expressed most abundantly in testis, and encodes a 79.6-kDa (719 amino acids) nuclear protein [108]. RAP80 comprises different functional domains, including Cys-X2-Cys-X11-His-X3-Cys zinc fingers near the C-terminal and UIMs (ubiquitin-interacting motif) at its N-terminal. UIMs are typically found in proteins with roles in (de)ubiquitination, endocytosis, transcription, replication and DNA repair, and are also known to bind ubiquitin [110]. The UIMs bind to K63-linked ubiquitin chain of H2AX, thereby recruiting the BRCA1 complex to the DNA damage site (**Figure 1.7**) [16, 111]. RAP80 acts upstream of BRCA1 and ABRAXAS and regulates localization and function(s) of BRCA1 after DNA Damage [112]. RAP80 deficient mice were found cancer prone and showed hypersensitivity to ionizing radiation [113].

RAP80 is found to be involved in DSB repair and cell cycle checkpoint control in response to IR exposure [111]. The reported crystal structure of the mouse RAP80-UIM1-UIM2 complex with K63-linked di-ubiquitin (PDB ID: 3A1Q) has a continuous 60Å long α -helical structure for the two UIMs. Individually, the UIMs are weak binders, but together they generate higher binding affinity through an avidity based mechanism. The linker region that joins the two UIMs determines the selectivity for the K63-linked chains [114]. The inter-UIMs region forms a 12Å -long α -helix and is an important determinant of specific binding with Lys 63-linked di ubiquitin. The binding affinity between UIMs and mono-ubiquitin varies from (Kd) ~0.1–2mM. However, the observed strong interactions between UIMs and Lys 63-linked polyubiquitin substrates is probably accomplished by co-operative binding among multiple UIMs. The binding of first UIM to one ubiquitin makes possible for the second UIM to adopt a conformation favorable for

interactions with a nearby ubiquitin in the chain through an avidity-based mechanism [114].

1.3.4 Coiled-Coil Domain Containing Protein 98 (CCDC98): CCDC98 is a 45 kDa nuclear protein which plays an important role in DNA damage repair mechanism. It is also called as ABRAXAS because of 39 % sequence identity with a protein KIAA0157. Recent finding has also predicted that ABRAXAS directly binds to BRCA1 BRCT repeats through its pSer-X-X-Phe motif to the mutual exclusion of earlier reported binding partners like BACH1 (BRCA1-associated C-terminal helicase) and CtIP(CtBP-interacting protein) [115]. ABRAXAS acts downstream to RAP80 but upstream of BRCA1 in the DNA damage response pathway [7].

ABRAXAS acts as a bridging molecule that mediates the interactions for both RAP80 and BRCC36 with BRCA1 (**Figure 1.7**). ABRAXAS binds to RAP80 through the N-terminal domain and to BRCC36 through its coiled-coil domain. Downregulation of ABRAXAS or RAP80 in cells displayed hypersensitivity to UV and IR (Ionization radiation), G2-M cell cycle checkpoint defects and defective homologous recombination (HR) repair [7].

1.3.5 Mediator of RAP80 Interaction and Targetting 40 (MERIT40): MERIT40 (C19orf62) is the recently discovered member of BRCA1- complex [19]. The putative *C19orf62* gene product consists of 329 residues (36.5 kD), although it migrates as a 40 kDa protein on SDS-PAGE. So far, no known functional domains of MERIT40 are reported. However, the central region (93–320) is highly conserved from sea anemone and Arabidopsis to human, and is surrounded by acidic residues [19]. MERIT40 has associations with RAP80, ABRAXAS, BRCC45, BRCC36 and BRCA1 as assessed by

coomassie staining of the purified complex [13]. MERIT40 is an integral member of the BRCA1 complex and mediates BRCA1 DSB recruitment through RAP80. IR induced foci formation of ABRAXAS and BRCC36 was also strongly reduced in MERIT40 downregulated cells, suggesting that MERIT40 mediates interaction of ABRAXAS to BRCC36 and RAP80 (**Figure 1.7**). MERIT40 knockdown reduces protein levels of ABRAXAS and BRCC36 as well as RAP80, thereby suggesting that MERIT40 is essential for the stability of BRCA1–RAP80 complex. The reported cancer-causing mutation of BRCA1 M1775R disrupts binding of MERIT40 with BRCA1 and further disrupts DSB localization of BRCA1 to IR-induced foci (IRIF) [116]. MERIT40 largely influences RAP80 stability and DSB localization. Down-regulation of MERIT40 results in loss of RAP80-associated DUB activity compared to BRCC36 depletion. MERIT40 or RAP80 knockdown significantly increases the sensitivity to IR in BRCA1 mutated HCC1937 cells [13]. This least characterized protein, MERIT40, opens the vast perspective to understand the structural and functional mechanism of stability of BRCA1 complex. It would be interesting to look whether it could be the main determinant of BRCA1 complex stability through its binding partners.

1.3.6 BRCA1/BRCA2-Containing Complex (BRCC45): BRCC45 acts as an adapter in the BRCA1-A complex and bridges the interactions between MERIT40 and the other members of the complex (**Figure 1.7**). It modulates the E3 ubiquitin ligase activity of the BRCA1-BARD1 heterodimer. It is also the component of BRISC complex that specifically cleaves 'Lys-63'-linked ubiquitin. It plays a role in cellular homeostasis and differentiation, and acts as a death receptor-associated anti-apoptotic protein, which

inhibits the mitochondrial apoptotic signaling pathway [107]. It regulates TNF- α signaling either directly or indirectly via its interactions with TNFRSF1A [107].

1.4 Ubiquitination and its role in cell signaling

Ubiquitin is a 76 kDa molecule and comprises seven lysine residues involved in lysine specific linkage during chain assembly [117, 118]. Most of the proteins undergo degradation by K48-linked chains in which second ubiquitin forms the isopeptide bond with the previous one [119]. These proteins undergo degradation because of specific recognition of K48 isopeptide linkage by 26 S proteasomes [120]. However, K63-linked chain present non-proteolytic signals that are involved in protein trafficking. [121].

Ubiquitination is a sequential process which involves at least three classes of enzymes: the E1 (ubiquitin-activating enzyme), an E2 (ubiquitin-conjugating enzyme) and an E3 (ubiquitin ligase). At the last stage of this process, an ubiquitin molecule get attached covalently to a Lys residue of the target substrate protein by an isopeptide bond. Monoubiquitinated substrate Lys residue then serves as a target for E3 to assemble next ubiquitin molecule and thereby forming polyubiquitin chain. Eucaryotes produce various types of E3 to ensure the specificity of ubiquitination [117]. E1 forms a thioester linkage between their active site cysteine and the C-terminus of ubiquitin through ATP molecule. The ubiquitin molecule is then transferred to the active site cysteine of E2 (**Figure 1.9**) e.g. In HECT family, ubiquitin is transferred from E2 to cysteine residue of HECT domain of the E3 before being attached to a substrate. Most E3s with the help of their RING domain or a structurally related Ubox bring a substrate and a charged E2 together, thereby activating the E2 to ligate ubiquitin molecule to a Lys in the substrate [122]. Ubiquitin generally attaches to the ϵ -amino group of Lys in substrates, and

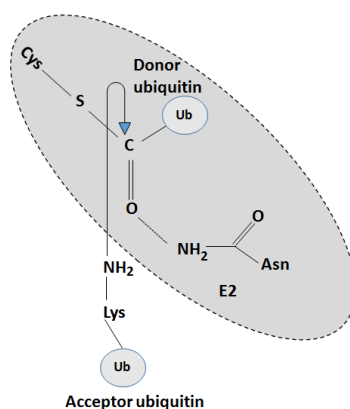


Figure 1.9: E2 Catalytic mechanism for addition of ubiquitin. The model illustrates isopeptide bond formation during ubiquitin chain synthesis.

transfer of a single ubiquitin can occur to one (monoubiquitylation) or multiple (multi monoubiquitylation) sites to recruit binding partners; this promotes or inhibits interactions, protein localizations and activities (**Figure 1.10**) [123].

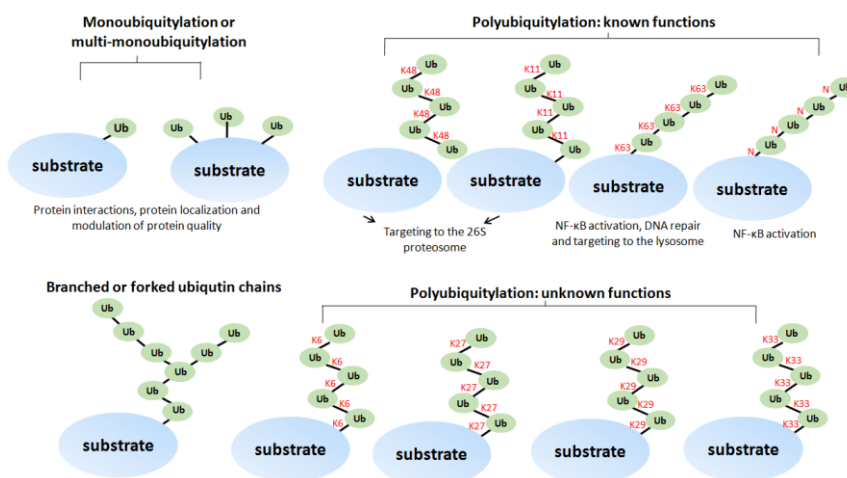


Figure 1.10: Various type of Lys residue specific linkage during (poly)ubiquitination and their fate in various cellular events.

One of the important functions of E2 is to ensure that it receives ubiquitin only and not the ubiquitin like modifiers (UBLs) on its active site (**Figure 1.10**). UBLs and ubiquitin are structurally similar, and the former can also be conjugated to the substrate Lys

residues with the help of E1s, E2s and E3s. Both can compete for the same Lys residue in substrates but modification with UBLs usually amends protein activation or localization, e.g. monoubiquitylated proliferating cell nuclear antigen (PCNA) endorses translation, DNA repair, whereas blocking same lysine with the UBL small ubiquitin like modifier (SUMO) blocks recombination between sister chromatids [124].

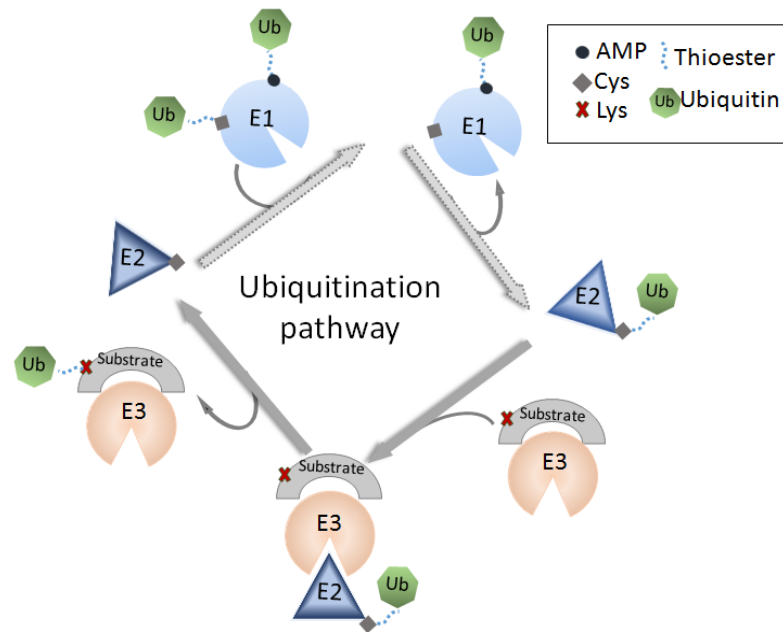


Figure 1.11: Importance of E2 in ubiquitylation. E2 interacts with the E1 that is preloaded with two ubiquitin molecules, one is Cys at its active site and other at adenylation domain as an adenylylate (~AMP). The activated ubiquitin is relocated to the Cys in the active site of E2. The E2 then dissociates from the E1 before it holds with E3, which further recruits substrates. After dissociation of E2 from E3, the former can be recycled for the next round of recharging with ubiquitin.

Once E2 is charged with a ubiquitin molecule, it employs catalysis of substrate ubiquitylation mediated through E3s (**Figure 1.11**). E2 can act together with several kinds of E3s, and is most predominantly seen in UBE2D family members. E2s recognize its counterpart E3s through the L1 and L2 loops and the N-terminal α -helix present on E2 surface, and sequence variations in these motifs determines the specificity of E3 binding (e.g. E2 UBE2L3) [125]. BRCA1 also undergoes polyubiquitination that is

independent of its own enzymatic activity, which involves other E3 ubiquitin ligases [126]. However, this Ubiquitination enhances BRCA1/BARD1 stability rather than targeting it for degradation [127].

Different proteins that contain ubiquitin binding motifs such as the ubiquitin-associated (UBA) domains and ubiquitin-interacting motif (UIM) can recognize substrates modified with ubiquitin [128]. These ubiquitin binding proteins may preferentially unite with different monoubiquitin or a multiubiquitin chain species to bring effectors downstream signaling pathways and thereby couple ubiquitylation to the particular biological function. For example, RAP80 binds to K-63 linked polyubiquitin chain and perform the protein trafficking function in DNA damage repair, whereas those bind to Lys 48-linked chains go for degradation (γ -tubulin) [129].

1.5 Conclusion

BRCA1-complex contains a five-member stoichiometric complex consisting of RAP80, ABRAXAS, BRCC36, BRCC45 and MERIT40. Although, *BRCA1* is the most studied gene, but least characterized for its association with cellular binding partners. It remains elusive how a downstream acting protein, MERIT40 affects the RAP80 and CCDC98 protein's levels. What are the binding partners associated with MERIT40 that are present in BRCA1 complex, through which it may extend its interaction network and maintain genomic stability of BRCA1- complex? Furthermore, the interacting regions of RAP80 and ABRAXAS are not well defined at the atomic level. Structural and functional mechanisms of disease susceptibility due to RAP80 and ABRAXAS mutations remain imprecise. These ambiguities inspired us to explore the structural aspects of wild type and mutant protein, and interaction analysis of BRCA1-complex members. Our findings

suggest that MERIT40 directly interacts with BRCA1 and ABRAXAS, thus helping in the stabilization of BRCA1-complex. These studies will help in elucidating the mechanism of repair defects, and their remuneration for structure based drug lead discovery.

Chapter 2

Materials and Methods

Materials: Restriction enzymes were purchased from Fermentas. All other chemicals were procured from Sigma-Aldrich, unless otherwise specified. Chemicals were analytical grade or molecular biology grade with more than 99.99% purity. Double distilled, 0.44 µM filtered Milli-Q water (Millipore, USA) was used for preparing buffers. Depending upon the requirements, protein and buffer solutions were either centrifuged or filtered through 0.44 micron filter (Millipore, USA) and degassed prior to use for any experiments.

Methods:

2.1. Gene Cloning

2.1.1 Preparation of target gene: For cloning different regions of gene of interest, the c-DNA was PCR amplified using primers incorporated with restriction sites and TEV protease sequence. Following are the steps for cloning:

2.1.2 Primer Designing: Primers were designed using Primer-X software online tool (www.bioinformatics.org/primerx/). The following parameters were taken into consideration.

Length: 18-25 (gene sequence), GC content: 40-60%

T_m: 50-65 °C, GC clamp: Yes (3')

RE site(s)	CAT ATG	GAG AAC CTG TAC TTT CAG GGT (TEV site)	Gene Sequence
------------	---------	--	---------------

2.1.3 PCR amplification: Different oligos were evaluated with the Primer-X suite and used for PCR amplification. The following PCR conditions were applied for amplification of genes of interest using gene specific primers.

1. Denaturation: 95°C for 5 minutes
2. Annealing: 55-65°C for 45 seconds
3. Extension: 72°C @ 0.5 kb/minutes
4. Repeat cycles from step 2 to 4 for 30 cycles
5. Final extension: 72°C for 10 minutes

The PCR amplified product was checked on 1% agarose gel and band of interest was excised. The gel piece containing product of interest was subjected to purification using commercially available kit from Qiagen. The purified DNA was further proceeded for digestion and insertion of gene of interest in a suitable bacterial expression vector.

2.1.4 Restriction digestion: The PCR amplified gene of interest and expression vector were digested with restriction enzymes such as BamH1 and EcoR1. The digestion protocol is as follows.

1. 1-2 µg (20µl) of DNA was taken in an eppendorf and digestion buffer was added corresponding to final volume of reaction mixture.
2. 1µl (10 unit) restriction enzyme was added to the reaction mixture.
3. Reaction volume was adjusted with DNAase and protease free distilled water.
4. The mixture was incubated at 37°C for 1.30 hours.
5. Reaction was terminated by heat inactivation of the enzymes at 85°C for 5 minutes.
6. Reaction product was purified using clean-up kit (Qiagen).

The digested product was then ligated using following protocol:

2.1.5 Ligation: Digested purified product of gene of interest (insert) and bacterial expression vector were ligated with T4 DNA ligase.

1. Vector and insert was taken in an eppendorf tube in 1:3 ratios.

2. 1 µl T4 DNA ligase was added to the reaction mixture.
 3. Ligation buffer was added corresponding to final reaction volume.
 4. Reaction volume was adjusted with distilled water and incubated at 16 °C for 3 hours.
- The reaction product was then used for transformation in bacterial host such as *E.coli*, DH5α strain.

2.1.6 Transformation: It is a process by which the genetic material from external environment is transferred into the organisms by means of extracellular stimulus such as temperature. In bacterial transformation, fragments or intact vector are taken up by a living bacterium under *in-vitro* condition. However, cells have to be made competent enough to enable easy uptake of the naked DNA into the cell.

Competency is a defined physiological state in which bacterial cell(s) uptake naked DNA more readily than in the normal growth conditions. This can be achieved by chemical or physical methods. Chemical method includes treatment of cells to divalent cations such as Mg^{2+} or Ca^{2+} . Physical treatment includes exposure of cells to heat shock treatment, or strong electric field called electroporation. All the described methods make bacterial cell membrane highly porous for easy uptake of foreign DNA. Under suitable condition, the transformation efficiency of most genes would be about 1 in every 10^3 cells [130].

2.1.7 Competent cells: The BL21 (DE3) bacterial strain conveniently expresses protein using the T7 promoter. The (DE3) designation indicates that the host is a lysogen of lambda-prophage (DE3), and therefore carries a chromosomal copy of the T7 RNA polymerase gene under the control of lacUV5 promoter (inducible by IPTG). This strain is most widely used for protein expression because it lacks the Lon and OmpT proteases.

Rosetta 2 (DE3) bacterial strain is a derivative of BL21 and is designed to enhance the expression of proteins containing rare codons. The Rosetta strains supply tRNAs for the rare codons such as AGA, CUA, AUA, AGG, CCC and GGA on a compatible chloramphenicol-resistant plasmid, called pRARE. In addition to the above mentioned rare codons, the strain also supplies the seventh rare codon (CGG). This strain provides a “universal” translation system, which otherwise would have been limited by the codon usage of *E. coli*, and their tRNA genes found under control of their own promoters. In the pLacI and pLysS derivatives of these bacterial strains, the rare tRNA genes are found on the plasmids that carry the T7 lysozyme and *lac* repressor genes, respectively.

Protocol of preparation of competent cells:

1. Streak the cells from glycerol stock or from a vial of competent cells onto a LB agar plate and allow them to grow over night.
2. Inoculate a colony from LB streaked plate into 2ml LB media and allow growing at 37 °C for six hours.
3. Inoculate the starting culture into 500 ml LB media without antibiotic or with antibiotics selective for the host, and incubate it at 37 °C with constant shaking till the O.D reaches between 0.6-0.8 at $\lambda=600$ nm.
4. Immediately stop the growth by placing culture in chilled condition for 5-10 minutes.
5. Centrifuge at 5000 rpm for 15 minutes at 4°C in pre-chilled rotor (sorvall tubes).
6. Discard supernatant and keep pellet on ice immediately (sterile condition).
7. Add 100 ml of pre-chilled TB1 buffer and resuspend the pellet (maintain on ice while re-suspending)
8. Centrifuge at 5000 rpm for 8 minutes at 4°C in pre-chilled rotor (sorvall tubes).

9. Discard the supernatant and keep the pellet on ice immediately (sterile condition).
10. Add 10ml of pre-chilled TB2 buffer and resuspend the pellet (maintain on ice while re-suspending)
11. Make aliquots of 100 µl in 0.5 ml microcentrifuge tubes.
12. Snap freeze with liquid nitrogen and store at -80 °C for further use.

Composition of TB1 (Transfer Buffer 1): 300 mM Potassium acetate, 50 mM Magnesium chloride, 100 mM Potassium chloride, 10 mM Calcium chloride, 15% glycerol

Make up the volume with autoclave double distilled water and store at 4 °C.

Composition of TB2 (Transfer Buffer 2): 10 mM MOPS, 10 mM Potassium chloride, 75 mM Calcium chloride, 15% glycerol

Make up the required volume with autoclaved double distilled water and store at 4 °C.

The competent cells can be tested by streaking on LB agar plates with and without ampicillin.

Protocol of DNA transformation:

1. Allow the competent cells to thaw on ice.
2. Add 1µl (50-100ng) of DNA to 100 µl of competent cells.
3. Incubate on ice for 30 minutes.
4. Give heat shock for 90 seconds at 42 °C.
5. Add 800 µl of sterile LB medium in aseptic conditions.
6. Incubate at 37 °C for 45 minutes on shaker incubator.

7. Centrifuge at 4000 rpm for 5 minutes and discard the supernatant, and re-suspend the pellet in 100 µl of LB and surface spread on LB agar plate containing the selective antibiotic.
8. Incubate the plate at 37 °C overnight to obtain well isolated transformed colonies.

2.2 Protein Expression and Purification

One of important application of recombinant DNA technology is to facilitate expression and purification of huge amount of proteins. One of the factors affecting the level of expression of a cloned gene is the strength of promoter. An efficient expression vector system can convert bacterial cells into a protein production system with yields as high as 10-20% of total cellular protein. Since foreign protein can be toxic to host cells, genes must be under very stringent promoter control. The expression vectors generally contain inducible promoters that can be controlled by an inducing agent, e.g. Lac operon is induced by addition of IPTG. In the absence of lactose in the growth medium, the lac promoter is repressed by Lac repressor protein. Induction or turning on of the Lac promoter is achieved by the addition of either Lactose or its analogue, IPTG (isopropyl-β-D-thiogalactopyranoside) to the medium. Either of these substances prevents the Lac repressor from binding to the Lac operator, thereby enabling transcription to occur.

In BL21 (DE3), the T7 polymerase is under the control of lacUV5 promoter. BL21 (DE3) carries the Lac I gene in its chromosome. Lac I represses the expression of T7 polymerase. Consequently the expression of T7 polymerase is induced by IPTG. Adding IPTG to the culture of BL21 (DE3) strain containing expression vector will induce the expression of T7 polymerase which in-turn will transcribe the gene of interest into corresponding protein.

Requirements: Sterile LB broth containing Ampicillin (100 mg/ml), IPTG, micropipette with sterile tips, sterile 10 ml pipette, table-top centrifuge.

Protocol for protein expression:

1. **Inoculation:** Pick a single transformed colony from antibiotic resistant LB agar plate and inoculate it in 100 ml LB broth containing 100 μ M of antibiotic. Incubate on shaker incubator at 37 °C overnight.
2. **Dilution:** Inoculate 10 ml of starting culture to 1000 ml (1: 100 ratios) of autoclaved LB broth containing 100 μ M of antibiotic. Incubate the flasks on shaker incubator at 37 °C until it has reached mid-log phase i.e. A_{600} between 0.6-0.8.
3. **Induction:** Cool down the flasks and add appropriate IPTG and incubate on shaker incubator at 24 °C for 16 hours.
4. **Harvesting:** The culture was transferred to centrifuge bottles and centrifuged for 10 minutes at 6000 rpm at 4°C. The pellet was resuspended in small amount of supernatant and centrifuged for 15 minutes at 5000 rpm, 4 °C.
5. **Storage:** The pellet obtained can be stored at -80 °C until further use.

2.2.1 Protein Purification: Protein purification is a sequence of practices projected to isolate a particular type of protein from a complex mixture. Highly pure protein is essential for the characterization of the structure, function and interactions of the protein with ligand, protein, DNA and RNA. The precursor substance is usually a microbial culture or biological tissue. The various steps during protein purification method may free the protein from undesired substances, separate the protein from debris of the mixture, and finally isolate the desired protein from the whole matrix. Separation of desired protein from all others is usually the most difficult aspect of purification.

Purification steps may utilize differences in physiochemical properties such as (for example) protein size, binding affinity and biological activity.

Generally, a purification procedure contains one or more chromatographic separation steps. The fundamental procedure in chromatography is to stream the medium containing the proteins through a packed column with various substances such as agarose. Different proteins interact in a different way with the column matrix and undergo separation during passage through the column. Regularly, proteins are detected as they elute from the column by their absorbance at 280 nm. Many different chromatographic methods exist including,

- Size exclusion chromatography: Separate the molecules on the basis of size
- Hydrophobicity chromatography: Separate the molecules on the basis of hydrophobicity
- Ion exchange chromatography: Separate the molecules on the basis of charge
- Affinity chromatography: Separate the molecules on the basis of affinity towards a tag

Affinity chromatography provides the highest specificity and selectivity for the purification of bio-molecules. Protein molecule has a three dimensional structure, and ligands having ability to recognize one of the structural motifs of a protein is used in affinity chromatography. This interaction can be exceptionally specific and the ligand molecule interacts only with the engineered molecule, or the ligand can be designed in such a way so that it recognizes a group of structural analogous. Ligand on the solid support should possess excellent avidity for its target selection; however, it should be reversible. If the adsorption of target bio-molecule(s) is too strong it will make it difficult

to separate it from bound ligand. If the affinity ligand is a peptide or protein, special care needs to be taken to keep them physiologically active during the process of purification. Since affinity chromatography is highly specific, it is considered as a powerful tool in purifying proteins from a complex mixture.

General Protocol for purification of Proteins:

1. **Re-suspension:** Re-suspend the pellet in 40 ml of purification buffer; 200 μ l of 200 mM PMSF and 20 μ l of protease inhibitor cocktail were added.
2. **Ultra sonication:** Transferred the suspension into centrifuge tube and sonicate at 50 pulse rate and 50 power with 1.45 minutes of duty cycle. This process is breaking open of the cells using high frequency sound waves. It involves application of sound energy using an ultrasonic bath or probe to disrupt the cell membrane and remove its cellular contents.
3. **Centrifugation:** Centrifuge at 18000 rpm for 45 minutes at 4°C to obtain pre-cleared lysate. Collect the supernatant and discard the pellet
4. **Equilibration of GST/Ni-NTA beads:** Give two column volume washes with distilled water to remove traces of ethanol as the beads are stored in 20% ethanol and then 3 column washes with purification buffer.
5. **Binding:** Bring pre-cleared lysate to room temperature and carefully mix with affinity resin, and incubate at room temperature for one hour.
6. **Washing:** After binding, give four column washes with wash buffer so as to remove non-specific bacterial protein. Take 40 μ l of beads to load on gel to observe the bound protein.

7. **Cleavage:** For removal of GST/his tag, add 400 μ l of TEV protease, 40 μ l of protease inhibitor and 100 μ l of PMSF in 20 ml of purification buffer and perform the cleavage step for 3 hours by passing the solution through column at interval of 1 hour. Take out 40 μ l of beads and run on SDS-PAGE to observe cleavage of protein. The TEV protease is highly site-specific cysteine protease that is found in Tobacco Etch Virus (TEV). The optimum recognition site for the enzyme is the sequence Glu-Asn-Leu-Tyr-Phe-Gln (Gly/Ser) and the cleavage occurs between the Gln and Gly/Ser residues.
8. **Elution:** After TEV cleavage, elute the protein in 30 ml of purification buffer.
9. **Equilibration of Ni-NTA resin:** Give two column washes with double distilled water followed by six column wash with washing Buffer.
10. **Metal Ion Chelate Affinity Chromatography:** After equilibration of Ni-NTA resin, pass the eluted fractions. This step is done to remove the residual his-tagged TEV enzyme from the solution of our protein of interest
11. **Equilibration of centricon:** Wash the centricon (10-50 kDa Molecular weight cut off) with water to remove traces of alcohol, calibrate it with wash buffer by centrifuging it at 4500 rpm for 10 minutes at 4°C.
12. **Concentrating the protein:** Transfer the eluted protein in the centricon and concentrate it upto 2 ml by centrifuging it at 4500 rpm for 10 minutes at 4°C. Check the concentration on Nanodrop spectrophotometer (280 nm). Centrifuge for 10 minutes at 130000 rpm at 4 °C to remove precipitates or aggregates, if any.
13. **Gel filtration:** Inject 2 ml of concentrated protein on AKTA- FPLC pre-equilibrated with FPLC buffer.

14. **Fraction collection:** Monitor the elution of the protein by recording the absorbance at wavelength 280 nm. Collect the fractions corresponding to the exact elution volume for protein, which is dependent on its molecular mass.
15. **Loading on SDS-PAGE:** Load 20 µl of FPLC fractions on SDS-PAGE, stain with commassie dye, and then destain to visualize the protein bands.
16. **Concentrate the protein:** The fraction which shows purified protein band at the expected size can be further concentrated as per the experimental requirement.

2.2.2 Regeneration of GST resin:

Composition of Wash buffer: 300 mM NaCl, 50 mM Tris

Buffer 1: 0.5 M NaCl, 0.1 M Tris (pH 8.5)

Buffer 2: 0.1 M Na-acetate, 0.5 M NaCl (pH 4.5)

Protocol:

1. Pass 20 mM reduced glutathione solution prepared in 50 mM Tris pH 8.0, twice, through the GST resin to remove traces of bound protein and recharge the resin with glutathione.
2. Give around two-three washes with distilled water.
3. One column washes with washing buffer followed by one column wash with 1% TritonX 100.
4. Pass one column volume of 70% ethanol followed by extensive washes with buffer-1 and buffer-2 alternatively.
5. Finally, give one column wash with Wash Buffer and a subsequent wash with distilled water.
6. Store the resin in 20% ethanol at 4°C.

2.2.3 Regeneration of Nickel resin:

Regeneration buffer: 6 M GnHCl, 0.2 M acetic acid

Protocol:

1. Pass 750 mM imidazole solution prepared in lysis buffer (pH 7.5), followed by two column washes with regeneration buffer.
2. Give five column washes with distilled water to remove traces of GnHCl with subsequent three column washes with 2% SDS.
3. Pass one column each of 25%, 50%, 75% and 100% ethanol. Repeat the step the inverse order.
4. Give one column wash with distilled water followed by five column washes with 100 mM EDTA (pH-8.0).
5. Re-charge the column by passing two column volumes of 100 mM NiSO₄ with subsequent water washes to remove unbound nickel.
6. Store the resin in 20% ethanol at 4 °C.

2.3 Protein characterization (RAP80, MERIT 40 and ABRAXAS)

2.3.1 Protein Estimation: Quantification of the protein was performed with Bradford (Expedon) and Nanodrop estimation protocol as per manufacturer's instructions.

1. BSA (0.1, 0.5, 1.0, 2.5, 5, 10 mg/ml) dilutions were prepared as a standard reference and absorbance was recorded at λ 595 nm using a spectrophotometer (Shimadzu).
2. For Nanodrop protein estimation, the absorbance was recorded at 280 nm and protein concentration was estimated by considering the extinction coefficient. The extinction coefficients of MERIT40, BRCT1-BRCT, RAP80 (1-130), RAP80 (1-405), CCDC98 (6-268), CCDC98 (6-301), CCDC98 (6-373), CCDC98 (6-409) were 28920, 36815, 125, 32470, 17670, 17795, 24785, 26275 M⁻¹ cm⁻¹, respectively.

2.3.2 Mass Spectrometry: The mass spectrometer has long been an indispensable tool in life science research. Mass Spectrometry provides information for proteomics profiling. The technique requires only miniscule amount of sample (1-10 ng) that can be readily applied from small amounts of protein extracted from a two-dimensional electrophoretic gel. Molecules to be analysed, referred as analytes, are first subjected in vacuum. When the charged molecules are introduced in an electric and/or magnetic field, their paths through the field are function of their mass-to-charge ratio, m/z . This measured property of the ionized species can be used to deduce the mass (M) of the analyte with very high precision.

Mass spectrometry can be used to sequence short stretches of a polypeptide using a technique called *tandem MS or MS/MS*. A solution containing the protein under investigation is first treated with protease or chemical reagent to hydrolyse it to a mixture of shorter peptides. The mixture is then injected into device that is mass spectrometers in tandem.

Types of Mass spectrometry: Mass spectrometry can be classified into various types on the basis of sample introduction/ionization methods-

- 1) Electron Impact (EI): This is a hard but versatile method which provides structural information for the molecule upto 1000 daltons.
- 2) Chemical Ionization (CI): It is a soft method for generation of molecular ion peak $[M+H]^+$ and can be employed for molecules upto 1000 daltons.
- 3) Electrospray (ESI): It is a soft method for generation of multiply charged species and can be employed for molecules upto 200,000 daltons.

- 4) Matrix Assisted Laser Desorption (MALDI): It is a soft method for generation of multiply charged species and can be employed for molecules upto 500,000 daltons.

Protocol for mass spectrometry sample preparation:

Destaining the gel particles-

1. Cut the gel into very small pieces and, transfer them into fresh destainer, and keep on vortex overnight for destaining.
2. Wash the gel band pieces with distilled water (3 changes for 10 minutes each) to remove destaining solution.
3. Wash the gel bands with 50 mM NH_4HCO_3 + 50% acetonitrile (1:1) mixture to provide buffering and destaining, and keep for 15 minutes at room temperature.
4. Remove the excess liquid and overlay it with sufficient amount of 100% acetonitrile to cover the gel allowing the gel plugs to shrink & stick together.
5. Remove the acetonitrile, and then rehydrate the gel pieces in 50 mM NH_4HCO_3 .
6. After 5 minutes add equal volume of 100% acetonitrile and allow the gel to shrink.
7. Dry down gel particles in vacuum spin dry (for 10-25 minutes) to evaporate the solvent.

Reduction and Alkylation-

1. Swell the gel particle (freshly prepared) in 10 mM DTT/50 mM NH_4HCO_3 (300 μl /tube) for reduction.
2. Incubate the tubes at 56 °C for 45 minutes. Immediately bring the tubes at room temperature.
3. Remove excess liquid and replace it quickly by nearly the same volume of freshly prepared 55 mM IAA (iodoacetamide) in 50 mM NH_4HCO_3 for alkylation.

4. Incubate for 30 minutes at room temperature in dark.
5. Remove the IAA solution and wash the gel particle with the mixture of 50mM NH_4HCO_3 and 50% acetonitrile (1:1).
6. Add enough of 100% acetonitrile (200 μl) to cover the gel pieces.
7. After the gel pieces have shrunk remove acetonitrile.
8. Dry down the gel particles in speed vaccum (5-6 minutes).

In Gel Digestion-

1. Add excess of freshly prepared enzyme solution (Trypsin + NH_4HCO_3). The concentration of trypsin can be prepared in the ratio of 1:100 or 1:10.
2. Remove excess of solution and incubate the tubes at 37 °C overnight in standing position (nearly 16 hours of incubation).

Extraction of peptides-

Composition of extraction buffer: 100% acetonitrile + 5% TFA

1. Use ~ 30 μl of extraction buffer each time during vortex (three times, keep extracted peptides in fridge for further use).
2. The solution is extracted from the sample and centrifuged for 1-2 minutes at 1000 rpm.
3. Dry the sample (45-50 minutes) using speed vaccum.
4. Store vaccumised samples at -20 °C until next step.

2.3.3 Circular Dichroism: Circular Dichroism (CD) refers to the differential absorption of left and right circularly polarized light and is revealed by the absorption bands of chiral molecules. As a result, CD is displayed by biological molecules because of their levo-rotary and dextro-rotary components. Secondary structural characteristics also

impart a distinct profile of respective molecules; therefore, the α -helix content of proteins has signatures in CD spectrum representative of their structures. The capability of CD to give a characteristics signature absorption peak makes it a powerful tool to study the structural components in protein-biochemistry. The far-UV spectra of proteins reveal important evidence of their secondary and tertiary structure pattern. Secondary structure can be monitored during protein unfolding which depicts the folding pathway and stability. Absorption minima at λ -208 nm and 222 nm indicate α -helical structure, whereas a minimum at λ 218 nm is a characteristic of β -sheets. CD signal is strongest at the aforesaid wavelengths; however, absorption spectra of buffer components may require monitoring at somewhat higher wavelengths. Routine monitoring of CD spectrum at 208, 218, 222 and 228 nm allow to obtained higher signal to noise ratio with respect to secondary structural changes at different urea concentration while avoiding amplification of the voltage at the higher concentrations of denaturant.

CD Polarimeter: Jasco J-815 spectropolarimeter fitted with a Peltier temperature controller.

Protocol:

1. The CD cuvette must be thoroughly washed with 0.1% triton, repeated distilled water washes followed by final wash with 100% ethanol and allowed to dry.
2. 10 μ M protein solution (in 2.5 mM HEPES, pH 7.5, 50 mM NaCl) is scanned in a wavelength range of λ 200-240 nm at 10 °C.
3. Molar ellipticity was calculated by considering the actual protein concentration and mass , and further data analysis was done using DichroWeb server (<http://dichroweb.cryst.bbk.ac.uk>).

2.3.4 Fluorescence Spectroscopy: The phenomenon of fluorescence is exhibited when a system absorbs light at a particular wavelength and shows emission spectra at higher wavelength. This phenomenon depends upon the molecular conformation and dynamics. According to quantum mechanics, the molecules have different electronic level due to the vibrational motion. Each electronic level further comprises bands of closely spaced energies which exist at different vibrational modes. A molecule typically exists in a ground state unless excited by beam of light, some of the energy is absorbed and electron reached to the excited state from where they may return to different vibrational levels. Furthermore, the electrons at higher electronic level dissipate some of their energy while going to a lower vibrational level. The electron emits radiation of lesser energy and longer wavelength during this process of returning to ground state. The molecule is thus said to have fluorescence.

The fluorescence of a well folded protein is a combination of the fluorescence from individual aromatic residues. The major contribution in intrinsic fluorescence emissions of a folded protein is due to tryptophan residue, with some counterpart from tyrosine and phenylalanine. Disulfide bonds also have considerable absorption in the same range. Emission scans are acquired from λ 310 to 400 nm, following excitation at λ 280 or 295 nm. The λ 280 is used to trace the environmental changes associated with tryptophan and tyrosine side chains due to the absorption wavelengths at 280 and 275 nm, respectively. Emission spectrum corresponding to λ 295 nm follows tryptophan emission selectively because of least contribution from tyrosines.

In general, fluorescence emission maxima use two methods for monitoring equilibrium unfolding. In one method, one can acquire an emission spectrum for the native and

unfolded protein using denaturant such as temperature and urea by using a single wavelength scan to monitor the structural changes. However, second method employs an emission spectrum at each concentration of denaturant, and estimates the average emission wavelength (AEW) for each sample.

Spectrofluorimeter: JOBIN YVON Horiba Fluorolog 3

Protocol:

1. The tryptophan and tyrosine positions were monitored using Fluorolog 3 (Horiba, USA) at excitation wavelength of $\lambda = 280$ nm and temperature of 20 °C.
2. Fluorescence emission scans were recorded from $\lambda = 310$ -400 nm wavelength.

The cuvette is thoroughly washed with 0.1% Triton-X, distilled water followed by 100% ethanol and, then allowed to dry.

There are two widely used methods to monitor the protein unfolding and for determining melting temperature.

1) Thermal Denaturation: A study of protein folding is very challenging and subject of extensive studies. Under physiological conditions, proteins exist in different conformation, where each conformation is shown to have a characteristic of closely related set of structures that oscillate around an energy medium. The forces that constrain proteins to assume their native ensembles generally defined as the difference between the entropy state of unfavourable chain and the favourable enthalpy interactions, where the stability of protein's structures can vary drastically. The kinetic and thermodynamic folding parameters of protein can be elaborated by utilizing intrinsic probes such as fluorescence and CD.

CD is one of the tools that help in to study the conformational stability of a protein under stress. Thermal stability can be assessed by monitoring the spectrum with increasing temperature. In some cases it is required to obtain the entire spectrum in the far- or near-UV CD region at different temperatures. Additionally, a single wavelength scan can provide alternate choice which monitors some structural feature of the proteins. Most often proteins aggregate or precipitate which in turn results in loss of protein folding. The reversibility of the unfolding transition can be measured by cooling the sample to its native conformation and then re-heating to observe the folding pattern. Another challenge is to find suitable buffer conditions and additives that make unfolding reversible and is important for long-term stability and storage.

Far-UV CD region can be used to evaluate the secondary structural changes in proteins. It can predict, whether protein is losing its secondary structural component at high temperatures or endures conformational change.

For thermal denaturation in CD, protein (10 μ M) was unfolded in a temperature range of 10-60°C at 208-230 nm wavelength. Fraction unfolded was calculated at the different temperatures—and plotted against each other to determine the thermal parameter and melting temperature .

In fluorescence, the tryptophan and tyrosine positions can be monitored at excitation wavelength of $\lambda=280$ nm and temperature of 20 °C. Emission scans were recorded from λ 310-400 nm wavelength. 2 μ M protein was unfolded in a temperature range of 15-80 °C and emission maxima were obtained.

Data fitting can be done in two to three states (depending upon transition in folding curve) and thermodynamic parameters can be calculated.

Protocol for RAP80 Thermal denaturation:

For thermal denaturation in CD-

1. Wild- type and mutant purified protein (10 μ M) were unfolded in a temperature range of 10-60 °C at λ = 218 nm wavelength.
2. Data fitting was done in two-state transition model and thermal parameters were calculated.

Protocol for MERIT40 thermal denaturation:

1. 2 μ M protein was unfolded in a temperature range of 15-80°C and emission maxima were obtained.
2. Data was analyzed by fitting in the data in a three-state transition model and thermal parameters were calculated.

2) Chemical denaturation: Fluorescence is an extremely sensitive technique which can detect even minor conformational changes, thereby making it a useful tool to evaluate protein tertiary structure. It requires very low protein concentrations, and have the ability to selectively monitor localize regions or motifs, and the use of a multitude of solution conditions. CD is also one of the useful tools to examine the alterations in protein's secondary or tertiary structures while a number of chaotropic agents are available to denature the protein, however we have selected urea for our study. Equilibrium-denaturation studies allow the calculation of the free energy, enlightening the stability of the native conformation and intermediates (if applicable). The kinetic study helps to decipher the folding pathway and intermediates that may not be detectable in equilibrium. For most of the biophysical study, it is very important to determine accurate protein concentration. Using protein sequence, the concentration of the native stock can be

readily obtained by determining the absorbance at $\lambda=280$ nm and considering extinction coefficient. The concentration required for measurement of equilibrium unfolding experiment is typically in the micromolar range but also depends on the number of aromatic residues. Fluorescence emission is recorded by considering tryptophans and tyrosines because of their high quantum yield at the wavelength of excitation. There are a range of buffers available that can be used for fluorescence and CD measurements, and their absorbance should be corrected to obtain protein spectra.

Following are the steps used to set up an equilibrium unfolding experiment:

1. Prepare 10 M urea stock.
2. Confirm that the protein is completely denatured.
3. Optimize instrument settings using native and denatured protein in their respective buffers.
4. Set up the samples in varying concentrations of denaturant.
5. Establish equilibration times and reversibility.
6. Perform equilibrium-unfolding experiments.
7. Repeat experiments at different protein concentrations to determine the model to be used for fitting the data.

For chemical denaturation using fluorescence, 10 M urea (Sigma-Aldrich) stock and protein dilutions were prepared. Emission scans were collected (as described above) and was blank subtracted in order to get high signal to noise ratio. Data fitting was done in two or three state (depending upon the number of transitions in folding curve) and further thermodynamic parameters were calculated.

Thermal and chemical denaturation data analysis and curve fitting: Data fitting (CD, Fluorescence) was done using Origin (OriginLab, Northampton, MA) for three state transition model and thermal parameters were calculated by considering native (N) to intermediate (I) and intermediate to unfolded (U) transition [131]



Here, K_1 and K_2 represent the equilibrium constants for the two reactions.

In order to visualize different spectroscopic signals on a same scale, data was normalized using following equation.

$$f_F = (Y_X - Y_U) \div (Y_N - Y_U) \quad (2)$$

In this case, Y_X and Y_U are the signal being normalized and the signal of the unfolded protein respectively, and Y_N is the signal of the native protein. f_F represents the fraction folded at a particular temperature or denaturant concentration. A plot of the normalized signal (fraction folded, f_F) versus the denaturant concentration or temperature, generate the unfolding curve for each of the spectroscopic probes.

The equilibrium constant (K) can be derived from f_F using following equations.

$$K = f_U/f_F = f_U/(1 - f_U) \quad (3)$$

Where, f_U represents the fraction unfolded at particular temperature of the denaturant concentration. Following equations were employed for calculation of free energy (ΔG) and enthalpy (ΔH) of unfolding reaction from K .

$$\Delta G = -RT \ln K \quad (4)$$

$$d(\ln K)/d(1/T) = -\Delta H/R \quad (5)$$

Where, R is the gas constant ($1.985 \text{ calK}^{-1}\text{mol}^{-1}$) and T is the temperature (Kelvin)

$$\Delta G = \Delta G^{\text{H}_2\text{O}} - m[\text{urea}] \quad (6)$$

Where, $\Delta G^{\text{H}_2\text{O}}$ is defined as an estimate of the conformational stability of a protein in the absence of denaturant (i.e Intercept of linear extrapolation of ΔG value), m is the slope of the plot ΔG vs. urea and it measures the dependency of ΔG on urea concentration [132] [133].

2.3.5 Differential Scanning Calorimetry: Differential scanning calorimetry (DSC) is a thermo-analytical method based on the principle of heat energy generated or dissipated during interaction process. The amount of heat required to keep the temperature of a sample constant with respect to reference cell is measured as a function of temperature. Therefore, both the sample and reference cell are maintained at almost the same temperature throughout the experiment. The reference cell must have a well-defined heat capacity over the range of scanning temperatures. When the sample undergoes physical transformation such as phase transitions, in order to maintain the sample cell at constant temperature, continuous heat energy needs to be supplied from external source. Depending upon the extent of heat energy exchange with respect to reference cell, the process can be defined as exothermic or endothermic. Thermal unfolding of protein of interest was done using Differential Scanning Calorimetry (Setaram μ DSC3 evo, USA).

Protocol for RAP80:

1. Protein and buffer were filtered and degassed prior to undertaking the scan.
2. A total of 2 mg protein (wild -type) and 0.2mg (Δ 81E) in solution form was allowed to unfold in 5-60 °C temperature range with a temperature increment rate of 1°C/min.
3. Data was fitted locally by “CALISTO” software in two-state transition model.

4. The thermodynamic reversibility of the protein unfolding was determined by heating the sample just above the transition maximum, cooling instantaneously, and then reheating.

Protocol for MERIT40:

1. Thermal unfolding of protein was done in the 15-80 °C temperature range.
2. Protein and buffers were degassed prior undertaking the scan in order to avoid air bubble interference.
3. A total of 1mg protein in solution form was run with a temperature increment rate of 1°C/min.
4. The experiment was replicated thrice independently and blank correction was done. Observed values of furnace temperature vs. heat flow diagram were fitted locally by “CALISTO” software (AKTS) and enthalpy value was calculated.
5. The thermodynamic reversibility of the transition was determined by heating the sample to a temperature that was little over the transition maximum, then cooling immediately and reheating.

2.3.6 Limited Proteolysis: Limited proteolysis is used to probe structural features of proteins. Proteolytic probes such as trypsin and chymotrypsin can locate the sites of a polypeptide chain characterized by enhanced backbone flexibility. Limited proteolysis was used to analyze the molten globule states of several proteins, such as alpha-lactalbumin, cytochrome c, apomyoglobin and human growth hormone. The target proteins would be induced to acquire the partially folded state under specific conditions, such as low temperature and pH. The protein conformational features obtained from other biophysical tools can be well correlated with those deriving from limited proteolysis. It

can also be useful to isolate the protein fragments that can fold autonomously and thus behave as domains. It can also find the stability of intact protein and further can be used to compare with mutant which might have altered structural conformation.

Protocol for RAP80 and Δ E81 mutant:

1. Equal concentration of RAP80 wild-type and Δ E81 (0.2mg/ml) was incubated with trypsin and chymotrypsin separately so that final concentration of proteases were 40 μ g/ μ l and 10 μ g/ μ l respectively.
2. Reaction mixture was incubated for different time frame 0, 10, 30, 60, 180, 360 minutes at 37°C (trypsin) and 25°C (chymotrypsin), respectively.
3. Reaction was terminated individually by adding 1mM PMSF .
4. Samples were heated by adding equal volume of laemmli buffer and analyzed by SDS-PAGE.

Protocol for MERIT40:

1. Trypsin and chymotrypsin were mixed to purified protein (1mg/ml) separately with their final concentration of 40 μ g/ μ l and 10 μ g/ μ l, respectively, and untreated protein was taken as a control.
2. Reaction mixtures were incubated at 37°C (trypsin) and 25°C (chymotrypsin) for different time period, 0, 10, 30, 60, 180, 360 minutes. Reaction was terminated by adding 2 μ l of 200 mM PMSF.
3. Samples were heated with laemmli buffer and analyzed on SDS-PAGE Band corresponding to 29 kDa was considered as a most stable fragment, which was further subjected to trypsin digestion followed by mass spectrometry.

4. Domain of interest was identified by Mascot analysis with Biotoool software (Bruker Daltonics, USA) [134].

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

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

Numerical approximations require:

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

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

Non-Bonded Interactions:

1) Electrostatic interactions

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

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

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

2) Van der Waals interactions

Attractive part: due to induced-dipole/-dipole; *Repulsive part:* due to Pauli's Exclusion Principle. Usually represented by the Lennard-Jones potential are obtained from the single atom parameter ϵ and σ .

$$V_{vdW} = \sum_{i>j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$

$$\epsilon_{ij} = (\epsilon_i \epsilon_j)^{1/2}, \sigma_{ij} = 1/2(\sigma_i + \sigma_j)$$

1) Hydrogen bonds (Hb)

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

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

2) Hydrophobic interactions

This type of interaction become favorable due to combined effect contact between water, polar medium and hydrophobic groups. These interaction are usually present upto a distance of 4.7 Å.

2.3.7 2 General steps for molecular modelling, simulation and docking:

1. Structure retrieval or generation: Crystal structures of bio-molecules can be obtained from Protein Data Bank (PDB) (<http://www.rcsb.org/pdb/>). Protein having least homology can be modeled using *ab-initio* method.
2. Structural visualization: The most efficient means to visualize the three dimensional structure molecule at atomic level is computer graphics. Several software (e.g. Pymol, Cn3D, RasMol, CCP4MG and KineMage) are available for visualization and formation of molecular structures.
3. Energy minimization: Energy level is the most basic property of molecules that can be calculated using three major theoretical methods, (i) molecular mechanics, (ii) semi-empirical (quantum mechanics), (iii) *ab-initio* (quantum mechanics) approaches. Energy minimization is the first step for geometry optimization of the molecular structure.
4. Dynamics simulation and conformation search: Quantum dynamics is defined as solving motion equation of nuclei in the average field of the electrons, whereas solving Newton's law of motion for the nuclei is known as molecular dynamics. Integration of latter for all atoms in the system generates molecular trajectories. Conformation search of bio-molecules can be conducted by repeating the process of rotating dihedral angles to obtain lowest energy conformations of molecular systems.

5. Calculation of molecular properties: Some computing properties are boiling point, thermodynamic quantities, solubility, molar volume, heat capacity, molar refractivity density, dipole moment, magnetic susceptibility, partial atomic charge, ionization potential, electrostatic potential, solvent accessible surface area and van der Waals surface area.
6. Structure superposition and alignment: to get the best energy minimized model, it often involves comparison of series of homologous molecules, and this requires superposition or alignment of structures.
7. Molecular interactions, docking: In intermolecular interaction studies, usually the receptor (e.g., protein) is kept rigid or partially rigid while the conformation of ligand/protein is dynamic.

Molecular modeling has become a valuable and essential tool to medicinal chemists in the drug design process.

Protocol for RAP80 molecular modeling & docking:

1. Protein structures RAP80 (Δ E81UIMs, 79-124 amino acids) was modeled using homology modeling server (Swiss modeler) by using the coordinate of NMR structure (PDB ID; 2RR9) as template.
2. Good-quality model was selected based on overall stereochemistry, Ramachandran plot and validated from server “**SAVES**” (Metaserver for analyzing and validating protein structures, <http://nihserver.mbi.ucla.edu/SAVES/>).
3. SAVES mainly comprises five programs, Procheck, What_check, Errat, Verify_3D and Prove. Modeled structure was simulated for 5 ns using Desmond software (Schrodinger) and superimposed onto wild-type complex.

4. PDBsum was obtained to analyze the interactions.

Protocol for MERIT40 molecular modeling, docking, DALI search and mutational analysis:

DALI server is an important tool that predicts the best structure homologs of the novel protein (Holm & Rosenstrom 2010). It has been reported that structurally similar proteins most likely play similar role in a particular biological occurrence [135].

1. Protein structure was modeled using Robetta server (<http://robetta.bakerlab.org/>) [136].
2. Best model was selected based on overall stereo chemical quality, Ramachandran plot and validated by “SAVES” server.
3. Protein was simulated for 5 ns using Desmond (Schrodinger).
4. Monomer-monomer geometrical docking was carried out using Patchdock & SymmDock server (<http://bioinfo3d.cs.tau.ac.il/SymmDock/>) [137].
5. Residues at the dimer interface were identified with the help of PDBsum software. Each residue was further mutated with bulky, charge repulsive and alanine residue followed by MD simulation for checking their ability to form dimer.
6. Monomer-monomer docking was carried out and best docked structures were evaluated and selected on the basis of docking score and atomic contact energy.
7. Monomer and dimer radius were measured from the docked structure using PyMol software. Structural homologs of MERIT40 were searched using DALI server [14].
8. For mutational studies, substitution was incorporated in the modeled structure corresponding to a particular variant and structural analysis was performed using

Ligplot (<http://www.ebi.ac.uk/thornton-srv/software/LIGPLOT/>) to observe changes in the molecular environment.

2.4 Protein-Protein Interactions

Protein–Protein Interactions (PPI) studies are very important to understand biological functions of complex molecules. Interactions profiles of proteins have been studied from the perspectives of quantum chemistry, biochemistry, chemical biology, molecular dynamics, signal transduction and genetic/epigenetic networks.

Molecular processes in the cell system are conceded by molecular mechanism of large pool of proteins organized by their protein–protein interactions. e.g. signals transduction from cell exterior to interior is accomplished by protein–protein interactions of the signaling molecules and these processes play a fundamental role in many biological activities and diseases such as cancer. Binding surface and energetics for protein-protein interactions are different for different proteins (e.g. nuclear pore importins transport form cytoplasm to nucleus or vice versa) and it also rely on several modification (e.g. a protein kinases add a phosphate group to the target protein). There is a auto-regulatory loop that monitor the outcome of a biological process and has the ability to undergoes changes in protein -protein interactions (e.g. SH2 domains containing protein bind to others only upon phosphorylation of tyrosines while bromodomains specifically recognize acetylated lysines). PPI is central to virtually every process in a living cell. The PPIs help us to understand the bioactive core located at the complex interface and design of inhibitor or provide leads for new therapeutic approaches.

As protein–protein interactions are highly important and there are different systems to analyze them, each system has its own merits and demerits. A highly sensitive method

detects all the interactions in great depth. Following are some of the widely used techniques for measurement of PPIs-

2.4.1 Isothermal Titration Calorimetry: Isothermal titration calorimetry (ITC) is a physio-chemical technique used to determine the thermodynamic parameters of molecular interactions in solution, and most often employed to calculate the binding affinity of different molecules like ligand/protein to another proteins, DNA etc. It is a quantitative method to determine the enthalpy changes (ΔH), binding affinity (K_a) and binding stoichiometry (n) of the interaction between two or more molecules in solution. These measurements then can be utilized to calculate Gibbs free energy changes (ΔG) and entropy changes (ΔS) using the equation:

$$\Delta G = -RT\ln K_a = \Delta H - T\Delta S$$

(where R represents the gas constant and T is the absolute temperature).

An isothermal titration calorimeter consists of two cells (reference cell and sample cell) to maintain isothermal condition. These cells comprise highly thermal conducting and chemically inert substance (Hastelloy alloy or gold), surrounded by an adiabatic jacket. The temperature difference between the reference cell containing water and the sample cell containing the macromolecule is detected and maintained using sensitive thermopile/thermocouple circuits. A known concentration of ligand is titrated into the sample cell which upon interaction results into heat dissipation or generation (exothermic or endothermic reaction). The change in heat during reaction is quantified in term of time-dependent input of power required to maintain identical temperatures between the sample and reference cells.

In an exothermic reaction, feedback power of the sample cell is decreased because of increase in the temperature of sample cell upon addition of ligand. In an endothermic reaction, to maintain a constant temperature, feedback circuit increases the power of the sample cell.

Protocol for MERIT40 and CCDC98

1. MERIT40 was taken as titrant (30 μ M) and ABRAXAS (290 μ M) as titer and the measurements were performed using Microcal ITC200.
2. The reaction was allowed to proceed at 25°C with a stirring speed of 1000 rpm. There was a total of 23 injections, 2 μ l each with an equilibration time of 4 seconds.
3. Blank correction was done and affinity was determined using one-site binding model in Origin software.

2.4.2 Surface Plasmon Resonance: Surface Plasmon Resonance (SPR) is the collective term used for oscillation of electrons in a solid or liquid field due to incident light. The resonance condition occurs when the natural frequency of surface electrons oscillating against the restoring force of positive nuclei coincide with the frequency of light photons. SPR measures adsorption of material onto inert substances such as gold or silver surfaces or onto the surface of metal nano particles. The basic principle involves the changes in the reflective index as a result of absorption (interaction) of molecules onto prefixed ligand.

Protocol for RAP80 interaction with Di-UB (K-63 linked)

1. Interaction studies between RAP80 (1-130), Δ 81E and di-ubiquitin (K63 linked) was performed using BIAcore 3000 (GE, Sweden).

2. A total of 5µg ligand was immobilized on CM5 sensor chip using amide coupling method.
3. Different concentration (0,100, 200, 400, 800, 1600 nM) of RAP80 (1-130) and Δ81E (analytes) were passed on the chip at a flow rate of 20 µl/min.
4. Sensor chip was regenerated with 2 M glycine, pH 2.0. Sensorgram was obtained after blank correction. The kinetic data was fitted in 1:1 Langmiur binding model to calculate the rate and binding constants.

2.4.3 Pull-down assay: Pull-down assays are generally used for qualitative measurement of interactions between a probe protein and unknown targets. The bait protein is a fusion such as GST, which is cloned into an inducible expression vector (e.g. pGEX). This fusion protein can be expressed in bacteria system and purified by chromatography procedure. Prey is usually obtained from the lysates of cells, either [³⁵S] methionine labeled or unlabeled, depending on the detection method. The cell lysate and the bound fusion protein on agarose matrix are incubated together and thereafter washed with buffer. Protein complexes are either eluted or directly loaded onto SDS-PAGE and are analyzed by western blotting, autoradiography or staining.

Protocol for RAP80 and Di-UB (K-63 linked) assay

1. Bacterial pellets of GST-RAP80 (1-130) and Δ81E were re-suspended in buffer and sonicated.
2. Soluble fusion protein(s) bound on glutathione resin (0.5 mg/ml) was used to capture the prey. Resin was washed with same buffer and loaded on SDS-PAGE. Complex was transferred to PVDF membrane (Millipore) and probed with anti-ubiquitin antibody (Abcam).

Protocol for MERIT40 and CCDC98

1. Bacterial pellets of GST and GST-MERIT40 were re-suspended in 10 mM HEPES buffer (pH 7.5) containing 300 mM NaCl, 10 mM β -ME, 0.1 mM EDTA and 5% glycerol, and sonicated.
2. Soluble fusion protein(s) bound on glutathione resin (0.5 mg/ml) was used to capture prey (his-ABRAXAS) (50 μ g).
3. Resin was washed with same buffer and loaded on to SDS-PAGE. Complex was transferred to PVDF membrane (Millipore) and was probed with anti-His antibody (Abcam).

2.5 Protein oligomerization and size characterization

Oligomerization is a common property found among disordered and exposed hydrophobic patch containing proteins. Such protein may form several oligomeric species which may or may not have biological significance. An oligomer is usually referred to a macromolecular complex formed by mostly non-covalent bonding of macromolecules like proteins. Homo-oligomerization takes place when few identical molecules assemble together, whereas, hetero-oligomer involves three or more different macromolecules. Protein oligomerization can be detected and characterized using different methodologies. Each technique has its own advantages and disadvantages, and therefore confirmation should be drawn by combining several approaches. The number of techniques available are analytical ultracentrifugation, size exclusion chromatography, scattering techniques, NMR spectroscopy and mass spectrometry to detect and quantify oligomerization [138, 139] [140].

2.5.1 Glutaraldehyde Cross-Linking: Protein-Protein Interactions (oligomerization) can be transient or stable depending upon their role in several biological functions; e.g. biosynthetic or signal-transduction pathway [140] . Most often transient and semi-stable association of protein in a complex exists in signal transduction pathways because of highly complex protein interaction networks. Stable interactions can be identified using various methods such as ITC, however, chemical cross-linking provide a direct method for identifying both transient and stable interactions [141]. Cross-linking involves the formation of covalent bonds between two proteins by using bifunctional reagents such as glutaraldehyde containing reactive end groups that react with protein functional groups (primary amines and sulfhydryls of amino acid residues). Two physically interacting proteins can be covalently cross-linked and the formation of cross-link bridge between two proteins suggest their existence in a close proximity. It can also reveal the regions of interacting interface between two proteins.

2.5.2 Native Gel Electrophoresis: Native PAGE is one of the most powerful techniques for studying the composition and structure of proteins in its native form, since the conformation and the biological activity of proteins remain intact during this technique. It is occasionally used for molecular weight (MW) measurements; however, SDS-PAGE is easier and in most cases are reliable than native PAGE .However, it is difficult to find standard proteins that resemble the shape, partial specific volume and degree of hydration of the native protein.

Protocol for MERIT40 glutaraldehyde cross linking and native gel electrophoresis

1. Purified protein was concentrated and various dilutions of different concentrations 0.1, 1, 2.5, 5, 7.5, 10mg/ml were prepared.

2. Reaction mixtures with 20 µg of protein in 10 mM HEPES buffer (pH 7.5) were treated with freshly prepared solution of glutaraldehyde (final concentration 0.1%) for 2 minutes at, 37 °C [142].
3. This reaction was terminated by addition of 5 µl of 2 M Tris-HCl, pH 8.0. Cross-linked product was mixed with equal amount of Laemmli buffer and analyzed on 12% SDS-PAGE.
4. For native gel same set of protein dilution (as above) without adding glutaraldehyde and Tris was used for running on 7.5% basic-native gel (non-denaturing).
5. Protein samples were mixed with laemmli buffer without SDS and β-ME and run at low voltage in cold room.
6. In order to see disulfide bond specific multimerization, one sample was incubated with β-ME before loading on gel (coomassie staining).

2.5.3 Dynamic Light Scattering: Dynamic Light Scattering (Photon Correlation Spectroscopy) is one of the most popular methods used to determine the molecular dimension of bio-molecule (particle). When a monochromatic light beam (laser) falls on the molecules which always remain in brownian motion, it causes a Doppler Shift on collision, changing the wavelength of the incoming light. This change in wavelength related to the size of the particle, gives a description of the particle's motion in the medium by measuring the diffusion coefficient of the particle. Most of the "particle size analyzed" operate at 90° and use red light of wavelength 675 nm. Dynamic scattering is less accurate for distinguishing small oligomers and dimer from monomer than is classical light scattering or sedimentation velocity, however, its advantage is in analysis of samples containing broad distributions of species having widely differing molecular

masses, and to detect very small proportion of oligomer ($<0.01\%$). DLS is also capable in measurements of several parameters such as molecular weight, radius of gyration, translational diffusion constant etc.

Protocol for MERIT40:

1. Molecular size measurement was done using Brookhaven 90 plus particle size analyzer (Brookhaven, NY).
2. 1 mg/ml protein was scanned at an interval of 5 minutes for 15 minutes and effective diameter of each measurement was considered.

2.6 Crystallization

Three-dimensional structural analysis of a protein is highly crucial for understanding how a protein's function(s) is regulated. In a globular protein, polypeptide chain(s) fold to generate a compact structure whose function depends on the arrangement of polypeptide chain in three dimensional space, termed as tertiary structure. Proteins tend to crystallize when provided with appropriate conditions, and in order to achieve crystallization, purified protein undergoes slow precipitation from an aqueous solution. The individual protein molecules assemble and align on each other in the form of lattice held together by non covalent interactions, and made up of smallest "unit cells" by adopting a dependable orientation. Protein crystallography is commonly used to determine protein's three-dimensional structure via X-ray diffraction method.

X-ray diffraction data cannot be obtained unless a well-ordered crystal appears which provide a diffraction pattern when exposed to X-rays. The diffraction pattern can then be processed to derive the three-dimensional structure of the protein. However,, protein crystallization is essentially very tricky due to highly fragile nature of protein crystals.

Protein crystals are highly occupied with large channels due to irregularly shaped surfaces, therefore, the non-bonding interaction that grip the lattice usually are formed through solvent molecule layers. In order to grow a well formed crystal, sufficient homogeneity with purity of the protein is required at an optimal pH conditions, as different pH's can result in different crystal packing [143]. For crystallization, protein should be present in appropriate buffer system and precipitants.

Several factors affect crystallisation of proteins which includes: Sample purity, concentration, temperature, pH, ionic strength and volume of crystallization solution, etc. Process of crystallization is differentiated into two steps: Nucleation process and crystal growth

Nucleation and growth can occur within the supersaturated regions as depicted in phase diagrams (**Figure 2.1**). The diagram mainly comprises of three regions; unsaturated region, saturated region, and supersaturated region. Growth can occur in saturated or supersaturated region while nucleation most often start in supersaturated region. The major focus of crystallization process is to obtain sufficiently large crystals having dimension between 0.1-0.5 mm. However, it is an extremely challenging task to obtain a diffraction quality crystal. The best approach to get crystallization is through systematic exposure of the protein solution to varieties of buffers which have different combinations of precipitants, such as salts, poly ethylene glycols. For successful crystallization, the protein-precipitant mixture must approach nucleation phase very slowly to provide sufficient time for crystal growth. Most commonly used precipitants in crystallization trail are salts ($(\text{NH}_4)_2\text{SO}_4$, NaCl, KH_2PO_4), organic polymers (PEG) and alcohols (methanol, ethanol, propanol, acetonitrile). Salt usually maintain the ionic strength of the

solution while organic polymers reduce protein solubility by lowering the dielectric constant of solvent. Solubility of a protein is found lowest at its isoelectric point (pI) since protein carries a net zero charge.

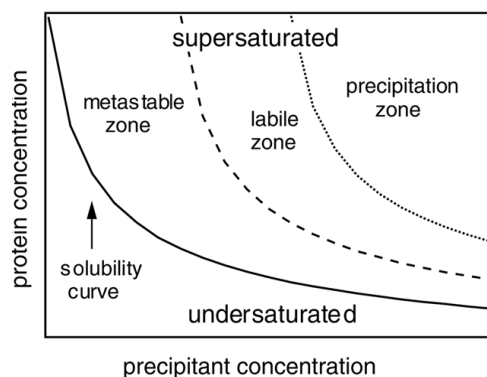


Figure 2.1: The phase diagram of crystallization.

2.6.1 Methods of crystallization: Different methods have been developed to crystallize functionally important proteins. Some of the more frequently used methods are vapor diffusion, micro-batch, dialysis and liquid–liquid diffusion method technique.

1) Vapour Diffusion method: Subcategories into hanging drop and sitting drop.

In both the methods a droplet of purified protein, buffer and precipitant is present in a closed system and allowed to equilibrate with a reservoir solution containing the same precipitant in higher concentrations. In the beginning, protein and precipitant concentration in the droplet is different, but as the system equilibrates the diffusion starts from higher concentration to lower concentration which in turn leads to nucleation, a level optimal for crystallization.

Vapor diffusion is a simple and powerful method since it consumes less protein, and several crystallization conditions can be explored with a limited amount of protein. It is also convenient for crystal growth monitoring, crystal manipulation and harvesting.

2) Microbatch: The problem of microcrystals can be overcome by using a low density paraffin oil (0.87 mg/ml) which float on the surface of protein-precipitant mixture, thus reduces the rate of evaporation [144]. In this method, a mixture of paraffin and silicon oil can also be used. This is an excellent method for refining the known crystallization condition and optimization of crystal growth [145].

3) Dialysis: The dialysis method employs slow diffusion of inorganic molecules through a semi permeable membrane which results in the formation of supersaturated state of protein solution. The system tries to maintain equilibration between precipitant and solute molecules by allowing selective passage of water and precipitant. The dialysis bag is kept in a chamber containing the reservoir solution with a precipitant [146].

4) Liquid-liquid diffusion method: In this method, protein and precipitant solution maintain direct contact keeping the dense solution at the bottom. A concentration gradient is generated due to the diffusion of protein and precipitant and crystallization may occur at appropriate protein and precipitant concentration. Free interface diffusion is a method of choice for fine-tuning the crystallization conditions.

2.6.2 X-Ray Diffraction: In an X-ray diffraction method, a crystal is exposed to X-ray beam, and diffractions are measured over a suitable detector. There are two kinds of X-ray sources used in protein crystallography, the conventional laboratory X-ray generator (home source) and the synchrotron radiation source. The high energy X-rays source generates the X-ray of wavelength 0.3-2 Å and used in the determination of three dimensional structures of proteins. The crystal is mounted usually in a loop which is attached with goniometer head. The goniometer allows proper centering of the crystal in

the X-ray beam so that during rotation crystal remain focused to X-ray beam. Commonly used detectors includes Image Plate, CCD and Pixel Array Detectors.

The electrons present on the periodically arranged macromolecules inside the crystal interact with the X-rays during exposure and diffract X-ray in certain directions. Diffracted rays are captured on the detector as discrete spots called as reflections. The directions of these reflections are given by Bragg's law which states that a reflection is produced only when difference in the path length for rays reflected from successive planes is equal to an integral number of wavelength of the incident X-rays.

$$2d_{hkl} \sin\theta_{hkl} = n\lambda$$

where λ is the wavelength, θ_{hkl} is the angle of incidence or reflection, n is an integer and d is the distance between the parallel lattice planes with Miller indices h,k,l . hkl are also the indices for the Bragg reflections. All the Bragg reflections contain information about any given atom in the three dimension structure, and every atom in the structure contributes to intensity of all the Bragg reflections. A Bragg reflection consist sets of parallel planes series and diffracted ray, as reflections from these atomic planes generate constructive interference. In all other cases, reflected waves from successive planes of the crystal become out of phase due to destructive interference and no diffracted beam emerges from the crystal. Intensities of several Bragg diffraction spots need to be collected to solve a protein three dimension structure. This can be achieved by using the rotation method, where the crystal is rotated by small angles and a series of diffraction images generated in which each image contains several reflections collected at a defined angles.

Diffraction data of mounted crystal can be collected either at room temperature or at ultra low temperature (100K) with the crystal bathed in a jet of liquid nitrogen. Room temperature data is generally collected to see the diffraction pattern but due to radiation damage mounted crystal become highly prone for damage. Data collection at cryogenic temperature significantly overcome this problem and X-ray induced radiation damage become minimal. To prevent the formation of crystalline ice and crystal dis-orderness, cryo-protectant can be applied at the surface of crystal by flash cooling the crystal in liquid nitrogen . Glycerol is the most commonly used cryo-protectant followed by PEG and other organic substances.

In the present work, X-ray diffraction data was collected at 100K at Home source facility ACTREC. Crystals were equilibrated briefly in cryoprotectant solution of 30% glycerol. A total of 227 frames were collected each for an oscillation angle of 1° and exposure time of 10 min.

Conditions: concentration: 10-25 mg/ml, Temperature: 22 °C, Crystal screen: Hampton
Protocol:

1. Protein of concentration 10-25 mg/ml is spinned in a refrigerated centrifuge at 13000rpm, for 10 minutes.
2. This protein is then transferred into a fresh eppendorf and neatly labelled.
3. Add around 1ml of the buffer in each well.
4. Take cover slips and put them in a row.
5. Pipette out 1ul of protein in each of the cover slips.
6. Carefully invert the coverslips and place them on the well and ensure that they are properly sealed.

Put the tray in a temperature controlled environment where it won't be disturbed and the tray is observed under the microscope at regular intervals.

3.0 Domain organization of expression constructs

RAP80: RAP80 (1-130) wild-type, RAP80 (1-130) Δ E81 and RAP80 (1-405) functional domains of RAP80 were sub-cloned in pGEX-kT vector (Amersham Pharmacia). RAP80 (1-130) comprising the N-terminal motif and UIMs (Ubiquitin Interacting Motifs) which binds to Ubiquitin at the site of DNA damage. RAP80 (1-405) comprises the UIMs and Abraxas Interacting Region (AIR). RAP80 (1-130) wild-type and RAP80 (1-405) were expressed in bacterial system *E.coli* BL21 (DE3) cells.

MERIT40: MERIT40 full length was sub-cloned into bacterial expression vector pGEX-kT using one step cloning method as described earlier.

ABRAXAS: ABRAXAS (6-268), ABRAXAS (6-301), ABRAXAS (6-373), ABRAXAS (6-409) regions were subcloned into pET28(a) vector. ABRAXAS (6-268) mainly comprises N-terminal region while ABRAXAS (6-301) and ABRAXAS (6-373) includes RAP80 interacting region.

BRCA1-BRCT: BRCA1 BRCT (1646-1859) was subcloned into pGEX-kT vector. This region is essential for binding to ABRAXAS directly in phosphodependent manner. The pGEX-kT BRCA1 BRCT (1646-1859) construct was transformed into bacterial expression host, *E. coli* BL21 (DE3) strain, and the fusion protein was expressed.

Chapter 3

*Structural Basis & Protein-Protein
Interaction Studies of RAP80*

3.1 Introduction

Cells have evolved a complex, dynamic and highly regulated network, called DNA damage response (DDR), to achieve extreme fidelity. In genotoxic stress, DDR coordinates numerous cellular processes like cell-cycle regulation, chromatin-remodeling, DNA-repair and transcriptions [20]. Sensing the DNA damage and promulgation of the DDR signaling cascade involve recruitment and assembly of many DDR mediators and effectors at the sites of damage [35, 147]. Double strand breaks elicit the activation of ATM and ATR kinases, which phosphorylate histone variants H2AX and MDC1 [17, 35, 86, 87, 148-150]. This event endorses the assembly of DDR mediators, which in turn facilitates the recruitment of UBC13/RNF8 to the DNA damage site [88, 151-153]. In the signaling pathways, this leads to the formation of polyubiquitin chains on H2AX, which are recognized by RAP80 (**Figure 3.1**) [17, 87, 149, 150]. RAP80 has two tandem UIM (Ubiquitin-Interacting Motif) at its N-terminal, ABRAXAS (CCDC98) Interacting Region (AIR) at the central domain, and two zinc -finger domains at its C-terminal [108]. It has been reported that RAP80 forms a stable complex with BRCA1 through an intermediate binding partner CCDC98 [6, 7, 16]. CCDC98 has a consensus sequence S-X-X-F motif at its C-terminus, which is involved in interactions with phospho-specific binding domain of BRCA1-BRCT [2, 6-8].

RAP80 acts upstream of CCDC98 and BRCA1 in DDR, and is required for the localization of the BRCA1 complex to ionizing radiation (IR)-induced foci (IRIFs)[6, 16, 112]. RAP80 knockdown cells have shown hypersensitivity to IR and ultraviolet (UV) light, cell -cycle dysfunction and defective homologous recombination (HR) repair [6, 7, 16, 87]. RAP80 and p53 auto-regulate each other and have influence on apoptosis [154].

Loss of RAP80 alleles (RAP80^{-/-}) increase the susceptibility to lymphoma, and promote tumor development in both p53^{-/-} and p53^{+/-} mice [155]. UIM1 and UIM2 motifs of RAP80 are very crucial since deletion of either or both motifs significantly perturb the foci formation of RAP80-BRCA1 complex at the DNA damage site [5].

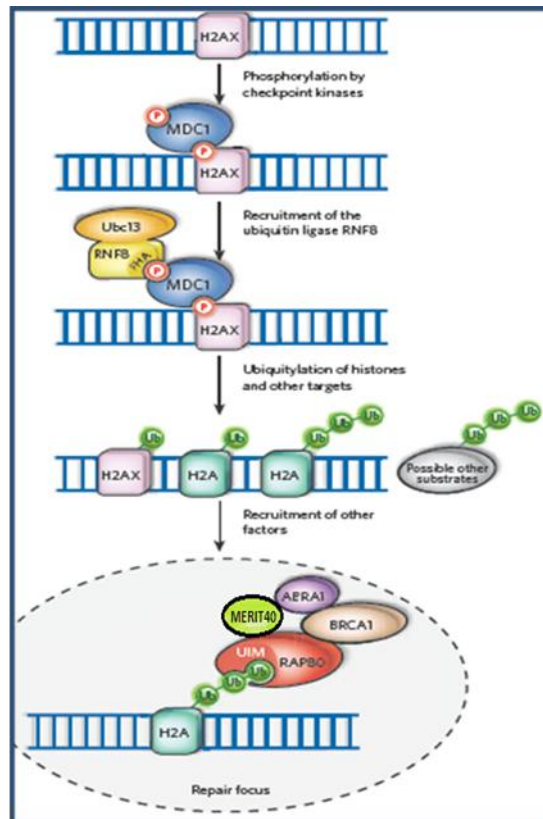


Figure 3.1: DNA damage response in repair pathway.

A novel alteration, c.241-243delGAA (Δ E81), which leads to an in-frame deletion of glutamic acid residue, has been identified at UIM1 motif of RAP80 [5]. RAP80 Δ E81 variant has been discovered in patient diagnosed with breast cancer, and the residue is found to be highly conserved among all the vertebrates. This variant showed an observed frequency of 0.9% (1/112) in the familial cases compared to 0.3% (1/325) in the controls ($P=0.45$; $OR=2.92$; $CI=0.18-47.1$). Another RAP80 Δ E81 carrier was also diagnosed with bilateral breast cancer in a group of 503 breast cancer cases (0.2%, 1/503) [5].

Furthermore, RAP80 Δ E81 expressing cells have shown abrogation of DSB localization of the RAP80–BRCA1 complex, and exhibited genomic instability (chromosomal aberration) [5].

A comparative structural stability and binding analysis of RAP80 (1-130) wild-type and RAP80 (1-130) Δ E81 was carried out to understand the functional implication(s) of this mutation. Our main goal is to analyze the structure of wild-type and functionally important mutants of RAP80 and its binding partners. The aim here is to use biophysical, biochemical and structural biology tools to disentangle the convolution of RAP80–BRCA1 complex.

3.2 Material and Methods

A comparative structural, stability and binding analysis of RAP80 (1-130) wild-type and RAP80 (1-130) Δ E81 was carried out. Different functional domains of RAP80 were sub-cloned in pGEX-kT vector (Amersham Pharmacia) as shown in figure 3.2. RAP80 (1-130) wild-type, RAP80 (1-130) Δ E81 and RAP80 (1-405) were sub-cloned from the mammalian vector (pEGFP-c1, a kind gift from J. Chen, USA) to the bacterial expression vector using one step cloning method. The basic scheme for sub-cloning, and nucleotide sequence of primers used in PCR reaction are shown in **Figure 3.3**. The gene of interest was amplified using following PCR condition: 95 °C denaturation (5 minutes), 95 °C denaturation (45 seconds), annealing 65°C for 35 seconds, extension 72°C at 0.5kb/min, final extension 72°C for 10 minutes, 25 cycles. The PCR amplified product was digested with EcoR1 and BamH1 and ligated in the pGEX-kT vector (a kind gift from J. Ladas laboratory). The ligation mixture was transformed into *E.coli* DH5 α cells. Colony screening was done by insert release from the plasmid using EcoR1 and BamH1

restriction enzymes and further confirmed by DNA sequencing. Positive clones were used for protein expression and purification.

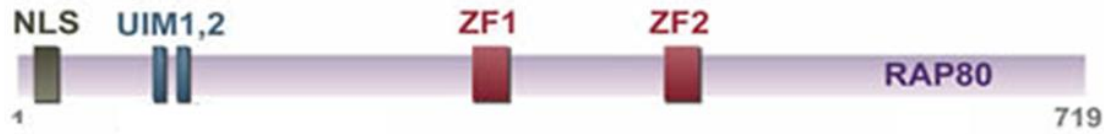


Figure 3.2: Different functional domains of RAP80

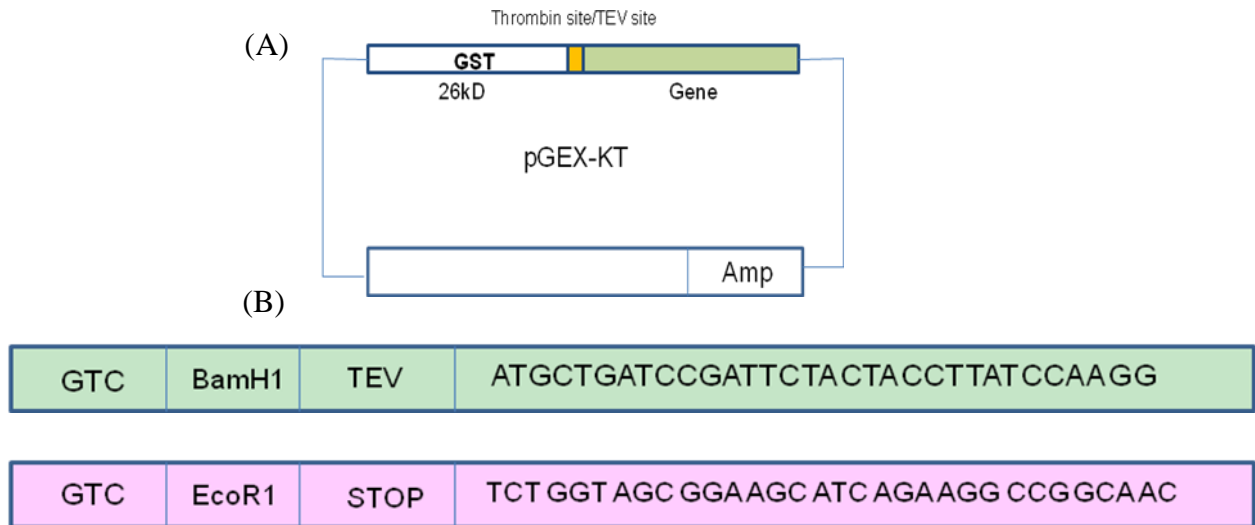


Figure 3.3: Basic scheme of sub-cloning of different functional domain of RAP80 in pGEX-kT. (A) Vector construct. (B) RAP80 primers used during PCR amplification.

RAP80 (1-130) wild- type and RAP80 (1-405) were expressed in bacterial system *E.coli* BL21 (DE3) cells. For protein expression, 100 µg/µl plasmid construct was transformed into *E.coli* BL21 (DE3) cells and grown on LB agar plate containing antibiotic ampicillin (100 µg/ml). Detailed protocol is described below.

Protocol for protein expression:

- Inoculation:** Pick-up a single transformed colony from antibiotic resistant LB agar plate and inoculate it in 100 ml LB broth containing 100 µg/ml of ampicillin. Incubate on shaker incubator at 37°C overnight.

2. **Dilution:** Inoculate 10 ml of starting culture to 1000 ml (1: 100 ratios) of autoclaved LB broth containing 100 µg/ml of ampicillin. Incubate the flasks on shaker incubator at 37 °C until it has reached mid-log phase i.e. A_{600} between 0.6-0.8.
3. **Induction:** Cool down the flasks and add 400 µl IPTG (stock 1M) and incubate on shaker incubator at 24°C for 16 hours.
4. **Harvesting:** The culture was transferred to centrifuge bottles and centrifuged for 10 minutes at 6000 rpm at 4°C. The pellet was resuspended in a small amount of supernatant and centrifuged for 15 minutes at 5000 rpm, 4 °C.
5. **Storage:** The pellet obtained can be stored at -80 °C for further use.

Proteins were purified by affinity chromatography followed by FPLC. RAP80 (1-130) Δ E81 mutant was expressed in *E.coli* (BL21) and purified as per wild- type protocol.

Protocol for purification of RAP 80 and Δ E81:

Purification buffer composition: 10 mM HEPES (pH 7.5), 300 mM NaCl, 0.1 mM EDTA, 2 mM β -ME, 5% ethylene glycol pH 7.5

FPLC buffer composition: 10 mM HEPES (pH 7.5), 300 mM NaCl, 0.1 mM EDTA, 2 mM BME pH 7.5

1. **Re-suspension:** Re-suspend the pellet of RAP80 in 40 ml of purification buffer; supplemented with 200 mM PMSF and 20µl of protease inhibitor.
2. **Ultra sonication:** Transfer the suspension into centrifuge tube and sonicate at 50 pulse rate and 50 power with 1.45 minutes of duty cycle.
3. **Centrifugation:** After sonication, the suspension was subjected to centrifugation at 18000 rpm for 45 minutes at 4°C to obtain cleared lysate. Collect the supernatant and discard.

4. **Binding:** The soluble fraction obtained is brought to room temperature and then mixed with pre-equilibrated affinity resin and incubated at room temperature for 1 hour. The sepharose beads in affinity column charged with glutathione binds specifically to those proteins having GST (Glutathione-S-transferase) tag.
5. **Washing:** After binding, the column is given 4 column washes with wash buffer so as to remove any non specific protein.
6. **Cleavage:** Add 400 μ l of TEV protease, 40 μ l of protease inhibitor cocktail and 100 μ l of PMSF in 20 ml of purification buffer and perform the cleavage step for 3 hours by passing the solution through column at interval of 1 hour. Take out 40 μ l of beads to observe the cleavage of protein.
7. **Elution:** After TEV cleavage the protein is eluted with 30 ml of purification buffer.
8. **Equilibration of Ni-NTA resin:** Give 2 column washes with double distilled water and then 5 to 6 column washes with wash Buffer.
9. **Metal Ion Chelate Affinity Chromatography:** After calibration of Ni-NTA resin, pass the eluted fractions through them to get rid of His-tagged TEV protease contaminant.
10. **Concentrating the protein:** Transfer the eluted protein in a 10 kDa pre-equilibrated centricon and concentrate the protein upto 2 ml by centrifuging it at 4500 rpm for 10 minutes at 4 °C. Check the concentration on Nanodrop spectrophotometer (280 nm). Centrifuge for 10 minutes at 13000 rpm at 4 °C for removal of insoluble aggregates or precipitates.
11. **Gel filtration:** Inject 2 ml of concentrated protein in AKTA- FPLC against FPLC buffer.

12. **Fraction collection:** Collect the purified protein obtained through FPLC in 1.7 ml eppendorf at its elution volume according to gel filtration spectra profile of the sample.

13. **Loading on SDS-PAGE 12% gel:** Load 20 µl of FPLC fractions on SDS-PAGE, stain with commassie dye, and then destain it to visualize the protein of interest.

14. **Concentrate the protein:** The fractions which showed purified protein band was further concentrated as per the requirement.

The purified proteins were used in various bio-physicochemical experiments. In addition, binding analysis of RAP80 (1-405) was performed with its binding partner. The complete details of protocol have been discussed in chapter 3 (Material and Methods).

3.3 Results and discussion

3.3.1 Cloning, expression and purification of RAP80 functional domain: Selected potential clones when digested with the EcoR1 and BamH1 restriction enzymes showed the insert release of appropriate size (**Figure 3.4**). Sequencing results confirmed the presence of ligated gene of interest in the vector with desired frame of codon sequence.

RAP80 (1-130) wild-type and mutant eluted at the same elution volume (98 ml) corresponding to monomer (**Figure 3.5**). This indicates that RAP80 Δ E81 mutation does not change its oligomeric property. However, the elution profile of RAP80 (1-405) showed that most of the protein is oligomeric in nature and forms aggregates. Very minuscule amount of RAP80 (1-405) protein was observed at proper elution volume (corresponding to monomer).

3.3.2 Structural insight into RAP80-Ub complex using *in-silico* analysis: In order to analyze the structural changes due to RAP80 mutation (Δ E81), RAP80 (80-120) Δ E81

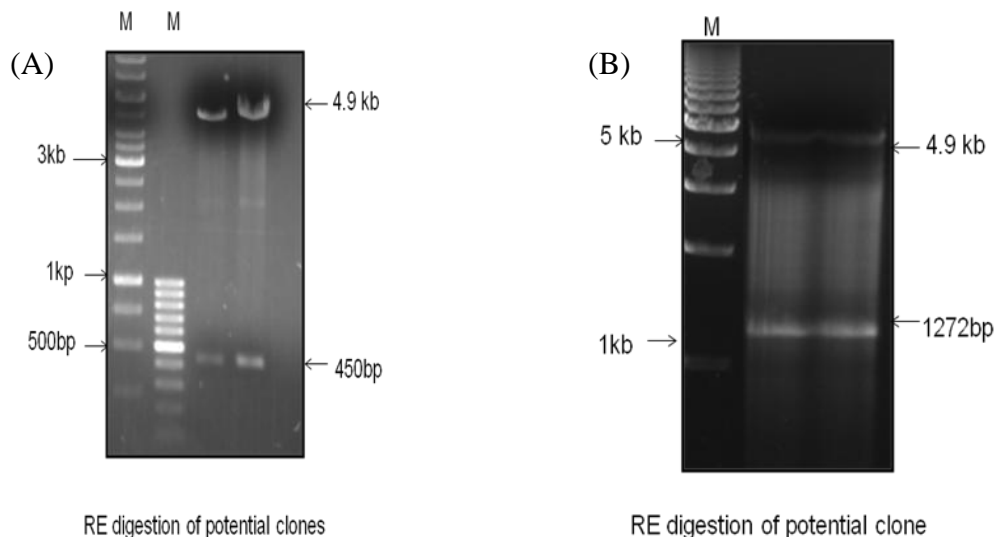


Figure 3.4: Agarose gel showing the insert release from potential clones' plasmids; (A) pGEX-KT-rap80 (1-130), (B) pGEX-KT-rap80 (1-405).

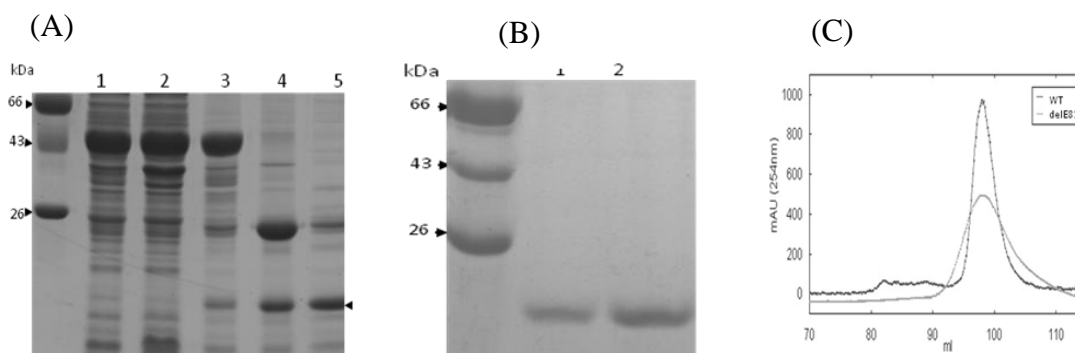


Figure 3.5: Expression and purification profile of RAP80 (1-130). (A) Lane 1-Total protein, 2-soluble protein, 3-fusion protein bound on beads, 4- protein after on beads cleavage and 5-elution fraction of affinity purified protein. Single arrow - RAP80 (1-130) protein, (B) Purified protein after gel filtration chromatography on SDS-PAGE. Lane 1- RAP80 (1-130) delE81, 2- RAP80 (1-130) wild-type (C)-Overlay of gel filtration spectra of RAP80 wild-type and $\Delta E81$ (Superdex 200) [156].

functional motif (called as UIM) was modeled using homology modeling server (Swiss Modeller, www.swissmodel.expasy.org/) considering, RAP80-Ub NMR structure (PDB ID: 2RR9) as the template [12]. A model was selected based on good stereo chemistry and better Ramachandran plot. The model was further validated using server “SAVES” (Metaserver for analyzing and validating protein structures, <http://nihserver.mbi.ucla.edu/SAVES/>).

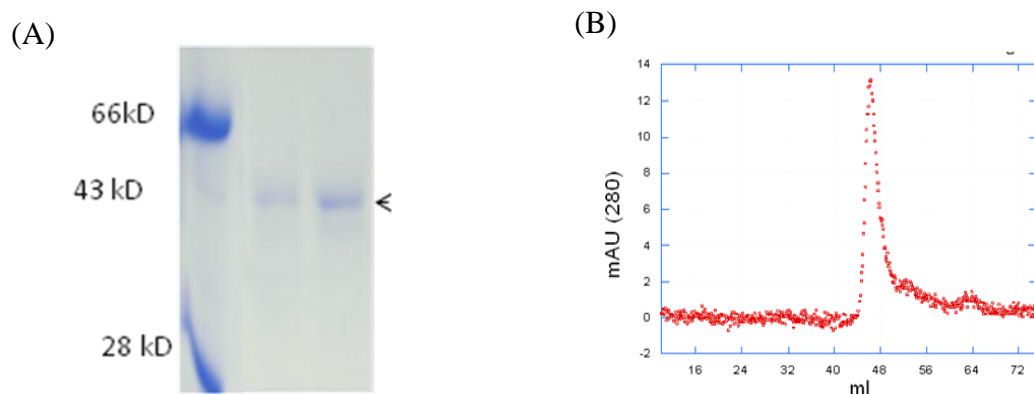


Figure 3.6: Purification of RAP80 (1-405). (A) Purified protein represented by single arrow head. (B) The gel filtration spectra of the same.

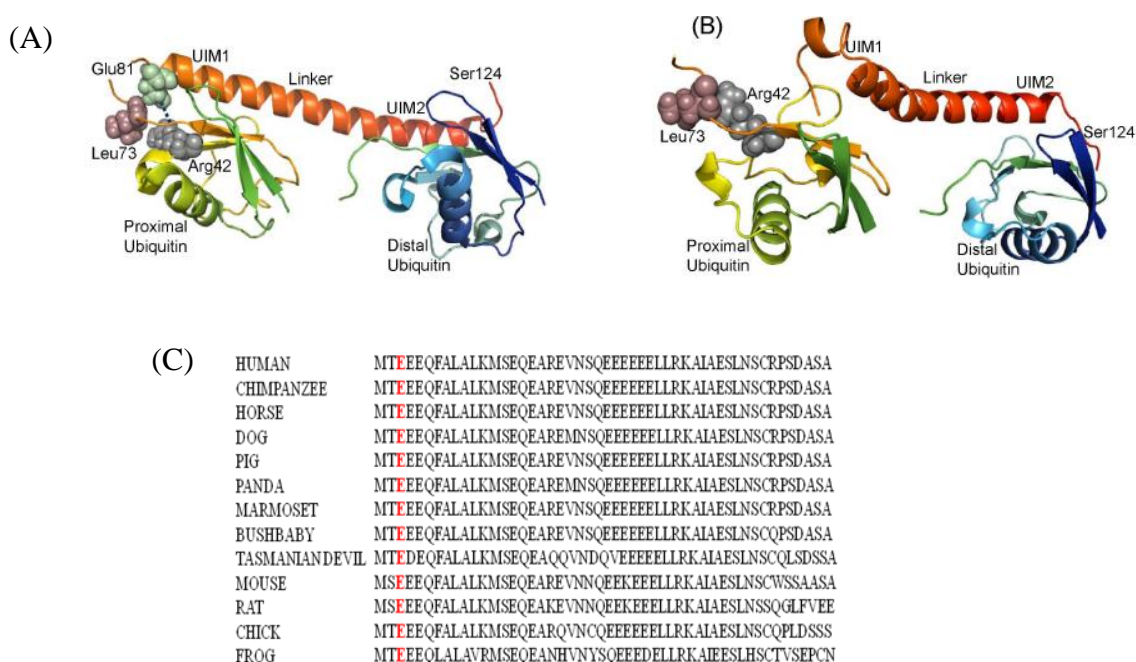


Figure 3.7: Binding interaction of RAP80 UIMs and Δ E81 with Di-Ub (K-63 linked). (A) Structure of Di-Ub (K-63 linked)-RAP80 UIMs (79-124) wild-type (PDB ID: 2RR9), and (B) Di-Ub (K-63 linked)-RAP80 (79-124) UIMs Δ E81. (C) Multiple sequence alignment of UIMs region showing Glu 81 residue highlighted in red color [156].

The server SAVES mainly comprises of five programs, Procheck, What_check, Errat, Verify_3D and Prove. The modeled structure was simulated for 5 ns using Desmond software (Schrodinger) and superimposed on wild-type complex. PDBsum software was used to analyze the interactions [157].

Modeled structure revealed that UIM1 and UIM2 motifs are connected with a linker in a head-to- tail manner. The three-dimensional structure of the wild-type was predominantly α -helical in nature and approximately 59 Å long . Strikingly, in case of Δ E81 mutant, α -helix is partly distorted and shortened to 45 Å. UIM1 and UIM2 bind to their respective proximal and distal ubiquitin of Di-Ub (K-63 linked) in 1:1 affinity ratio [158, 159]. Glu residue at 81 position was found to be highly conserved (**Figure 3.7**) and forms ionic bond and hydrophobic interaction with Arg42 and Leu73 residue of proximal ubiquitin, respectively. However, a drastic conformational change in RAP80 UIMs Δ E81 was found which significantly altered the intermolecular interactions with ubiquitin. Structural distortion in RAP80 UIMs Δ E81 probably renders its binding interactions unfavorable with Di-Ub (K-63 linked).

3.3.3 Structural insights into RAP80-Ub complex using *in-vitro* analysis: In order to determine the domain stability of RAP80 (1-130) wild-type and Δ E81 against the protease digestion, limited proteolysis was performed using trypsin and chymotrypsin. After treating the wild-type and mutant for limited time period with equal enzyme concentration, it has been observed that domain stability of RAP80 (1-130) wild-type is significantly higher than that of Δ E81. This imply that wild-type exists in a well formed structure which made it resistive towards protease digestion (**Figure 3.8**). However, high susceptibility of Δ E81 towards protease digestion indicates its probable deformed structure.

Susceptibility towards limited proteolysis of RAP80 (1-130) Δ E81 impart the supposition of deformities in the native structure. Hence, the secondary structural components with far-UV Circular Dichroism were investigated (**Figure 3.9**). RAP80 (1-

130) wild-type showed an estimated percentage of α -helices, β -sheets, turns and random structure as 45, 13, 16, and 26 % respectively, while for Δ E81 the corresponding values were 5.1, 4.7, 3.2 and 88 %. It suggests that secondary structural component of RAP80 is severely altered due to mutation, resulting in the loss of α -helical characteristic. β -sheet proportion was also moderately altered, and the mutant showed majority of random coil structure. Earlier report suggests that UIMs motif of RAP80 is found in equilibrium between α -helix and random structure [160]. Δ E81 mutation probably alters the α -helical conformation of RAP80 UIMs which leads to a shift in the equilibrium towards a random structure pattern.

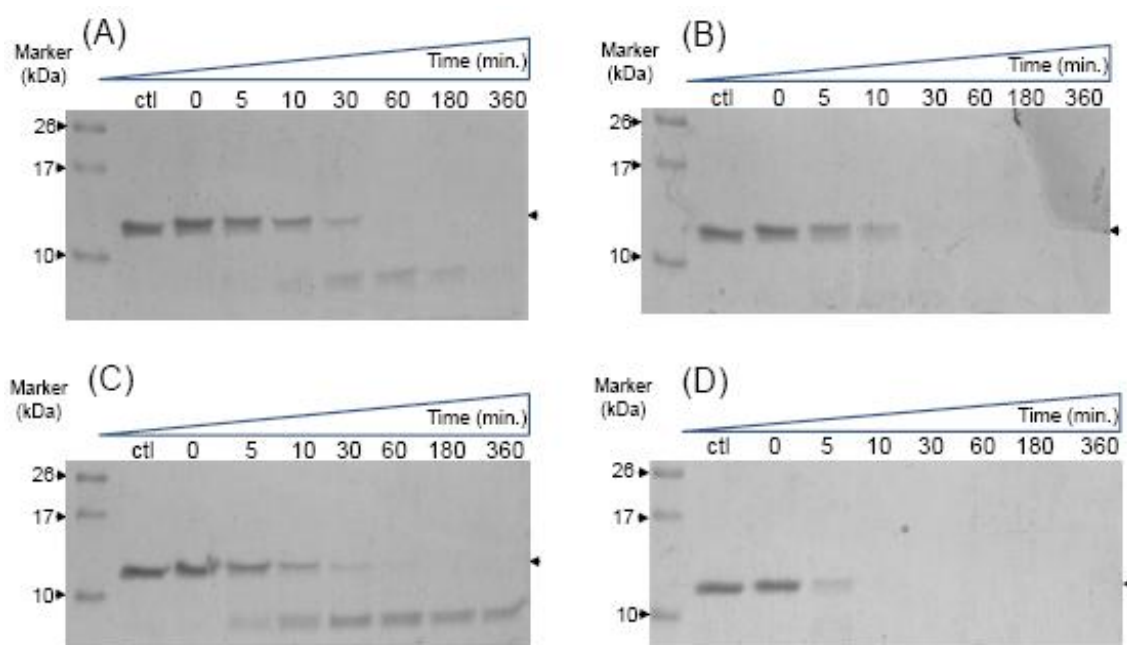


Figure 3.8: Resistivity profiles of RAP80 wild-type and Δ E81 towards protease digestion. Limited proteolysis of RAP80 wild-type (A, C) and Δ E81(B, D) using trypsin (A, B) and Chymotrypsin (C, D) as proteases. Ctl- control was taken as untreated with proteases [156].

Since limited proteolysis and Far-UV spectra deciphered changes in structural organization of Δ E81, it would be interesting to explore the domain stability and three dimension folding profile of wild-type and the mutant. Stability profiles of RAP80 wild-

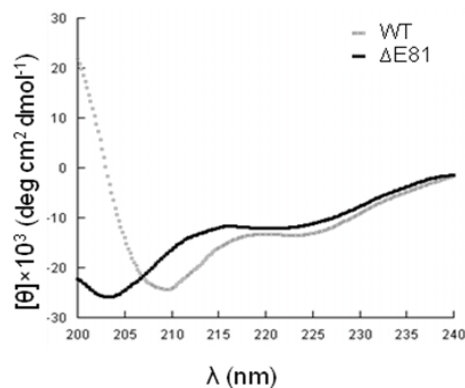


Figure 3.9: Comparison of secondary structural components of RAP80 (1-130). Overlay of Far-UV Circular Dichroism spectrum of wild-type and $\Delta E81$.

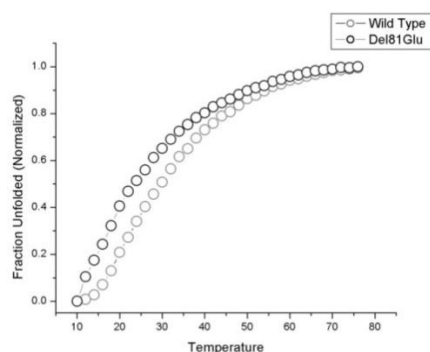


Figure 3.10: Thermal stability of RAP80 (1-130). Thermal denaturation of RAP80 (1-130) wild-type and $\Delta E81$ in Circular Dichroism [156].

type and $\Delta E81$ were compared at secondary (CD) and tertiary (DSC) structure levels. The spectra obtained from Circular Dichroism corresponding to λ at 218 nm showed the maximum changes in ellipticity and high signal to noise ratio (**Figure 3.10**). Thermal stability of RAP80 $\Delta E81$ (T_m 22°C, $\Delta G^\circ_{H_2O}$ 1.3 \pm 0.2 Kcal/mol, ΔH 1.0 \pm 0.5 Kcal/mol) was found significantly low compared to wild -type (T_m 29°C, $\Delta G^\circ_{H_2O}$ 2.0 \pm 0.5 Kcal/mol, ΔH 5.0 \pm 2.0 Kcal/mol) (**Figure 3.10**). Thus, protein most likely unfolds without any intermediate species.

These findings were further supported by Differential Scanning Calorimetry, which gave a T_m value of 28°C for RAP80 wild-type (**Figure 3.11**). However, a defined transition

for $\Delta E81$ could not be obtained due to lesser stability and saturation concentration. Collectively, these results suggest that three-dimensional folding of RAP80 $\Delta E81$ is impaired which may be due to helix to random structure transition of UIMs. $\Delta E81$ mutation probably shifts this transition equilibrium towards the random structure.

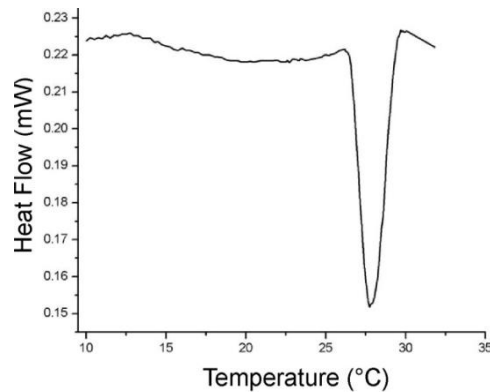


Figure 3.11: Differential Scanning Calorimetry profile of RAP80 wild-type [156].

3.3.4 Binding interaction of RAP80 wild-type and $\Delta E81$ with di-Ub (K-63 linked):

It is well reported that RAP80 UIMs bind to K-63 linked polyubiquitin chain(s) on the H2AX and recruit the RAP80-BRCA1 complex at the DNA damage site [6, 112]. Considering alterations in folding pattern & domain stability of RAP80 $\Delta E81$, it can be suspected that it would further impair the binding affinity to the polyubiquitin chain. Binding analysis between RAP80 wild-type and $\Delta E81$ with Di-Ub (K-63 linked) was performed using Surface Plasma Resonance (SPR) and GST pull-down assay. The observed binding affinity for RAP80 $\Delta E81$ (K_D : 0.459 μ M) was several fold less as compared to wild-type (K_D : 36.5 nM) in SPR (**Figure 3.12**).

GST pull-down assay also corroborated with the findings obtained using SPR (**Figure 3.12**). Collectively, it can be concluded that RAP80 wild-type has higher binding affinity for the polyubiquitin chain, besides, it associates faster than $\Delta E81$. Mutant protein

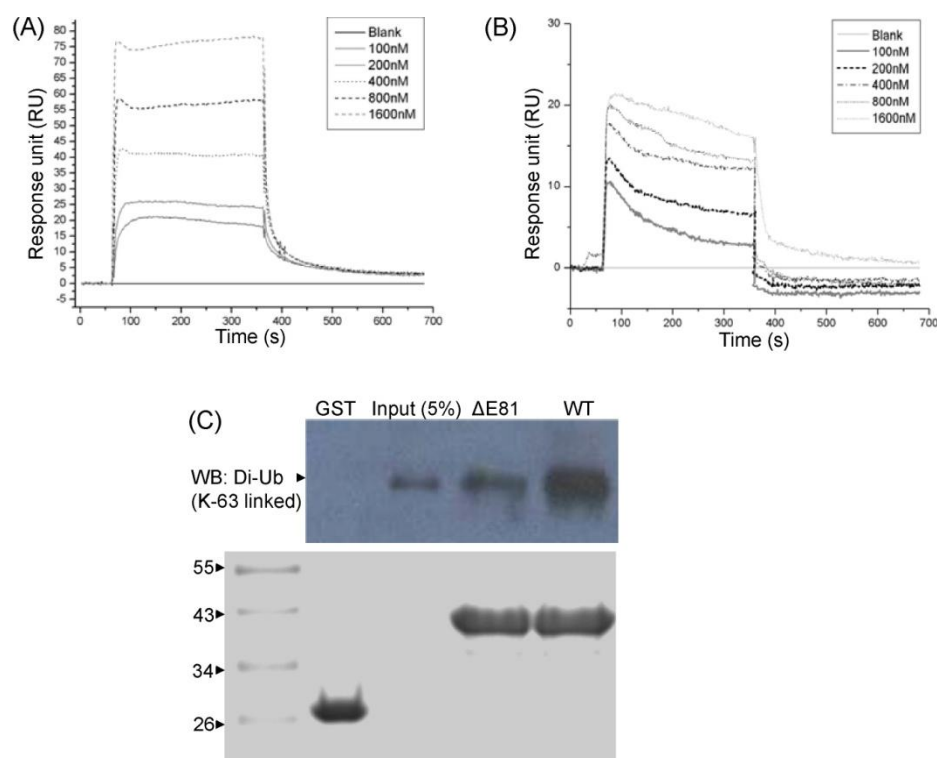


Figure 3.12: Binding analysis of RAP80 wild-type and $\Delta E81$ with Di-Ub (K-63 linked). Sensograms of RAP80 wild-type (A) and $\Delta E81$ (B) determined by Surface Plasma Resonance. (C) GST pull-down assay followed by western blotting. GST-RAP80 wild-type and $\Delta E81$ were used as a bait and Di-Ub (K-63 linked) as prey [156].

complex $\Delta E81$ -Di (Ub) is not stable due to high dissociation rate and lower binding affinity, which may be due to its structural deformation.

3.4 Crystallization of functional domain of RAP80

Protein crystallization was performed using commercially available reagents from Hampton Crystal Screen 1, 2 and SaltX screen. We tried both the hanging and sitting drop vapor diffusion methods. Initial crystallization trials for RAP80 (1-130) protein was set with 5 mg/ml protein concentration. The protein and mother liquor solution was mixed in 1:1 ratio (1 μ l+1 μ l) and allowed to equilibrate at 22°C with 500 μ l reservoir solution. Unfortunately, RAP80 (1-130) protein could not be concentrated more than 5 mg/ml due to high precipitation. A clear drop or light precipitation was observed in most

of the drops. Some crystallization conditions were selected for optimization on the basis of crystal conditions observed but all the effort for crystallization were futile. Crystal trial for mutant protein could not be performed due to its lower stability (0.2 mg/ml).

3.5 Conclusion

RAP80 wild-type and $\Delta E81$ are moderately soluble proteins with a molecular weight of about 14 kDa. Thermal and proteolytic stability of wild -type was found to be significantly higher as compared to $\Delta E81$, but both unfold via two-state transition. RAP80 UIMs are found in an equilibrium between random-coil and helical structure. Dynamic nature of UIMs provides immense flexibility for dissociation and association of ubiquitin molecules during the protein trafficking process.

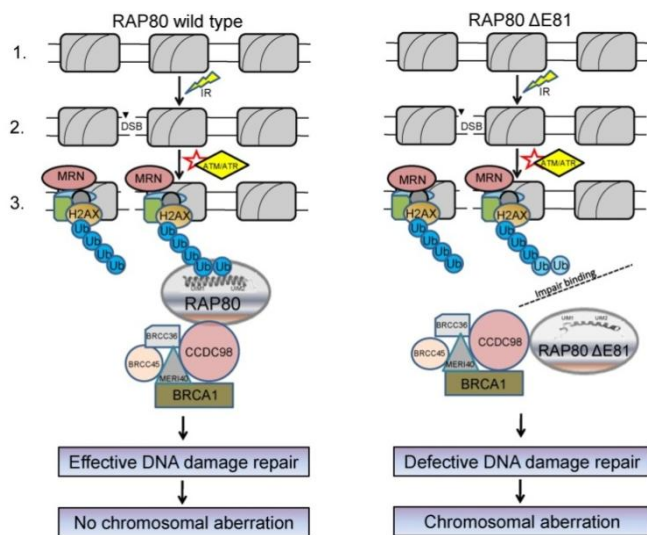


Figure 3.13: Mechanism of consequence for RAP80 wild type and $\Delta E81$. The model elucidates mechanisms of chromosomal aberration due to RAP80 $\Delta E81$ mutation. In the wild-type RAP80: **Step1**, showed the intact nucleosome complex; **Step 2**, double strand break due to ionization radiation; **Step 3**, ATM/ATR kinase activation and assembly of various damage repair proteins at DNA double strand break (DSB) site followed by formation of polyubiquitin chain(s) on histone(s) (H2AX) [156].

This dynamic nature is essential for a flexible and transient initiation mechanism of the DNA damage repair process. Deletion of 81E residue alters the helical conformation, thus shifting the equilibrium towards a random structure. Helical to random structure

transition results in the loss of weak intermolecular hydrogen bonds and hydrophobic interactions between the UIMs and Di-Ub (K-63 linked), thereby making the binding interactions unfavorable for ubiquitin. Since binding affinity of individual UIM for mono-ubiquitin is low [161], an avidity-based mechanism may play a very important role in the interaction between RAP80 and Lys 63-linked polyubiquitin. Co-operative binding between multiple UIMs and ubiquitin chains likely occurs, which may favor the interaction of second UIM with ubiquitin after positioning of the first [159]. It has been reported [5] that expression of RAP80 Δ E81 allele abates the recruitment of BRCA1 complex at DSB site, which further augments chromosomal aberration (chromatic breaks). Deletion of 81 glutamic acid residue significantly obliterates RAP80 structure and impairs its binding with the polyubiquitin chain. The unstable nature of mutant and di-ubiquitin complex may be responsible for defective recruitment of RAP80-BRCA1 complex to the DNA damage sites. Defective DNA damage repair perhaps leads to chromosomal aberration as shown in the model (**Figure 3.13**). Comparison of RAP80 Δ E81 with wild-type will help in understanding its role in various diseases and repair defects. It will further explore the possibility of structure based inhibitor design for therapeutic application that can compensate the effect of such mutations.

Chapter 4

Biophysical & Structural Characterization of MERIT40

4.1 Introduction

BRCA1 interacts with numerous molecules and consequently plays an essential role in multiple pathways such as the DNA damage repair, cell-cycle regulation and transcription regulation [13, 19, 20][2, 8, 89, 162, 163]. It acts as a tumor suppressor and has been found to be associated with hereditary breast and ovarian cancers [3, 164]. Furthermore, BRCA1 BRCT deletion mutant showed perturbed behavior towards sub-nuclear co-localization with H2AX, thus highlighting its importance in DNA damage repair [85, 116, 147, 165].



Figure 4.1: Schematic representation for structure of MERIT40. Willebrand Factor A (vWFA) motifs are named for the prototype member of the family, von Willebrand factor A, a platelet glycoprotein with critical roles in blood clotting.

MERIT40 (Mediator of RAP80 Interaction and Targeting 40) is an essential molecule of the BRCA1 complex and helps to maintain its protein complex integrity (**Figure 4.1**). Knockdown of MERIT40 significantly reduces the RAP80 and ABRAXAS protein levels, consequently affecting the integrity of BRCA1 complex [13, 19]. MERIT40 downregulated cells shown lower the stability of BRCA1 complex compared to any other member of this complex [13]. To elucidate its role in BRCA1 complex, structure of MERIT40 and its correlation with functional properties were analyzed.

The aim here is to use the biophysical, biochemical, *in-silico* and structural biology tools to dissect the structure aspect of MERIT40. We have performed cloning, expression and purification of MERIT40 and bio-physicochemical characterization using multi-model approach.

4.2 Material and Methods

In the present study, a structural stability, protein folding and thermodynamic analysis of MERIT40 was carried out. MERIT40 was sub-cloned into bacterial expression vector pGEX-kT using one step cloning method, as described in chapter 3 (**Figure 3.3**). The nucleotide sequence of primers used in PCR reaction for cloning are shown in **Figure 4.2**. The gene of interest was amplified using following PCR condition: 95 °C denaturation (5 minutes), 95 °C denaturation (45 seconds), annealing 65°C for 35 seconds, Extension 72°C at 0.5kb/min, final extension 72°C for 10 minutes, 25 cycles. The PCR amplified product was digested with EcoR1 and BamH1 and ligated in the pGEX-kT vector. The ligation mixture was transformed into *E.coli DH5α* cells. Colony screening was done by insert release from the ligated vectors using EcoR1 and BamH1 restriction enzymes, and final confirmation was done by DNA sequencing. Sequence verified clones were then proceeded for protein preparation.

GTC	BamH1	TEV	ATG ATGGAAGTGGCAGAGCCCAGCAG
GTC	EcoR1	STOP	CCATTGAGGTTGAGGCCACTGTCTGA

Figure 4.2: Primer used during PCR amplification of MERIT40.

4.2.1 Protein expression: MERIT40 was expressed in bacterial system using *E.coli* Rosetta (2DE3) strain. For protein expression, 100 µg/µl plasmid construct was transformed into *E.coli* Rosetta (2DE3) cells and grown on LB agar plate containing antibiotic ampicillin (100 µg/ml). Further protocol is described below.

Protocol for protein expression:

1. **Inoculation:** Pick-up a single transformed colony from antibiotic resistant LB agar plate and inoculate it in 100 ml LB broth containing 100 µg/ml of ampicillin. Incubate on shaker incubator at 37°C over night.
2. **Dilution:** Inoculate 10 ml of starting culture to 1000 ml (1: 100 ratios) of autoclaved LB broth containing the 100 µg/ml of ampicillin. Incubate the flasks on shaker incubator at 37 °C until it has reached mid-log phase i.e. A_{600} between 0.6-0.8.
3. **Induction:** Cool down the flasks and add 400 µl IPTG (stock 1M) and incubate on shaker incubator at 24°C for 16 hours.
4. **Harvesting:** Transfer the culture to centrifuge bottles and centrifuge for 10 minutes at 6000 rpm at 4°C. Resuspend the pellet in a small amount of supernatant and centrifuge for 15 minutes at 5000 rpm, 4 °C.
5. **Storage:** The pellet obtained can be stored at -80 °C until further use.

Protein purification was carried out using affinity chromatography followed by size exclusion chromatography by AKTA-FPLC to get pure and homogenous protein.

4.2.2 Purification of MERIT40:

Purification buffer composition: 300 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, 10 mM β-ME, 5% glycerol, pH 7.5

FPLC buffer composition: 300 mM NaCl, 10 mM HEPES, 0.1 mM EDTA, 10 mM BME, pH 7.5

Protocol for purification of MERIT40:

1. **Re-suspension:** Re-suspend the pellet of MERIT40 in 40 ml of purification buffer supplemented with 200 µl of 200 mM PMSF and 20 µl of protease inhibitor.

2. **Ultra sonication:** Transfer the suspension into centrifuge tube and sonicate at 50 pulse rate and 50 power with 1.45 minutes of duty cycle.
3. **Centrifugation:** After sonication the suspension is subjected to centrifugation at 18000 rpm for 45 minutes at 4°C.
4. **Equilibration of GST beads:** Give two column volume washes with distilled water to remove the traces of ethanol as the beads are stored in 20% ethanol and then 3 column washes with purification buffer.
5. **Binding:** The soluble fraction obtained is brought to room temperature and then mixed with affinity resin properly and incubated at room temperature for 1 hour.
6. **Washing:** After binding, the column is given 4 column washes with wash buffer so as to remove non specific protein. Take 40µl of beads to load on gel to observe the bound protein.
7. **Cleavage:** Add 400 µ of TEV protease enzyme, 40 µl of protease inhibitor cocktail and 100 µl of PMSF in 20 ml of purification buffer and perform the cleavage step for 3 hours by passing the solution through column at interval of 1 hour. This step is performed to cleave the GST tag from fusion protein to get the purified native protein. Take out 40ul of beads to observe the cleavage of protein.
8. **Elution:** After TEV cleavage the protein is eluted in 30 ml of purification buffer.
9. **Calibration of Ni-NTA resin:** Give 2 column washes with double distilled water and then 5 to 6 column wash with washing Buffer.
10. **Metal Ion Chelate Affinity Chromatography:** After calibration of Ni-NTA resin, pass the eluted fractions through them. This step was done to remove his-tagged TEV enzyme from protein of interest.

11. **Concentrating the protein:** Transfer the eluted protein in pre-equilibrated 10 kDa centricon and concentrate the protein upto 2 ml by centrifuging it at 4500 rpm for 10 minutes at 4°C. Check the concentration on Nanodrop spectrophotometer (280 nm). Centrifuge for 10 minutes at 5000 rpm at 4 °C to remove any soluble aggregates or precipitates.
12. **Gel filtration:** Inject 2 ml of concentrated protein on AKTA- FPLC against FPLC buffer.
13. **Fraction collection:** Collect the purified protein obtained through FPLC in 1.7 ml eppendorf at its elution volume according to gel filtration spectra profile of sample.
14. **Loading on SDS-PAGE 12% gel:** Load 20 µl of FPLC fractions on SDS-PAGE, stain with commassie dye, and then destain to visualize the protein of interest.
15. **Concentrate the protein:** The fractions which showed purified protein bands of interest were further concentrated as per requirement.

The purified proteins were used in various bio-physicochemical experiments. The complete details of protocol have been discussed earlier in chapter 3 (material and methods).

4.3 Results and discussion

4.3.1 Cloning, expression and purification of MERIT40: Selected potential clones when digested with the EcoR1 and BamH1 restriction enzymes showed the insert release of appropriate size (**Figure 4.3 B**). Sequencing result confirmed the presence of ligated gene of interest in the vector with desired frame of codon sequence.

MERIT40 eluted from superdex 200 column into the monomeric and dimeric form with some proportion of aggregated protein (**Figure 4.4**). The elution profile of MERIT40

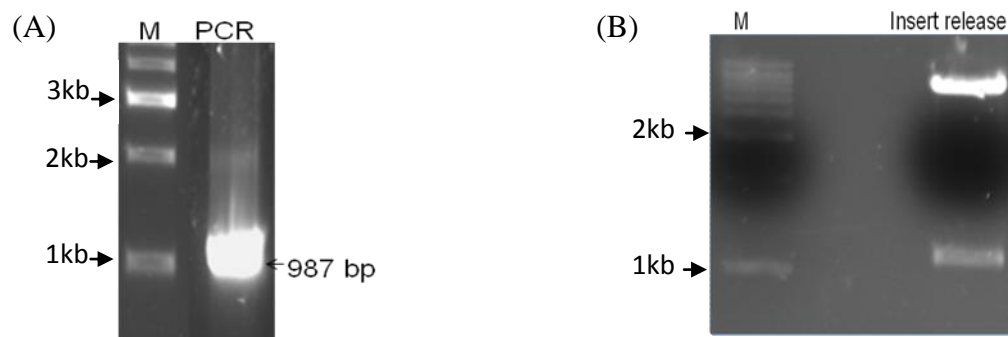


Figure 4.3: Agarose gel showing the PCR amplified product (A) and insert release from potential clones' plasmids (B) of pGEX-kT-MERIT40.

suggests that most of the protein is monomeric in nature, while dimeric proportion was limited to ~40%. The monomeric fractions were proceeded for biophysical characterization and crystallization.

4.3.2 Molecular modeling: Predication of protein function often requires knowledge of structure and physiochemical properties [166, 167]. Till date, MERIT40 is the least characterized member of BRCA1- complex, and literature search could not find any structure that exists in the protein data bank. It showed least homology with existing proteins in database and hence could not be modeled using homology modeling.

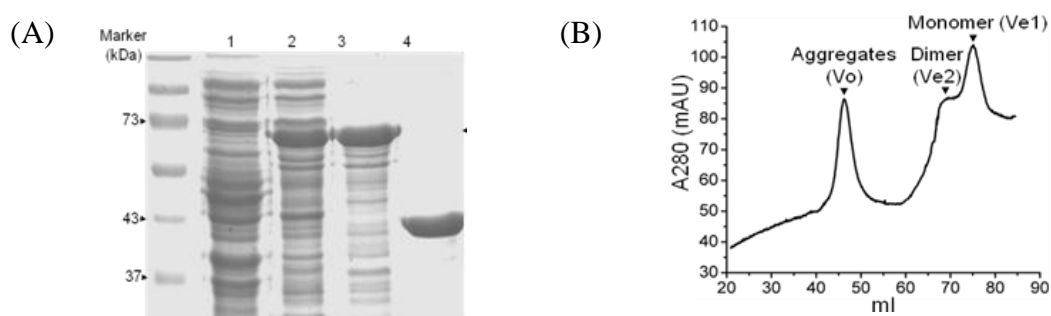


Figure 4.4: Expression and purification of MERIT40. (A) SDS-PAGE showing expression and purification profile of MERIT40. Lane 1-uninduced, 2-soluble protein, 3-protein bound on affinity resin, 4- purified protein. Single arrow- fusion protein. (B) Gel filtration chromatography elution profile of MERIT40 (AKTA-FPLC Superdex 200). V_e represents the elution volume and V_o is for void volume [168].

Considering the fact of non-availability of suitable template for homology modeling, we decided to perform ab-initio modeling using Robetta server which employs combinatorial approaches for model generation. [136] The model built for the full-length MERIT40 sequence is shown in **Figure 4.5**. The overall quality and reliability of the model was assessed based on the energy and stereo chemical geometry of the structure (PROVE_Plot Z-score RMS: 30.936, overall quality factor (ERRAT): 84.936, Verify_3D (of 87.88% residues) score>0.2).

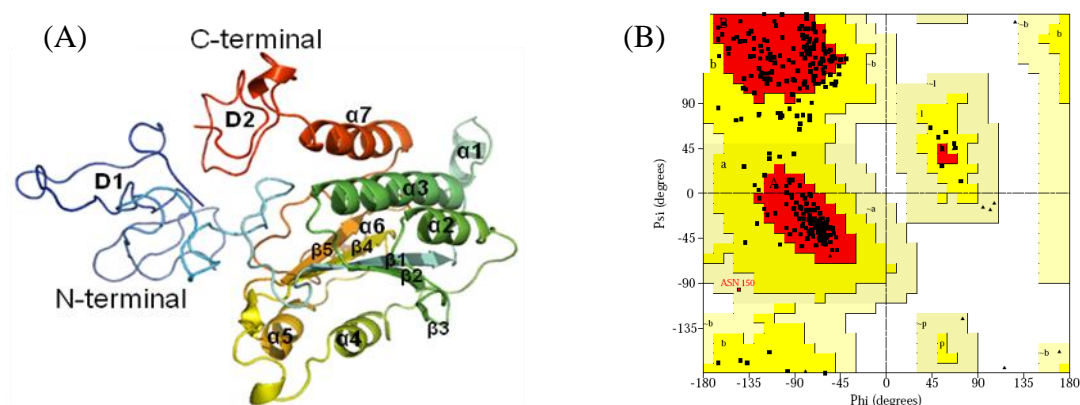


Figure 4.5: Modeled structure of MERIT40 and its validation. Full length protein structure was modeled from Robetta server (A); Modeled structure having 5- β sheets, 7- α helices and disordered region at N-terminal and C-terminal (α 1-7: alpha helices, β 1-5 : beta sheets, D1-Diordered 1, D2-Diordered 2). (B) Ramachandran plot of modeled structure [168].

The Ramachandran plot has shown 99.7% of the residues to be present in the allowed region [169] and none were in disallowed region (**Figure 4.5**) [170, 171]. The observed coordinates in Procheck analysis were found to be satisfactory (>80%), Verify 3D profile revealed that the average score remained above 0.2 and never dropped to below zero, which is significant [172] [173]. Furthermore, structural geometry was evaluated using Mol Probity web server (<http://molprobity.biochem.duke.edu/>) which showed all the parameters in required range (**Table 4.1**). The model quality was found reliable, and can be used for further studies such as docking and domain organization.

Table 4.1: Protein geometry evaluation using Mol Probiy web server [168]

Structural parameter	Value (%)
Poor rotamers	00.00
Ramachandran outliers	00.00
Ramachandran favored	92.66
C β deviations >0.25Å	00.00
Bad backbone bonds	00.00
Bad backbone angles	00.00
Asn's, Gln's, and His's orientation (correctly)	100.0

4.3.3 Domain Organization of MERIT40: Intrinsic dis-orderedness is a fairly common phenomenon in many proteins, especially those related to cancer and testis [174]. It has been reported that intrinsically disordered proteins lack the rigid three dimensional structures either for entire length or in localized regions [175]. Due to the difference in conformations and interactions, they play an important biological role in major cellular processes such as transcription, cell cycle regulation, signal transduction and DNA repair [176]. Since MERIT40 is a member of BRCA1-complex, and RAP80 has reported association with retinoid-related, testis associated receptor (RTR), it would be more likely that MERIT40 has disordered regions [108]. To assess this possibility and map the disordered region, the structure of MERIT40 was analyzed using *in-silico* and *in-vitro* approaches. From the model of MERIT40, it has been observed that N- and C-terminal regions are dis-ordered while its middle domain exhibits a well folded β/α conformation. Disordered regions of the N-terminal (~ 70 amino acids) and C-terminal (~ 30 amino acids) displayed a random structure pattern (**Figure 4.5**). The ordered region of the MERIT40 (~71–299 residues) encompasses four parallel, one anti-parallel β -strand and seven α -helices. The modeled structure provides the crucial information for disorderness, hence it would be interesting to see whether the same scenario exists in an *in-vitro* condition as well. Limited proteolysis is one of the widely used approach for

identification and mapping of disordered regions of proteins in solution [177]. A compact globular domain of protein should significantly resist the protease digestion, whereas disordered region undergoes rapid digestion due to more accessibility of protease cleavage sites [177].

To confirm the dis-orderness of the N- and C-terminal regions and identify the stable domain, MERIT40 was purified using affinity chromatography (**Figure 4.4**) and subjected to limited proteolysis assay. Trypsin and chymotrypsin proteases are commonly used to locate the compact globular domain(s) in solution [178]. It was observed that both proteases rapidly digested MERIT40 (amino acids 1–329) to a proteolytically stable fragment. MALDI-TOF/TOF analysis indicated that chymotrypsin digested fragment lacked the first 70 amino acids from the N-terminal region and a short stretch of 31 amino acids from the C-terminal (**Figure 4.6**). Furthermore, trypsin digestion generated similar pattern at the N-terminal leaving the C-terminal intact (**Figure 4.6**). Mass spectrometry identification of the stable fragment obtained after limited proteolysis suggested that it belongs to a region from ~ 71–298 residues (**Figure 4.6, Figure 4.7**). This stable fragment is likely defined in the range of MERIT40 vWFA-like domain (95–298), which has sequence homology with the von Willebrand factor (vWA) domain of the Rpn10 proteasome subunit [20, 179]. Based on the structural similarity, it seems that the stable fragment of MERIT40 (71–298) may have proteasome-like activity, thus raising the possibility of its involvement in the ubiquitination and deubiquitination processes along with RAP80 UIM.

4.3.4 Role of MERIT40 in DNA damage repair: Unraveling the function of a novel protein using multidisciplinary approach is a challenging task. Structure based function

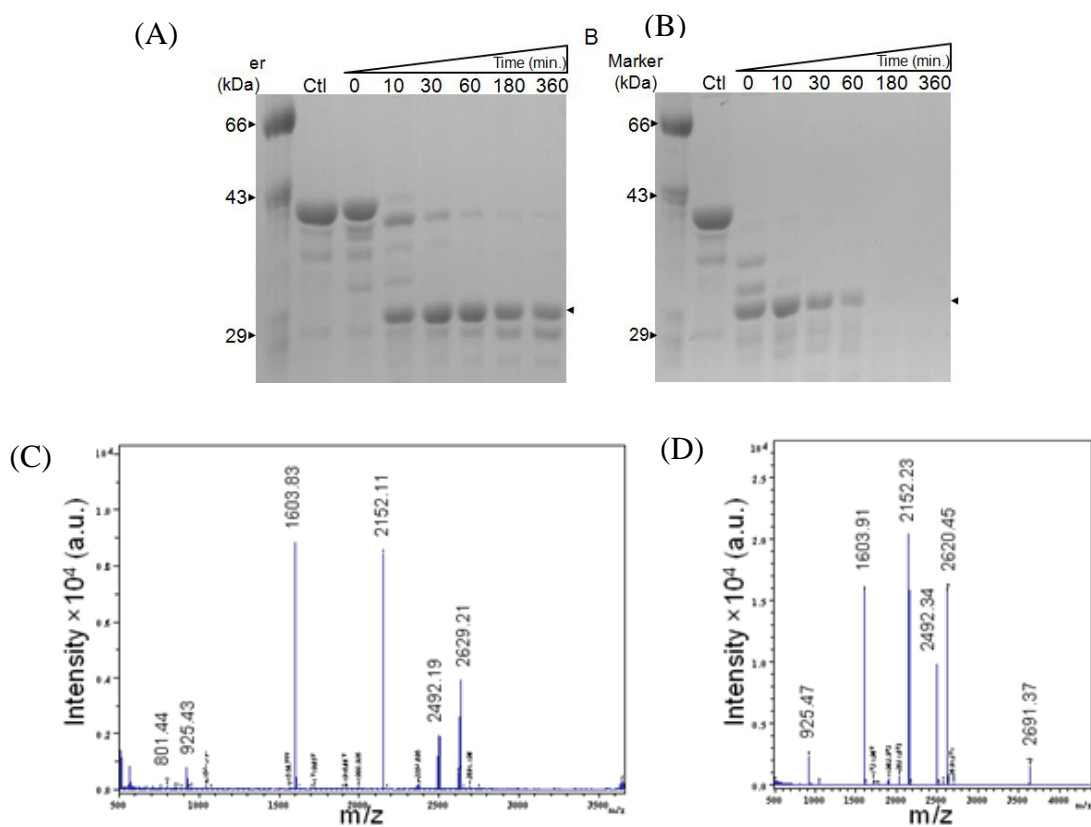


Figure 4.6: Limited proteolysis followed by mass spectrometry of MERIT40. Time dependent digestion of purified MERIT40 with (A) chymotrypsin and (B) trypsin (single arrow-digestion product). Peptides generated in mass spectrometry after digestion with (C) chymotrypsin and (D) trypsin digestion [168].

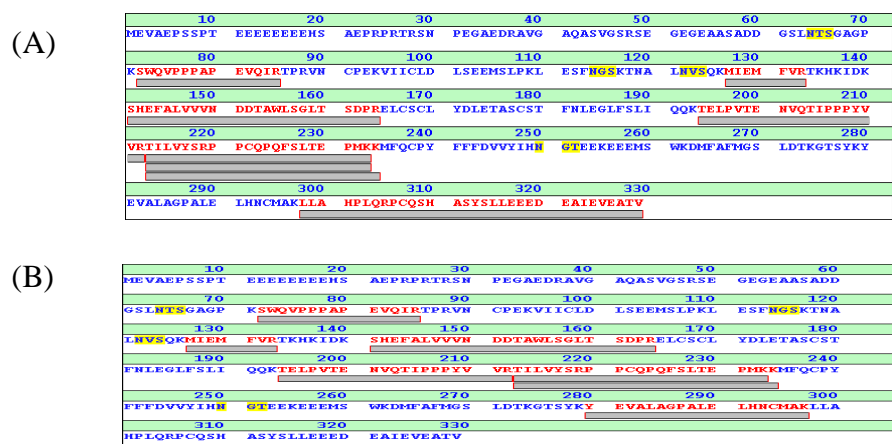


Figure 4.7: Matched peptides of proteolytically stable region in case of (E) trypsin and (F) chymotrypsin digestion.

prediction is one of the approaches for identifying the function of proteins in biological processes [135]. DALI server predicts the probable structural homologs of the protein (Holm & Rosenstrom 2010).

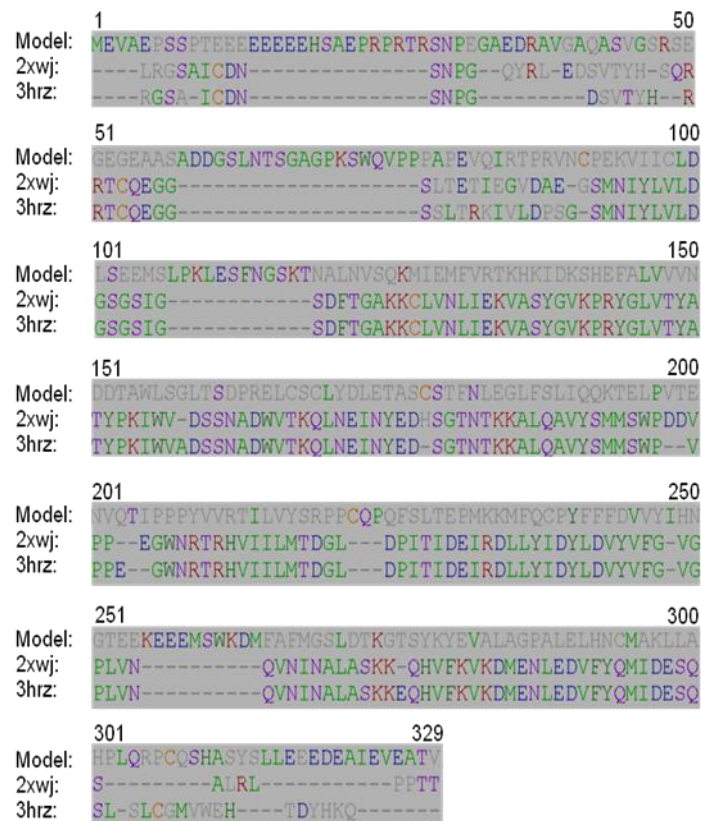


Figure 4.8: Multiple structural alignment of MERIT40 model with DALI searched templates [168].

DALI search was carried out to find the structural homologs of MERIT40 to predict its possible function [14]. Our search could find only three structural homologs, (i) Cobra Venom Factor (PDB ID: 3hrz), (ii) Complement Factor B (PDB ID: 2ok5) and (iii) Complement C3 Beta Chain (PDB ID: 2xwj) (**Figure 4.8**). Cobra Venom Factor (CVF) is a complement-activating protein in Cobra venom which plays a crucial role in host defense [180]. It resembles structurally and functionally with Complement C3 that plays an important role in complement activation and immunity (**Table 4.2**) [181]. The β -chain of Complement C3 is homologous to the α -chain of CVF, and the C-terminal and N-

terminal portions of Complement C3 α -chain are homologous to the α - and β - chains of CVF, respectively [182, 183]. Complement fragment (C3b)/CVF interacts with Factor B (FB), a single-domain serine protease which circulates in plasma, and forms the pro-convertase complex during complement activation. FB consists of a central vWA domain and a C-terminal trypsin like serine protease (SP) domain that form the protease segment. Structural similarity of MERIT40 vWFA like domain with a protease complex highlights its possible role in proteasome-like activity and involvement in ubiquitination processes.

Table 4.2: Structural comparison of MERIT40 with structurally similar protein using DALI [168]

Protein	PDB ID	Z-score	Sequence identity (%)	RMSD
Complement C3 Beta Chain	2XWJ	15.2	11	3.7
Cobra Venom Factor	3HRZ	14.9	10	4.1
Complement Factor B	2OK5	13.7	10	4.2
Complement C2a Fragment	2I6S	13.2	10	3.8
Complement C2	2ODP	13.0	10	3.9

4.3.5 Oligomeric properties of MERIT40: Oligomerization is a common property found among disordered and hydrophobic patch containing proteins [138]. These proteins exist in several oligomeric form which may or may not have biological significance. To determine the magnitude of oligomerization and types of different oligomeric species, MERIT40 has been characterized using size exclusion chromatography, native gel electrophoresis, glutaraldehyde cross-linking, mass spectrometry, Dynamic Light Scattering (DLS) and different *in-silico* tools. Gel filtration profile showed that the protein predominantly exists as a monomer, however, a second peak corresponding to dimer was also observed (**Figure 4.4**). Furthermore, to confirm the presence of dimeric population and to delineate the multimeric behavior of MERIT40, native gel electrophoresis was performed. The presence of two distinct populations on native gel

corresponding to monomer and dimer provided evidence that MERIT40 also exists in a dimeric form. In order to evaluate the proportion of dimeric population and to confirm whether it occurs in a concentration-dependent manner, different concentrations of MERIT40 were analyzed on native gel. We observed that MERIT40 exhibit a concentration-independent formation of dimers and the dimeric proportion remained same irrespective of increasing concentration from 0.1mg/ml to 10mg/ml (**Figure 4.9**). To support these finding and re-evaluate the concentration independent dimer formation, different concentrations of protein (0.1-10 mg/ml) were treated with glutaraldehyde and run on SDS-PAGE. Dimeric fractions remained same, irrespective of the changes in concentration, thus validating MERIT40 dimerization is concentration independent (**Figure 4.9**).

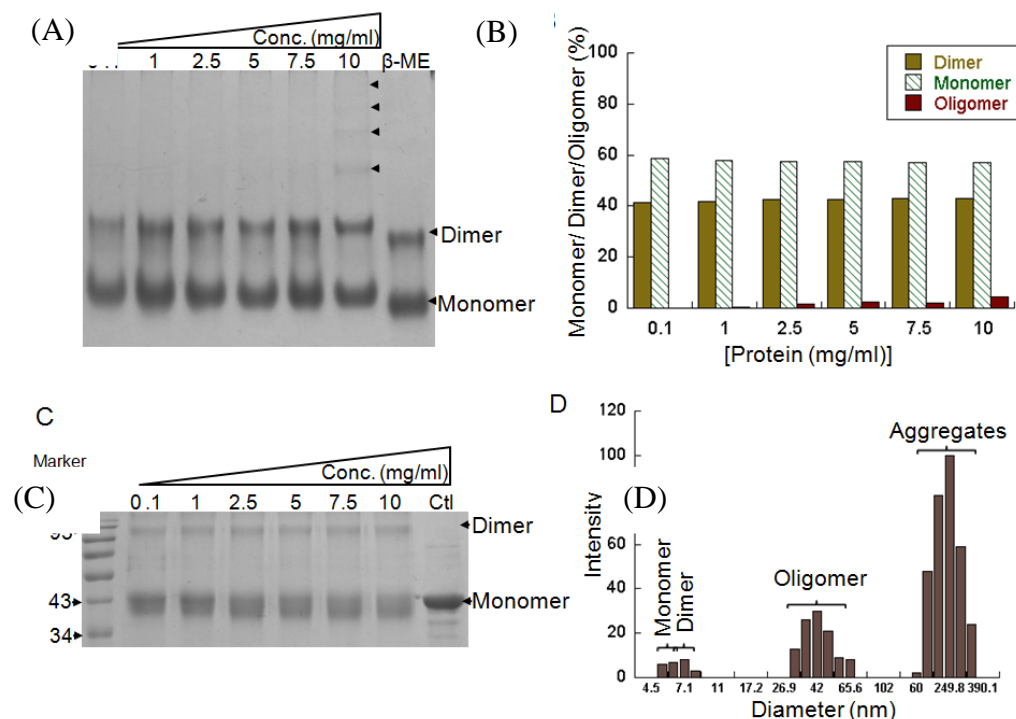


Figure 4.9: Oligomeric characterization of MERIT40. (A) Native gel (non denaturing), (B) densitometry profile native gel, (C) Glutaraldehyde Cross-linking reaction product on SDS-PAGE and (D) Dynamic light scattering. Last lane in native gel: β -ME treated sample Ctl: control as untreated with glutaraldehyde [168].

To further comprehend these findings, mass spectrometry MALDI-TOF was performed which revealed the presence of monomeric and dimeric population (**Figure 4.10**). It is often observed that molecules undergo dimerization due to inter-molecular disulfide linkage [184]. To investigate further whether dimer formation is occurring due to disulfide linkage, the purified MERIT40 protein was treated with β -ME and loaded onto native page. There was no effect found on the dimeric population even after β -ME treatment, indicating dimer formation probably does not occur due to inter-molecular disulfide linkage (**Figure 4.9**). Densitometric analysis suggested that the monomeric and dimeric protein fractions were found to be approximately 60% and 40% of the total protein, respectively, in a range of 0.1–10 mg/ml concentrations (**Figure 4.9**). Similar proportion was also observed on gel filtration chromatography (**Figure 4.4**). Hence MERIT40 exists as a dimer but no relative concentration dependency exists for dimer formation.

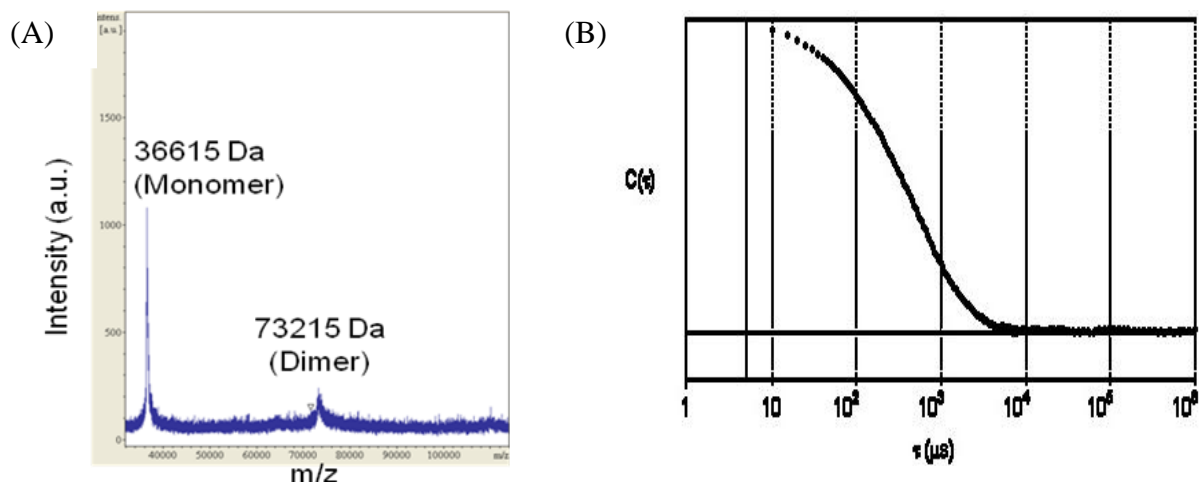


Figure 4.10: (A). MALDI-TOF mass spectrometry profile of full length MERIT40 showing presence of monomer and (B) dimer Correlation coefficient obtained in Dynamic Light Scattering analysis of MERIT40 for molecular size measurement [168].

To determine the heterogeneity present in purified MERIT40 protein, Dynamic Light Scattering (DLS) experiment was performed. It was observed that majority of samples are composed of three different clusters. The first cluster corresponds to the monomeric and dimeric populations, while the second and third clusters correspond to higher-order oligomers or aggregates (**Figure 4.9, Figure 4.10**). The diameter of the monomer determined from modeled structure using *in-silico* approach was average 5–6 nm and the effective hydrodynamic diameter determined by the DLS was found in 5–6 nm range, which showed good concurrence with each other (**Figure 4.5, Figure 4.9**).

4.3.6 Dimer Interface of MERIT40: In order to determine the *in-silico* symmetrical dimer formation, blind docking of MERIT40 monomeric molecules was performed after molecular dynamic simulation (MDS) using Patchdock and Symmdock servers [137]. Only dimeric (diameter: 6–8 nm) form was observed in the docked structure (**Figure 4.11**). The dimeric structure was stabilized by involvement of Gly115, Gln125, Lys126 and Arg133 residues at the interface. To validate these findings, *in-silico* substitution of these residues with bulky and charge repulsive (Gln115, Ile125, Glu126 and Glu133 respectively), as well as Ala residues were made, followed by docking of monomers. Substitution of either of the residue in the monomers, disfavored the *in-silico* dimerization of MERIT40 (docking scores [185] of various substitution: wild-type 14640, Gly115Gln 13430, Gly115Ala 13408, Gln125Ile 14620, Gln125Ala 14614, Lys126Glu 14596, Lys126Ala 14308, Arg133Glu 14110, Arg133Ala 13392). The energy (kJ/mol) of dimeric complex due to Gly115Ala (-19.56), Gln125Ala (-6.78), Lys126Ala (0.84) and Arg133Ala (-19.43) was found higher as compared to wild-type (-47.92),

indicating relatively decreased stability of dimers due to substitution with the alanine residue.

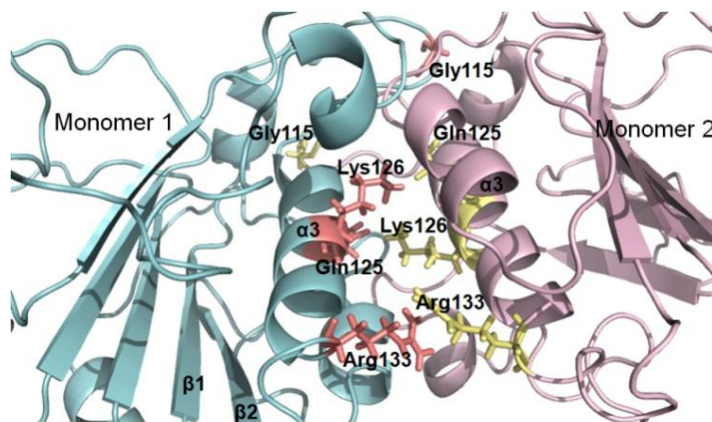


Figure 4.11: Dimer interface of MERIT40. Monomer1 interface residues are with light red color and monomer 2 with light yellow [168].

4.3.7 Structural analysis of MERIT40: To characterize the secondary structural constituents of MERIT40, FPLC purified MERIT40 protein was used to record far-UV Circular- Dichroism (CD) spectrum (**Figure 4.12**). The CD spectra were analyzed using DichroWeb server. It was observed that the α -helices and β -sheets are approximately 20 and 16.6%, respectively [186]. Furthermore, secondary structural components of modeled structure was determined using the PROSS server (<http://roselab.jhu.edu/utis/runpross.html>), which showed the percentage of α -helices and β -sheets to be ~ 23 and 15%, respectively. These results from the two independent approaches were in good agreement.

Tryptophan and tyrosine residues are the most frequently used intrinsic fluorophores to study the conformational changes induced in the protein due to external agents like temperature and chaotrophs [187]. These aromatic residues are generally buried inside the protein and are extremely sensitive to modifications around their micro-environment.

To study the three-dimensional structure of MERIT40, Trp and Tyr microenvironment were monitored using fluorescence spectroscopy.

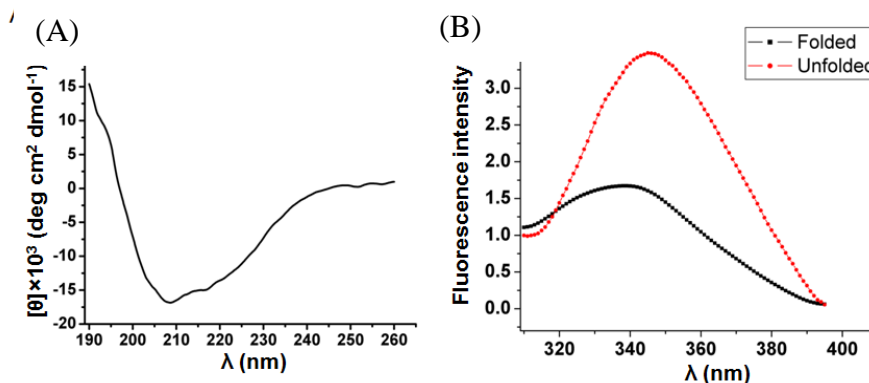


Figure 4.12: Far-UV CD spectra (A), and urea unfolding (B) using fluorescence, of MERIT40 [168].

An emission maximum at λ 338 nm was obtained for the native protein, while a red shift at λ 346 nm was observed for the unfolded protein (**Figure 4.12**). This indicates that most of the tryptophan and tyrosine residues of the native protein are buried inside the hydrophobic core.

4.3.8 Assessment of stability and unfolding pathway: To obtain the stability parameters and folding pattern of MERIT40, thermal denaturation was studied using CD and fluorescence spectroscopic probes. In Far-UV CD data, maximum change in ellipticity at different temperature was observed at λ 208 nm, which was therefore used to monitor the changes in protein's secondary structure at varied temperatures (15-80°C) (**Figure 4.13**). Upon data fitting the T_m (melting temperature) was calculated to be $48 \pm 4.1^\circ\text{C}$, and thermodynamic parameters were $\Delta G^\circ_{\text{H}_2\text{O}}$ 4.8 ± 1.0 kcal/mol, ΔH 144 ± 2.5 kcal/mol. To further validate these findings, protein was unfolded in a temperature gradient of 15-80°C and unfolding was monitored by recording fluorescence emission for Trp, Tyr upon excitation at λ 280 nm. The change in emission maxima of protein at various temperatures was used to determine its folding pathway and thermodynamics

parameters. Fluorescence spectroscopy studies showed a T_m value of $52 \pm 3.2^\circ\text{C}$ and energy parameter ($\Delta G^\circ_{\text{H}_2\text{O}}$ 4.6 ± 1.5 kcal/mol, ΔH 142 ± 2.3 kcal/mol (**Figure 4.13**). These parameters indicate that MERIT40 possessed moderate stability, and its enthalpy and free energy values are similar to globular protein [188].

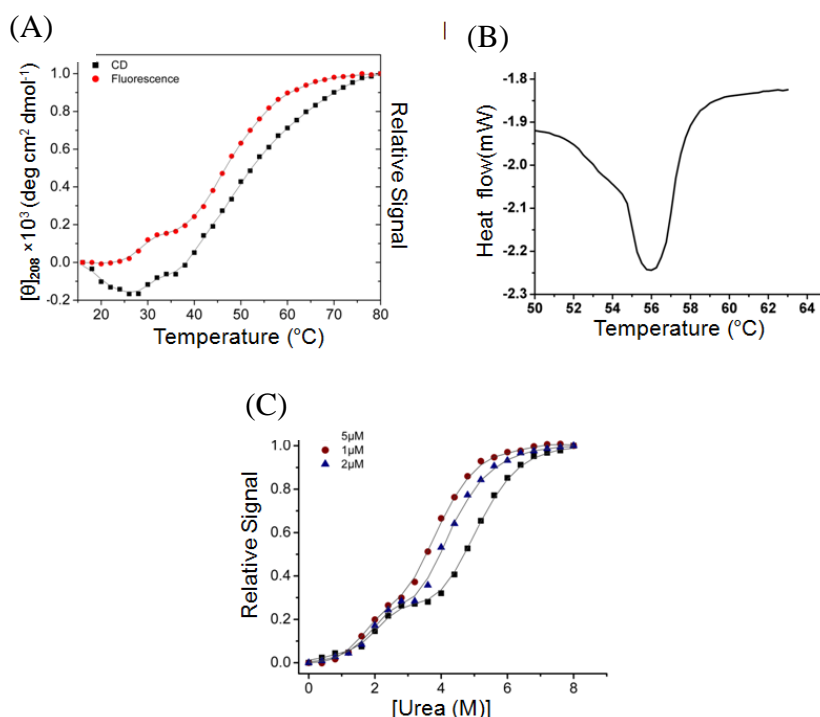


Figure 4.13: Denaturation profile of MERIT40 using CD, DSC and fluorescence spectroscopy. (A) Overlay of protein fractions unfolded at different temperature in CD and fluorescence spectroscopy, (B) DSC thermogram and (C) Chemical unfolding using fluorescence at different urea concentration. Relative signal represent fraction unfolded [168].

Both methods demonstrate that the protein most likely unfolds via the formation of an intermediate species as indicated by abrupt increase of signal in the unfolding curve. In order to validate these findings, the unfolding profile of native protein was assessed using differential scanning calorimetry (DSC). Purified MERIT40 protein in the concentration of 1 mg/ml was allowed to unfold in a temperature gradient of 15-80°C in calorimetry, and unfolding transition was obtained. A peak maximum during unfolding transition was considered as melting temperature (T_m). DSC thermogram suggested overlapping of two

transitions, which could be corresponding to intermediate species formed during the unfolding process (**Figure 4.13**). The melting temperature ($55\pm0.9^{\circ}\text{C}$) and enthalpy parameters ($\Delta H\ 150\pm2.0\ \text{kcal/mol}$) were derived using Calisto software. The T_m value determined by DSC and fluorescence were similar and unfolding profile indicated that MERIT40 protein forms an intermediate species. Considering the dimeric behavior, it can be concluded that MERIT40 forms the dimeric intermediate during unfolding process. To test this hypothesis, a concentration dependent urea denaturation study was performed to find the type(s) of intermediate species formed during unfolding process. Protein was incubated in different urea concentration (0-8M) till equilibrium was achieved and emission maximum using fluorescence were recorded. Change in average emission wavelength at different urea concentrations was plotted to calculate unfolded fraction. A biphasic unfolding curve with two transitions was observed and protein concentration dependency in the second transition indicated that dissociation of MERIT40 subunits occurs due to a dimeric intermediate [189] (**Figure 4.13**). The chemical denaturation using urea as the unfolding agent substantiates the thermal denaturation findings.

4.3.9 Clinical implication of MERIT40 variants: Mutations or variants are known to exist almost in every member of BRCA1-complex. However, predicting the pathogenicity of mutations discovered in the cohort of patients is still challenging and is a subject of extensive studies [5, 190]. Mutations reported on BRCA1 has been reported to have role in cancer susceptibility [191]. Disease implication caused by these variants arise either due to alteration of whole domain structure or localized alteration in weak inter- or intra-molecular interactions [191]. To characterize the pathogenicity of sequence variants identified in the patient cohort and further evaluate the importance of weak intra-

molecular interactions in disease susceptibility, seven recently reported sequence variants of MERIT40 (c.342C > T, c.344 + 41A > T, c.393C > T, c.787 – 6C > T, c.821A > G, c.837G > A, c.87G > A) were analyzed.

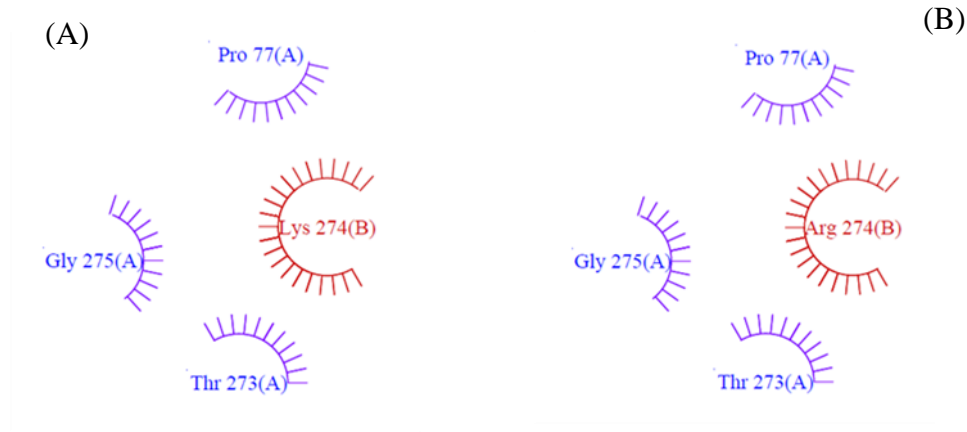


Figure 4.14: Mutational analysis of MERIT40. (A) ligplot of Lys 274. (B) ligplot of Arg 274 [168].

It is reported that none of the above variants have any role in predisposing the individual to disease [190]. All the variants are silent in nature except of c.821A > G in exon 9, which lead to the substitution of Arg to Lys at 274 position. However, change in Lys to Arg doesn't bring any pathological implication as observed in cohort of patients [190]. We decided to explore the cause of MERIT40 K274R variant neutrality. The molecular environment around the Lys274 residue and changes as a result of substitution with Arg was analyzed. Structural analysis of modeled structure unraveled that Lys274 is buried in the hydrophobic core of the protein which probably restricts its binding with other intra amino acids (**Figure 4.14**). Lys274 forms several non-bonded interactions with Pro77, Gly75 and Thr273. There were no changes found in the structural conformation and hydrophobic environment around Lys274 residue of MERIT40 as a result of substitution with Arg and majority of interactions with adjacent residues remained unaltered (**Figure 4.14**). Since intramolecular interactions around the mutant remained unaffected, it could

be assumed that MERIT40 K274R mutant is clinically insignificant. This assumption is in agreement with the findings of a previous report which states that mutations predisposing to breast cancer are either very rare or absent in the coding region of MERIT40 [190].

4.4 Crystallization of MERIT40

Initial crystallization trials for MERIT40 protein was set with 15 mg/ml protein concentration. The protein and mother liquor solution was mixed in 1:1 ratio (1 μ l+1 μ l) and allowed to crystallize at 22°C with 500 μ l reservoir solution in a closed system. A clear drop or light precipitation was observed in most of the drops.

4.5 Conclusion

MERIT40 can be classified as an intrinsically disordered protein due to the presence of N-and C-terminal disordered region. However, its middle region showed well defined compact structure which indicates a good possibility of crystallization of middle region for X-ray diffraction study. Structural homologs of MERIT40 suggests its plausible role in complement activation pathway, nevertheless, definite function determination needs further experimental evidences. Its central region shows remarkable stability towards protease digestion and has structural similarity with vWA-like region, a domain mainly present in complement activation factors. MERIT40 undergoes a three-state unfolding transition pathway with a dimeric intermediate. It exhibits concentration independent dimerization and unfolds through a dimeric intermediate. MERIT40 could be a multifunction molecule having role in DNA damage repair and complement activation.

Chapter 5

Structural & functional analysis of ABRAXAS

5.1 Introduction

ABRAXAS acts as a bridging molecule among the BRCA1-complex family which comprises of RAP80, ABRAXAS, BRCA1, MERIT40, BRCC36 and BRCC45 proteins [6, 15-18]. RAP80 can form DNA damage induced foci of BRCA1 complex in two different manners, one is Ubiquitin Interacting Motif (UIM) dependent while the other is UIM independent but AIR (ABRAXAS Interacting Region) dependent. BRCA1 complex foci formation is essential for execution of homologous recombination repair process after DNA damage. siRNA mediated knockdown of ABRAXAS reduces the BRCA1 complex foci formation in IR induced DNA damage, and the BRCA1 level was found significantly lower compared to the RAP80 UIM truncation [87]. RAP80 double mutants lacking UIM and AIR completely abolish the foci formation, thus indicating potential role of ABRAXAS in BRCA1-complex foci formation [87]. Furthermore, ABRAXAS and RAP80 knockdown cells showed defective homologous recombination repair and increased hypersensitive to IR and UV radiation [20, 24]. Role of ABRAXAS is also suspected in G2/M check point activation since depletion of RAP80 and ABRAXAS display defective G2/M check points [6, 7]. ABRAXAS comprises pS-X-X-F (SPTF) binding motif at its C-terminal which interacts with BRCA1-BRCT phospho-peptide binding domain and brings about foci formation (**Figure 5.1**). The Ser 404 and Ser406 on the SPTF motif are phosphorylated by ATM/ATR after IR treatment and is required for the phosphodependent interaction between ABRAXAS and BRCA1 BRCT domains [6, 7]. Moreover, S406A mutation abolishes the foci formation between the ABRAXAS and BRCA1, suggesting phosphodependent interaction between them [192]. ABRAXAS acts upstream to BRCA1 and its knockdown significantly reduce accumulation of BRCA1-

complex at DNA damage site. There are possible augmentations that loss of ABRAXAS or BRCA1 could lead to similar phenotype, thereby projecting ABRAXAS as an excellent candidate gene for familial breast cancer [6, 16, 192, 193].

Several reported mutation in BRCA1 associated proteins either disrupt their interaction with BRCA1 or with other binding partner, thereby causing various pathogenic implications [4, 5]. Familial mutation at C-terminus of ABRAXAS in Finnish population (Arg361Gln) is suspected to disassemble the BRCA1 complex [11]. c.1082G>A alteration was observed among 3 of 125 (2.4%) studied breast cancer families but was found to be absent in 868 healthy female volunteers [11]. This variant impairs nuclear localization of BRCA1 complex and its DNA damage repair function, thereby predisposing the individual to repair defects [11]. Immunoprecipitation of epitope tagged ABRAXAS R361Q mutant with BRCA1 and other core complex in cytoplasm displays its failure to form foci in the nucleus after DNA damage, whereas nuclear localization of wild-type is retained [11]. Failure to achieve nuclear retention adversely affects G2/M checkpoint and homology-directed DNA repair, and reduces nuclear retention to DSB site of ABRAXAS interacting partners [11]. Moreover, expression of the ABRAXAS R361Q variant causes hypersensitivity to ionizing radiation and reduced BRCA1 localization at sites of DNA damage in several cell lines[11].

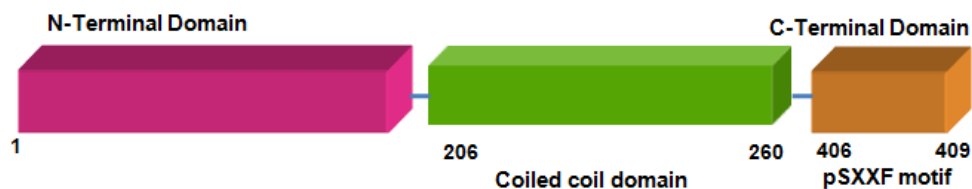


Figure 5.1: Domain organization of ABRAXAS.

These reports substantially list ABRAXAS as a new susceptibility gene to cancer predisposition and opened the vast perspective to study the mutant like R361Q discovered in cohort of patient in disease progression.

To understand the deleterious effects of mutations and possible consequences, biophysicochemical characterization and interaction analysis of ABRAXAS wild- type and R361Q mutant with BRCA1 were performed. This includes cloning, expression and purification of functional domains of ABRAXAS wild-type and mutant and their biophysical analysis

5.2 Material and Methods

A comparative structural, stability and binding analysis of ABRAXAS (6-373) wild- type and ABRAXAS (6-373) R361Q mutant was carried out. In addition, different regions of ABRAXAS (6-268 a.a), ABRAXAS (6-301 a.a), ABRAXAS (6-373 a.a), ABRAXAS (6-409 a.a) were also sub-cloned to perform binding analysis with MERIT40. Nucleotide sequence of primers used in PCR amplification for ABRAXAS (6-268), ABRAXAS (6-301), ABRAXAS (6-373), ABRAXAS (6-409) regions are shown in **Figure 5.2**, respectively. PCR condition is as follow; 95 °C denaturation (5 minutes), 95 °C denaturation (45 seconds), annealing 68°C for 35 seconds, Extension 72°C at 0.5kb/min, final extension 72°C for 10 minutes, 25 cycles. The PCR amplified product was digested with Nco1 and Xho1 and ligated into pET28(a)+ using one-step cloning protocol. The ligation mixture was transformed into *E.coli DH5α* cells. Selected potential clones were screened for the presence of insert by digesting with enzymes Nco1 and Xho1 and loading on agarose gel. Final confirmation was done by DNA sequencing.

Sequentially confirmed clones were used for protein expression and purification. The methods and materials have been described in detail in chapter 3 (material and methods). Functional domains of ABRAXAS were expressed in bacterial system *E.coli* BL21 (DE3) cells. For protein expression, 100 ng/μl plasmid construct was transformed into *E.coli* BL21 (DE3) cells and grown on LB agar plate containing antibiotic kanamycin (100 μg/ml). Further protocol is described below.

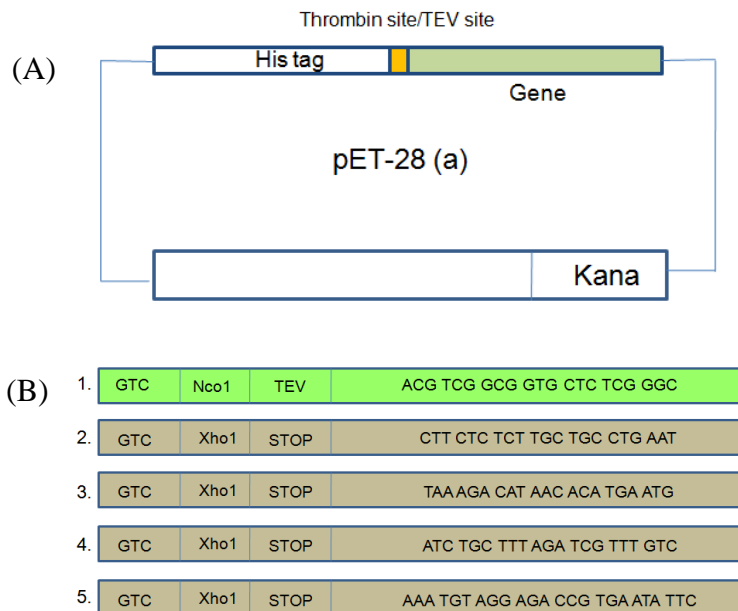


Figure 5.2: (A) Basic scheme for sub-cloning of different functional domain of ABRAXAS in pET28. (B) ABRAXAS primers used for PCR amplification; 1-(Abraxas_6 amino acids_ F), 2-(Abraxas_268 amino acids_R'), 3-(Abraxas_301 amino acids _R'), 4-(Abraxas_373 amino acids _R'), 5-(Abraxas_409 amino acids _R). F stand for forward and R for reverse primer.

Protocol for protein expression:

- Inoculation:** Pick-up a single transformed colony from antibiotic resistant LB agar plate and inoculate it in 100 ml LB broth containing 100 μg/ml of kanamycin. Incubate on shaker incubator at 37 °C over night.

2. **Dilution:** Inoculate 10 ml of starting culture to 1000 ml (1: 100 ratios) of autoclaved LB broth containing the 100 µg/ml of kanamycin. Incubate the flasks on shaker incubator at 37 °C until it has reached mid-log phase i.e. A₆₀₀ between 0.6-0.8.
3. **Induction:** Cool down the flasks and add 400 µl IPTG (stock 1M) and incubate on shaker incubator at 24°C for 16 hours.
4. **Harvesting:** Transfer the culture to centrifuge bottles and centrifuge for 10 minutes at 6000 rpm at 4°C. Resuspend the pellet in small amount of supernatant and centrifuged for 15 minutes at 5000 rpm, 4 °C.
5. **Storage:** The pellet obtained was stored at -80 °C for further use.

Further, proteins were purified with affinity chromatography followed by FPLC. ABRAXAS (6-373) R361Q mutant was expressed in *E.coli* (BL21) and purified as per wild- type protocol.

Protocol for purification of ABRAXAS (6-373) wild-type and R361Q mutant:

Purification buffer composition: 300 mM NaCl, 10 mM HEPES, 0.1mM EDTA, 2mM BME pH 7.5

FPLC buffer composition: 300 mM NaCl, 10 mM HEPES, 0.1mM EDTA, 2mM BME pH 7.5

1. **Re-suspension:** Re-suspend the pellet of ABRAXAS in 40 ml of purification buffer; containing 200 µl of 200 mM PMSF and 20ul of protease inhibitor were added.
2. **Ultra sonication:** Transfer the suspension into the centrifuge tube and sonicate at 50 pulse rate and 50 power with 1.45 minutes of duty cycle.

3. **Centrifugation:** After sonication the suspension is subjected to centrifugation at 18000 rpm for 45 minutes at 4°C. The aim here is to obtain cell free extract. Collect the supernatant and discard the pellet as it contains cell debris.
4. **Calibration of Ni-NTA beads:** Give two column volume washes with distilled water to remove the traces of ethanol as the beads are stored in 20% ethanol and then three column washes with purification buffer.
5. **Binding:** The soluble fraction obtained is brought to room temperature and then carefully mixed with affinity resin, and incubated at room temperature for an hour.
6. **Washing:** After binding, the column is given 4 column washes with wash buffer so as to remove non specific protein. Take 40µl of beads to load on gel to observe the bound protein.
7. **Cleavage:** Add 400µl of TEV protease enzyme, 40ul of protease inhibitor cocktail and 100µl of PMSF in 20 ml of purification buffer and proceed the cleavage step for 3 hours by passing the solution through column at interval of 1 hour. This step is performed to cleave the His-tagged from fusion protein to get the purified native protein. Take out 40µl of beads to observe the cleavage of protein.
8. **Elution:** After TEV cleavage the protein is eluted in 30 ml of purification buffer.
9. **Calibration of Ni-NTA resin:** Give 2 column washes with double distilled water and then 5 to 6 column wash with washing Buffer.
10. **Metal Ion Chelate Affinity Chromatography:** After equilibration of Ni-NTA resin, pass the eluted fractions through it.

11. **Equilibration of centricon:** Wash the centricon (Millipore 10 kDa Molecular weight cut off) with water to remove traces of alcohol, rinse it with washing buffer by centrifuging it at 4500 rpm for 10 minutes at 4 °C.
12. **Concentrating the protein:** Transfer the eluted protein in the 10 KDa centricon and concentrate the protein upto 2 ml by centrifuging it at 4500 rpm for 10 minutes at 4 °C. Check the concentration on Nanodrop spectrophotometer (280 nm). Centrifuge for 10 minutes at 5000 rpm at 4 °C for removal of any traces of impurities.
13. **Gel filtration:** Inject 2 ml of concentrated protein on AKTA- FPLC against FPLC buffer.
14. **Fraction collection:** Collect the purified protein obtained through FPLC in 1.7 ml eppendorf at its elution volume according to gel filtration spectra profile of sample.
15. **Loading on SDS-PAGE 12% gel:** Load 20µl of FPLC fractions on SDS-PAGE, stain it with commassie dye, and then destain to visualize the protein of interest.
16. **Concentrate the protein:** The fractions which showed purified protein bands of interest were further concentrated as per requirement.

The purified proteins were used in various bio-physicochemical experiments. The complete details of protocol have been discussed in chapter 3 (material and methods).

5.3 Results and discussion

5.3.1 Sequence alignments: Multiple sequence alignment of ABRAXAS protein sequences among various species showed a highly conserved nature of Arg 361 which indicates that it could be a critical residue for maintaining the protein structure and function (**Figure 5.3**). Hence, *in vitro* study was performed to analyze the effect of substitution mutation in ABRAXAS protein.

HUMAN	WQFKRS R LLDTQDKR
CHIMP	WQFKRS R LLDTQDKR
BOVIN	WQFKKS R LGGIQNRP
FERRET	WQFKKS R PLEIQNKP
RAT	WPAKRP R LLESESRP
MOUSE	GSVKRP R LLETESRP
CHICK	KSLKKL R SLQLDQEL
FROG	TLTKKS R LLQLQKQH

Figure 5.3: Multiple sequence alignment of ABRAXAS showing highly conserved nature of Arg361 [194].

5.3.2 Cloning, expression and purification of ABRAXAS functional domain: All deletion construct of ABRAXAS showed the insert release at proper size on Agarose gel after digestion with Nco1 and Xho1 (**Figure 5.4**). ABRAXAS wild-type eluted from Superdex- 200 column as monomer, with some proportion of aggregated protein. (**Figure 5.5**). ABRAXAS(6-373) wild -type and R361Qmutant eluted at the same elution volume corresponding to monomer.

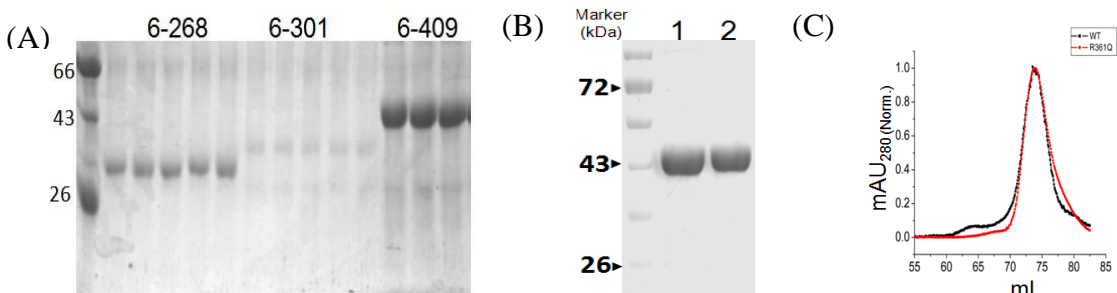
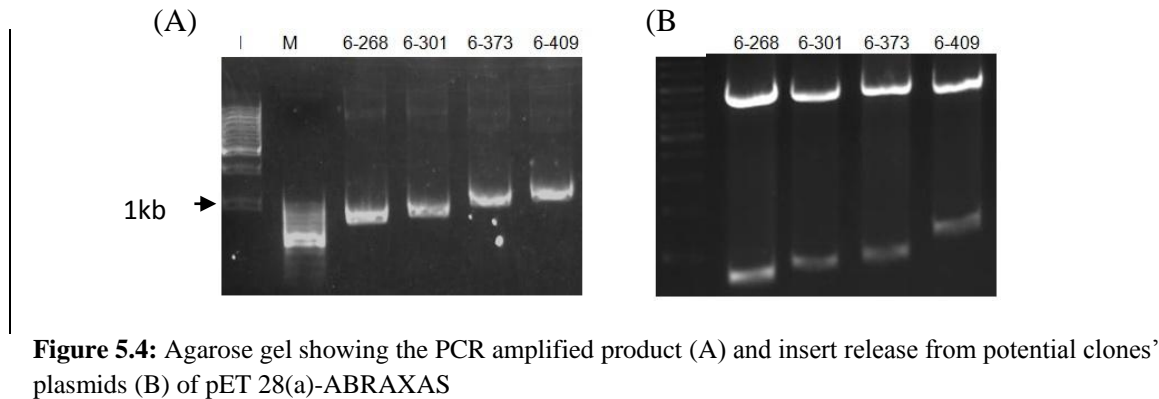


Figure 5.5: (A) Purification of ABRAXAS. SDS-PAGE showing purified protein of ABRAXAS (6-268) lane 2-6, ABRAXAS (6-301) lane 7-11 and ABRAXAS (6-409) lane 12-14. (B) ABRAXAS (6-373) wild-type (lane 1), mutant (lane 2). (C) Overlay of gel filtration spectra of ABRAXAS (6-373) wild-type and R361Q (Superdex 200) [194].

5.3.3 Oligomeric characterization of ABRAXAS (6-373) wild-type and R361Q

mutant: Substitution mutation results into distortion of secondary or tertiary structure which may lead change in the oligomeric behavior of the protein [195, 196]. ABRAXAS R361Q mutation involves substitution of a basic residue with non polar charged residue, which might change the oligomeric property of the protein.

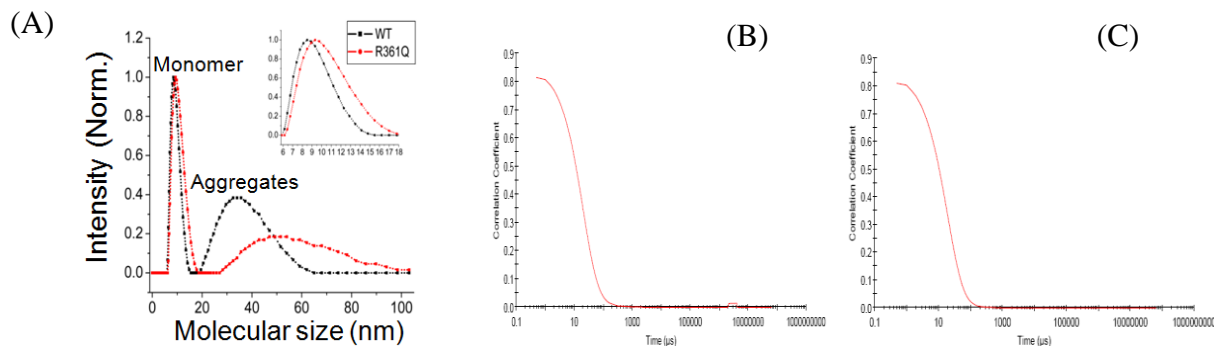


Figure 5.6: Dynamic Light Scattering profile of ABRAXAS (6-373) wild-type and R361Q. Inset showing the magnified region of the peak corresponds to monomeric population (A). Correlation coefficient of wild-type (B) and mutant (C) [194].

The observed hydrodynamic diameter of wild-type was 9.0 ± 0.3 nm while mutant showed a hydrodynamic diameter of 10.0 ± 0.5 nm (**Figure 5.6**). Relative increase in the effective diameter of mutant protein may be due to expansion of molecule dimension, nevertheless, the increment was not due to oligomerization. Moreover, the aggregates detected in case of mutant showed a broader variety as compared to wild-type and spread over a long molecular weight range. This indicates that there may be some structural changes or disordered region present in mutant which are different from wild-type. These findings supported that ABRAXAS R361Q mutation probably incorporates short or long range structural changes in the protein without changing its oligomeric behaviour.

5.3.4 Structural Organization of ABRAXAS (6-373) wild-type and R361Q: Limited

proteolysis is one of the widely used approaches for determining any structural alteration in protein due to acquired mutation [177]. In general, a compact globular domain of

protein considerably resists the protease digestion, while disordered region undergoes rapid digestion [177]. To understand domain integrity and determine the resistivity of ABRAXAS wild- type and mutant against the protease digestion, limited digestion with trypsin and chymotrypsin proteases was performed. Wild-type and mutant proteins were treated with same concentration of proteases for limited time and results were analyzed on SDS-PAGE (**Figure 5.7**). ABRAXAS wild -type resistance towards protease digestion was comparable to that of mutant protein which indicates existence of similar structural domain(s). Parallel susceptibility of ABRAXAS mutant towards protease digestion suggests that substitution of R361Q is not

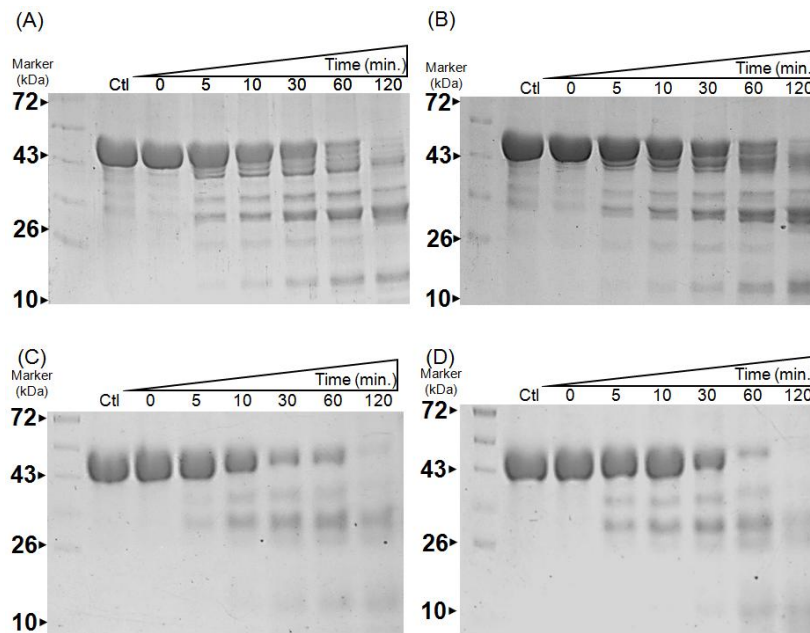


Figure 5.7: Resistivity profile of ABRAXAS wild-type and mutant towards Protease digestion. Limited proteolysis of ABRAXAS (6-373) wild-type (A, C) and mutant (B, D) using trypsin (A, B) and Chymotrypsin (C, D) as proteases. Ctl- control as untreated with trypsin and chymotrypsin, respectively [194].

destabilizing the domain integrity of ABRAXAS. Since the observed domain's organization in limited proteolysis is similar it indicates that wild- type and mutant proteins may have similar or slightly changed secondary and tertiary structures. To

evaluate this, secondary structure of ABRAXAS wild -type and mutant was compared using far-UV Circular Dichroism (**Figure 5.8**). It has been observed that ABRAXAS wild- type and mutant have well-defined α/β characteristics, however, β -sheets characters were more prominent. Data analysis using Dichroweb server suggested that wild -type and mutant are having α -helix (wild-type 15%, mutant 15.5%) and β -sheets (wild-type 24%, mutant 23%) [186]. Secondary structural comparison of wild-type and

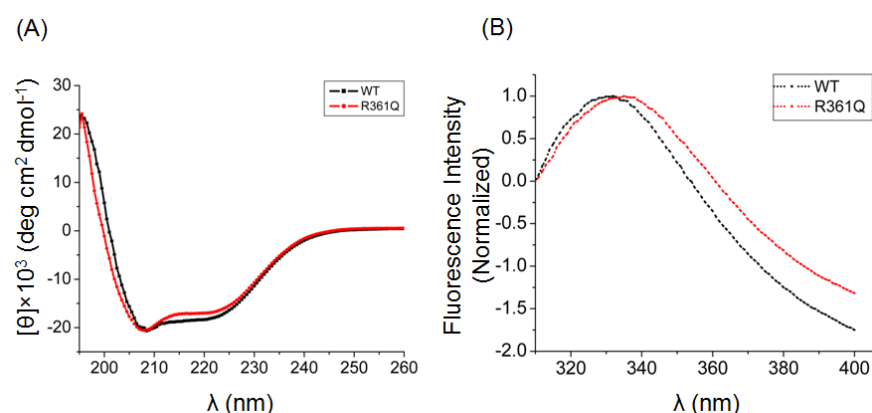


Figure 5.8: Secondary and tertiary structure evaluation of ABRAXAS wild-type and mutant. (A) Overlay of Far-UV Circular Dichroism spectrum of wild-type and mutant. (B) Overlay of fluorescence emission spectrum of wild-type and mutant [194].

mutant deciphered the similar α -helical and β -sheets characteristics which illustrates less effect of R361Q on the ABRAXAS structure. ABRAXAS R361Q mutation may be inducing short range 2° structural changes which might be collectively responsible for tertiary structural modification.

To study the three dimensional structure of ABRAXAS wild-type and mutant, Trp and Tyr microenvironment were monitored using fluorescence spectroscopy. An emission maxima at λ 332 nm obtained for wild- type and 334 nm for mutant indicated slight changes in Trp and Tyr microenvironment (**Figure 5.8**). Thus, it can be concluded that

tertiary structural components become moderately altered without causing drastic conformational changes.

5.3.5 Thermal stability and folding pathway of ABRAXAS (6-373) wild- type and

R361Q Mutant: Thermal stability of ABRAXAS wild-type and mutant was compared at secondary (CD) and tertiary structure (fluorescence) levels. The spectrum obtained from CD corresponding to $\lambda 222$ nm showed the maximum change in ellipticity and high signal to noise ratio. Change in ellipticity was plotted against the different temperatures and fraction unfolded was determined (**Figure 5.9**). Thermal stability of ABRAXAS wild-type (T_m 35°C, $\Delta G^\circ_{H_2O}$ 2.1 \pm 0.5 Kcal/mol) was found equivalent to mutant (T_m 34.8°C, $\Delta G^\circ_{H_2O}$ 1.53 \pm 0.6 Kcal/mol), and endures similar folding pathway. For stability assessment of tertiary structure, emission maxima of wild-type and mutant were monitored at different temperatures corresponding to $\lambda 280$ and fraction unfolded was calculated. Fluorescence spectroscopy revealed comparable unfolding pattern but different stability, and derived T_m value 23 \pm 3.1°C for mutant ($\Delta G^\circ_{H_2O}$ 1.25 \pm 0.5 Kcal/mol) and 27 \pm 2.5°C for wild-type ($\Delta G^\circ_{H_2O}$ 1.08 Kcal/mol) (**Figure 5.9**). Furthermore, it revealed that proteins most likely unfold through an intermediate species and undergo three state transition(**Figure 5.9**). To determine the chemical stability of wild -type and mutant, protein samples (2 μ M) were incubated in different concentrations of urea (0-8 M) till equilibrium was reached and emission maxima were obtained for the wavelength of $\lambda 280$. Thermodynamic parameters (wild- type $\Delta G^\circ_{H_2O}$ 2.53 \pm 0.45 kcal/mol, mutant $\Delta G^\circ_{H_2O}$ 1.91 \pm 0.44 kcal/mol) were calculated by plotting fraction unfolded

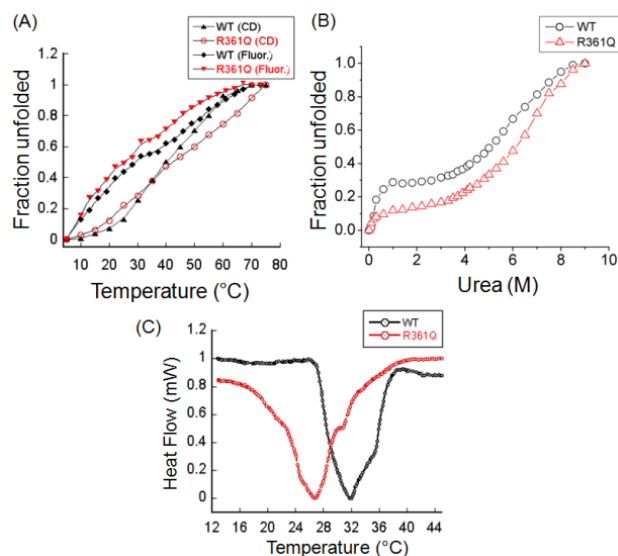


Figure 5.9: Thermal and chemical stability of ABRAXAS (6-373) wild-type and mutant. (A) Thermal and (B) chemical stability profiles of wild-type and mutant showing overlay of protein fraction unfolded at different temperature/urea concentration. (C) Differential scanning calorimetry of wild-type and mutant protein [194].

against urea concentration. Chemical stability of mutant was found different, nevertheless, both unfold through acquiring an intermediate species which was more predominant in case of wild-type (**Figure 5.9**). To substantiate these findings and re-evaluate the reversibility of unfolding of wild-type and mutant, Differential Scanning Calorimetry (DSC) was performed. A peak maxima was considered as temperature of melting and area under the curve was used to determine enthalpy contents and folding pathway. DSC data has shown overlapping of two transitions which could be corresponding to an intermediate formed during unfolding process, and displayed three state reversible folding pathway of wild-type and mutant. The observed T_m values for wild-type was significantly higher compared to mutant (wild-type T_m 32 ± 2.5 , ΔH 440 ± 11.6 kcal/mol; mutant T_m 27 ± 3.1 , ΔH 660 ± 15 kcal/mol) (**Figure 5.9**). Furthermore, the observed intermediate species was reluctant to exist in case of mutant while it was predominant in the wild-type. These results suggest that three-dimensional structural

stability of ABRAXAS is different in R361Q mutation and undergoes different folding pathways.

5.3.6 Binding interaction of RAP80 with ABRAXAS (6-373) wild-type and R361Q

mutant: ABRAXAS is a member of BRCA1 complex and is involved in direct interaction with RAP80 and BRCA1[7]. ABRAXAS binds directly to RAP80 AIR and through (pS)-P-T-F motif located at the C-terminal to BRCA1 BRCT, thus facilitating the recruitment of BRCA1 complex to the DNA damage site. Secondary structural component of wild-type and mutant remained equivalent irrespective of mutation. However, the three dimensional folding pattern and unfolding pathway indicate lesser stability of mutant. Since, structural components and folding pathway monitored at global structural level has shown differences, we speculate that the binding interaction between ABRAXAS wild-type and mutant with RAP80 would be affected. To test this hypothesis, His pull-down assay of native RAP80 (1-405) with His tagged ABRAXAS wild-type and mutant was performed. We observed that the binding of wild-type and RAP80 is significantly higher as compared to mutant (**Figure 5.10**). Dissimilar binding profile suggested that due to mutation, structure alteration hinders the association between RAP80 and ABRAXAS.

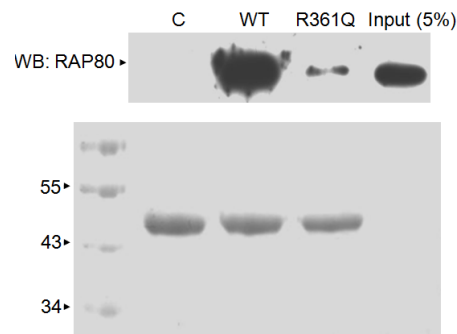


Figure 5.10: Binding analysis of ABRAXAS wild-type and mutant with RAP80. Histidine pull-down assay followed by western blotting [194].

5.4 Crystallization of ABRAXAS

ABRAXAS functional domain protein crystallization trial was set with 15 mg/ml protein concentration. The protein and mother liquor solution was mixed in 1:1 ratio (1 μ l+1 μ l) and allowed to equilibrate at 22°C with 500 μ l reservoir solution in a closed system. A clear drop or light precipitation was observed in most of the drops.

5.5 Conclusion

ABRAXAS is the key member of BRCA1 complex and acts as a bridging molecule of the complex. Knockout studies of ABRAXAS have shown defective recruitment of BRCA1 complex to the DNA damage site and hence defective DNA repair [6, 7]. Thus, ABRAXAS is a multifaceted molecule which plays an important role in cancer progression, and BRCA1 mediated homologous recombination repair.

Multiple sequence alignment of ABRAXAS in different species in phylogenic order has showed highly conserved nature of Arg 361 residue (**Figure 5.3**). Wild-type and mutant showed similar CD spectra and secondary structural composition, while the relative orientation of Trp and Tyr were slightly disturbed. This indicates that R361Q mutation brings several localized changes in structure pattern of ABRAXAS which altogether furnish a different conformational stability of overall structure. These conformational changes are very minor and hence could not be detected at secondary structural level while their relative effect was traced during three dimension unfolding pathway. The relative redundancy of intermediate species in wild-type suggests the existence of different unfolding pathway which may be partially followed by mutant, furthermore unfolding of wild-type was found to be more cooperative. Altogether, the localized changes in the mutant structure brings down its thermal and chemical stability which

further perturbs its interaction with RAP80. The cumulative global changes in mutant structure was sufficient to disturb the critical interaction necessary for BRCA1 complex integrity and localization. Therefore in the presence of R361Q mutation, ABRAXAS could not extend its bridging interaction through RAP80 which perhaps affects the recruitment of BRCA1 complex to the DNA damage site. Consequently, the nuclear retention of BRCA1 is adversely affected which further agitates G2/M checkpoint and homology-directed DNA repair (**Figure 5.11**) [11].

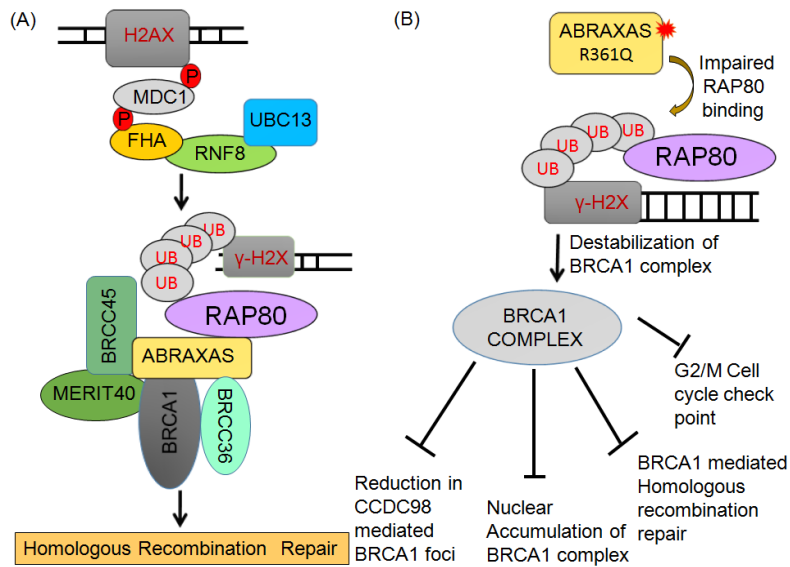


Figure 5.11: The figure illustrates the plausible mechanism of defective homologous recombination repair and other consequences due to ABRAXAS R361Q mutation [194].

Studies of Abraxas wild-type and R361Q mutation describes the role of protein folding in disease progression such as cancer. It will further explore the opportunity of inhibitor design for therapeutic application that can recompense the effect of such adverse mutation.

Chapter 6

Structural Associates of BRCA1-complex

6.1 Introduction

BRCA1 interacts with numerous molecules and consequently plays an essential role in multiple pathways such as DNA damage repair, cell-cycle and transcription regulation [2, 8, 89, 162, 163]. It acts as a tumor suppressor and has been found to be associated with hereditary breast and ovarian cancers [3, 164]. Furthermore, BRCA1 BRCT deletion mutants showed perturbed behavior towards subnuclear co-localization with H2AX, thus highlighting its importance in DNA damage repair [85, 116, 147, 165].

Protein-Protein Interactions in BRCA1 complex play very important role in its structural stability, and alteration in interactions profiles lead to destabilization of whole complex [13]. MERIT40 is an essential component of BRCA1 complex, but how it stabilizes this complex remain elusive. A literature search could not find any direct association between MERIT40 and BRCA1 that stabilizes the complex. Knockdown of MERIT40 significantly reduces the RAP80 and ABRAXAS proteins levels, consequently affecting the integrity of BRCA1-complex [13, 19]. MERIT40 down regulated cells considerably lower the stability of BRCA1-complex compared to any other members [13]. Hence, MERIT40 is a mediator protein required for stable BRCA- complex formation.

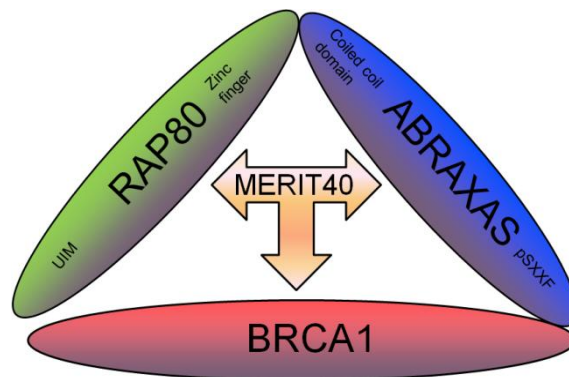


Figure 6.1: Schematic representation of members of BRCA1 complex.

MERIT40 is essential for integrity and localization of BRCA1-complex by performing multiple interactions within the complex at DSB site [13]. As, MERIT40 knockdown has a negative effect on the RAP80, BRCC36 and CCDC98 levels irrespective of BRCA1 protein level. It further demonstrates that the MERIT40 has a greater influence on the BRCA1-complex [13]. However, it remain elusive how each member form association together and bring about the stabilization of BRCA1 complex. Recombinant MERIT40, BRCA1 BRCT domain and CCDC98 were purified to establish their interacting relationship. To further explore the possibility of complex formation between MERIT40, CCDC98 and BRCA1 BRCT, GST pull-down assay and Isothermal Titration Calorimetry was performed.

The overall goal is to analyze structure of BRCA1 complex and determine the binding association among the various members

1.2 Material and Methods

In the present study, structural and binding analysis of RAP80, BRCA1, ABRAXAS and MERIT40 were carried out. The following plasmid constructs were used for the expression and purification of proteins. RAP80 (1-405), BRCA1-BRCT (1649-1859), ABRAXAS (6-409), ABRAXAS (6-373), ABRAXAS (6-301), ABRAXAS (6-268), and MERIT40 full length. Expression and purification of RAP80 (1-405) and MERIT40 have been described in chapter 3 and 4, respectively. Similarly, expression and purification of different functional domain of ABRAXAS (6-409), ABRAXAS (6-373), ABRAXAS (6-301) and ABRAXAS (6-268) have been described in chapter-5

BRCA1-BRCT was expressed in bacterial system using *E.coli* BL21 (DE3) strain. For protein expression, 100 ng/μl (1ul) plasmid construct was transformed into *E.coli* BL21

(DE3) cells and grown on LB agar plate containing antibiotic ampicillin (100 µg/ml). Single colony was inoculated in LB broth and bulk culture (10 litres) was grown, as mentioned in case of MERIT40 (chapter 4). Purification protocol of BRCA1 BRCT is the same as that of MERIT40 except the purification buffer composition.

Purification buffer composition: 300 mM NaCl, 50 mM Tris pH 7.5

FPLC buffer composition: 300 mM NaCl, 10 mM HEPES

Purified proteins were passed through gel filtration column (Superdex 200) in order to get pure and homogenous protein. They were further used in various bio-physicochemical experiments. The complete details of protocol have been discussed earlier in chapter 3 (material and methods).

6.3 Results and discussion

MERIT40, BRCA1-BRCT, RAP80 and ABRAXAS native proteins were FPLC purified and analyzed on SDS-PAGE (**Figure 6.2**). Purified proteins were gel extracted and subjected to MALDI-TOF/TOF mass spectrometry analysis for identification [197] [198]. MERIT40, BRCA1, RAP80 and ABRAXAS have shown correct identity in ExPASy server (<http://www.expasy.org/>) through mascot analysis (**Table 6.1**). It illustrates the authenticity of purified protein for further use in structural and functional analysis.

Table 6.1: Mass spectrometry profiles of MERIT40, BRCA1, ABRAXAS and RAP80 using Mascot analysis [199]

Protein	Match	Score	Nominal Mass (M _r , Dalton)	Calculated pI	Sequence Coverage (%)
MERIT40	h-MERIT40	119	37050	4.6	35.0
ABRAXAS	h-ABRAXAS	164	47033	6.6	31.5
RAP80	h-RAP80	85	80875	5.3	20.1
BRCA1	h-BRCA1	63	208000	5.0	24.0

h: Human

6.3.1 Structural Characterization: Secondary and tertiary structural characteristics are the important benchmark for defining protein folding and functional properties. A well-folded protein shows typical secondary structure signature as well as localized behavior of aromatic and hydrophobic residues. Tyrosine and tryptophan are the most extensively used intrinsic fluorophores to study the micro-environmental changes induced in the protein due to external stimulus such as temperature and chaotrophs [187]. In a folded protein, these residues are generally buried inside the hydrophobic core and are extremely sensitive to modification around their micro-environment.

In order to evaluate the secondary structural characteristics of MERIT40, RAP80, ABRAXAS and BRCA1, far-UV CD spectra were recorded. Secondary structural component of purified

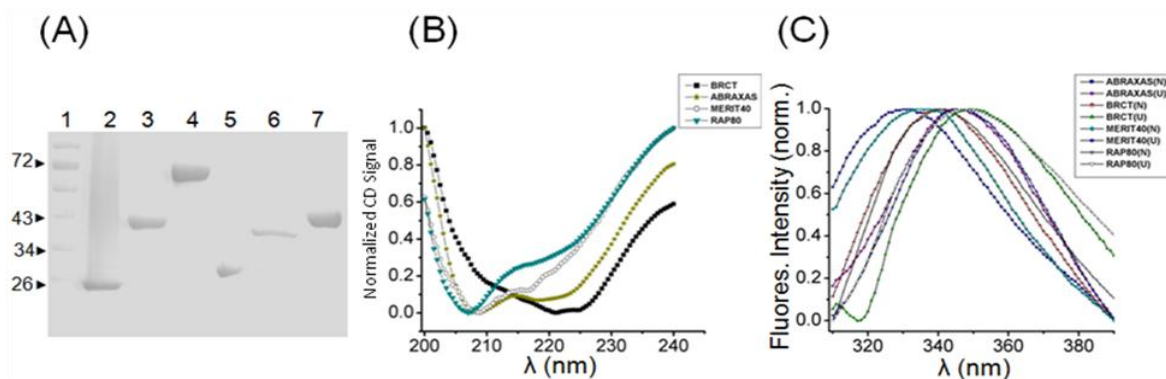


Figure 6.2: Protein purification and structural characterization of ABRAXAS, BRCA1-BRCT, MERIT40 and RAP80. (A) Lane 1-marker, purified proteins in lane 2-BRCA1-BRCT, 3-MERIT40, 4-RAP80, 5-ABRAXAS (6-268), 6-ABRAXAS (6-301) and 7-ABRAXAS (6-373). (B) Overlay of far-UV CD spectrum and (C) emission maximum of BRCA1-complex proteins. N-native, U-unfolded with urea [199].

protein was analyzed using DichroWeb server (<http://dichroweb.cryst.bbk.ac.uk>) [186, 200-203]. In MERIT40, percentage of α -helices and β -sheets were found to be 20 and 16.6% while BRCA1 BRCT showed it 22 and 27.1%, respectively (**Figure 6.2B**). RAP80 and ABRAXAS showed the presence of fewer α -helices (21.1 and 15%), and displayed

the prevalence of β -sheets (30.7 and 24%) (**Figure 6.2B**). Tertiary structures of purified proteins were evaluated by monitoring positions of Tyr and Trp residues. Emission spectrum of native and unfolded protein was monitored. Native MERIT40, RAP80, ABRAXAS and BRCA1 showed emission maxima at λ =338, 340, 332 and 339 nm respectively, and these maxima undergo substantial red shift when the protein is unfolded (**Figure 6.2C**). It indicates Trp and Tyr residues are buried in three dimensional hydrophobic environments [187]. Altogether, these results also deciphered a well folded state of MERIT40, RAP80, ABRAXAS and BRCA1.

6.3.2 Qualitative interaction analysis of MERIT40 with BRCA1, RAP80 and ABRAXAS: Pull-down interaction analysis suggested direct interaction of MERIT40 with BRCA1-BRCT or ABRAXAS (**Figure 6.3**). However, we did not observe significant binding between MERIT40 and RAP80, indicating their interaction is either absent or highly transient.

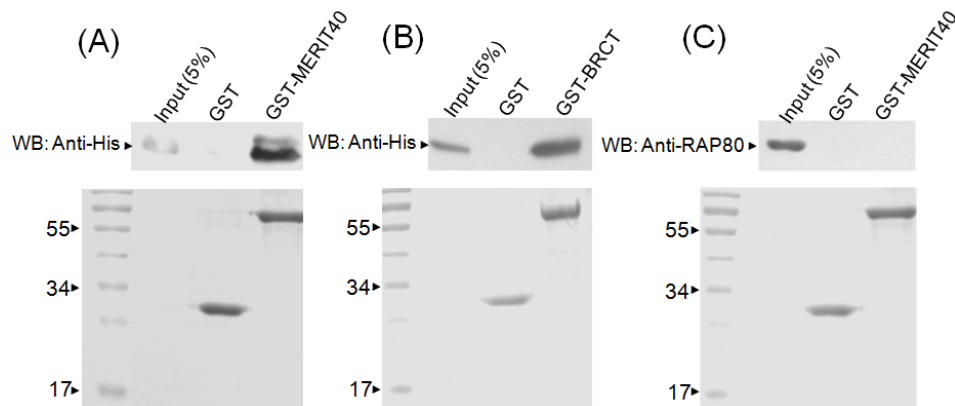


Figure 6.3: Pull-down assay to determine direct binding partner(s) of MERIT40. (A) ABRAXAS, (B) BRCA1-BRCT and (C) RAP80. Lower panel showing the loaded bait protein on SDS-PAGE stained with ponceau. Western blotting was performed against the his-ABRAXAS (prey), his-MERIT40 (prey) and RAP80 (prey), respectively [199].

6.3.3 Quantitative Interaction analysis of MERIT40 and BRCA1: To establish role of MERIT40 in stability of BRCA1-complex, interaction study between MERIT40 and

BRCA1 was performed using MicroCal ITC 200. All the samples were degassed thoroughly (30 min) before proceeding to the titration reactions, and a control/standard was run before performing the individual experiment. 30 μM BRCA1 BRCT in cell was titrated with 300 μM MERIT40 in the syringe, the reaction was allowed to proceed at 25°C. A total of 22 injections (spacing of 2 seconds) were given with constant stirring at 1000 rpm to achieve equilibrium in the system. The interaction between BRCA1-BRCT domain and MERIT40 was found with an affinity constant of $K_d = 521 \pm 7 \mu\text{M}$. Hence, MERIT40 mediates a direct interaction with BRCA1-BRCT, however, the affinity of interaction was found to be weak (**Figure 6.4**).

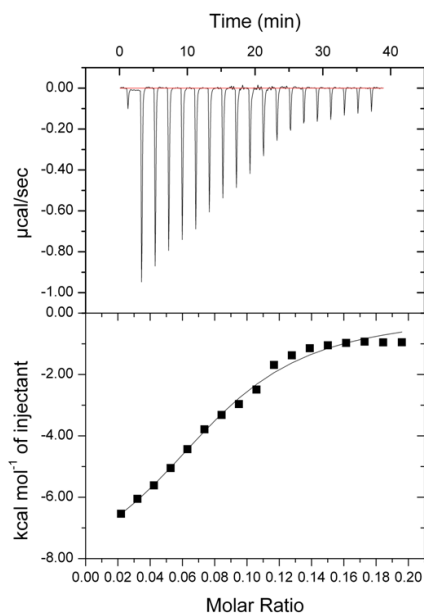


Figure 6.4: ITC thermogram for interaction analysis between MERIT40 and BRCA1-BRCT [199].

6.3.4 Interaction analysis of MERIT40 and ABRAXAS: Interaction analysis of MERIT40 with ABRAXAS (6-409) was performed using Isothermal Titration Calorimetry. MERIT40 was titrated with ABRAXAS (10-fold molar ratio) at constant temperature and heat changes were determined.

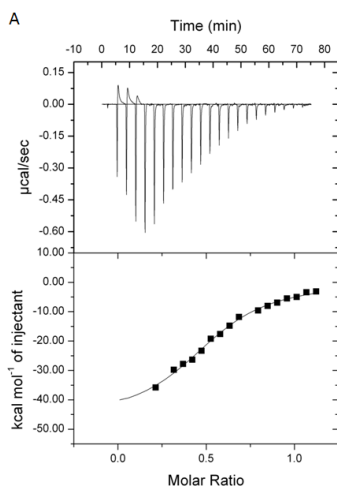


Figure 6.5: Interaction analysis of MERIT40 with ABRAXAS using Isothermal Titration Calorimetry [199].

MERIT40 was taken as titrant (30 μM) and ABRAXAS (290 μM) as titre in a Microcal ITC200. The reaction was allowed to proceed at 25°C with a stirring speed of 1000 rpm. A total of 23 injections were given (each 2 μl) with an equilibration time of 4 seconds. MERIT40 showed direct interaction with ABRAXAS as indicated with significant heat change during reaction. The binding affinity between MERIT40 and ABRAXAS was found $K_d = 2.5 \pm 0.24 \mu\text{M}$ (**Figure 6.5**). Pull-down assay also confirmed the interaction between MERIT40 and ABRAXAS (**Figure 6.5**).

6.3.5 Determining the minimal region of ABRAXAS binding to MERIT40: In order to determine the minimal binding region of ABRAXAS, isothermal titration calorimetry was carried out using different deletion regions of ABRAXAS, such as (6-373), (6-301) and (6-268) amino acids. ITC of different ABRAXAS fragment was performed with full length MERIT40. Interaction profile of MERIT40 and ABRAXAS (6-373) suggests that ABRAXAS C-terminal region might be involved in interaction with MERIT40 with a binding affinity $K_d = 1.7 \pm 0.28 \mu\text{M}$ (**Figure 6.6**). To further narrow down the minimal

binding domain of ABRAXAS with MERIT40, we have performed interaction analysis with shorter region 6-301, 6-268 amino acids. However, no binding was observed between MERIT40, and ABRAXAS (6-301), (6-268) amino acids. From these experiments we conclude that the BRCA1 BRCT binding domain of ABRAXAS lies in the C-terminal region, while N-terminal of ABRAXAS might be mainly involved in interaction with RAP80 (**Figure 6.6**) [7].

6.3.6 Interaction analysis of MERIT40 and RAP80: RAP80 is an upstream player

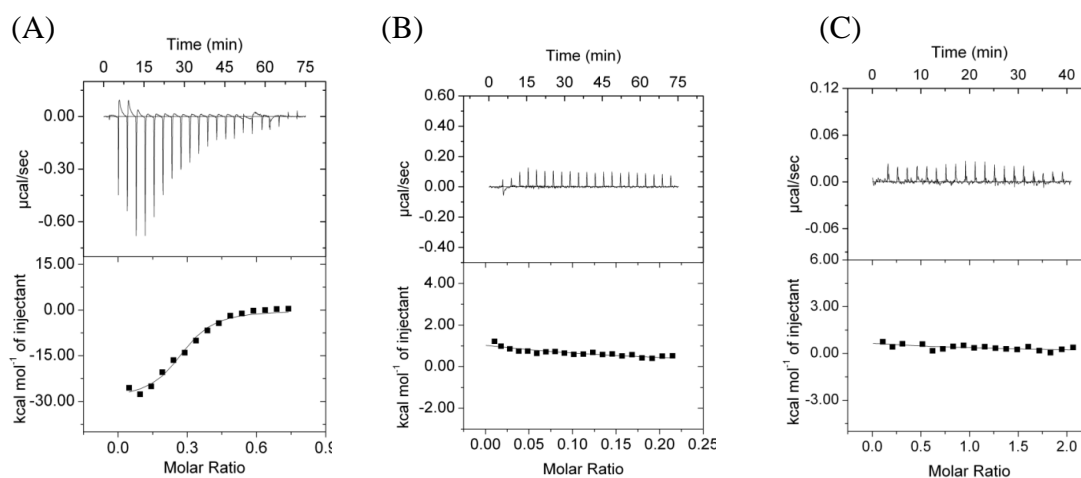


Figure 6.6: Isothermal titration calorimetry of MERIT40 with (A) ABRAXAS (6-373), (B) ABRAXAS (6-301), (C) ABRAXAS (6-268) amino acids [199].

in BRCA1-complex, hence it would be interesting to observe the binding between RAP80 and MERIT40. Surprisingly, MERIT40 did not show interaction with RAP80 in isothermal titration calorimetry (**Figure 6.7**).

6.4 Co-crystallization of MERIT40 with ABRAXAS and BRCA1 BRCT

For co-crystallization, both the proteins were mixed in equal proportion (0.8 mM) and crystallization trials were set. The protein and mother liquor solution was mixed in 1:1 ratio (1 µl+1µl) and allowed to equilibrate at 22°C with 500 µl reservoir solution in a closed system. A clear drop or light precipitation was observed in most of the drops.

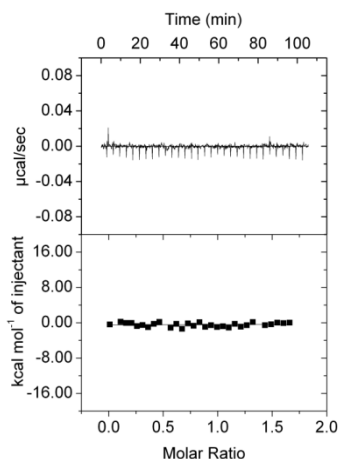


Figure 6.7: Isothermal titration calorimetry of MERIT40 with RAP80 (1-405) amino acids [199].

6.5 Conclusion

A direct interaction of BRCA1-BRCT and ABRAXAS with MERIT40 was observed which list MERIT40 as an essential molecule in BRCA1-complex. It perhaps set up an interaction network in BRCA1 complex which is being utilized for stabilization of ABRAXAS, since knockdown of MERIT40 significantly reduces the ABRAXAS and RAP80 levels (**Figure 6.8**)

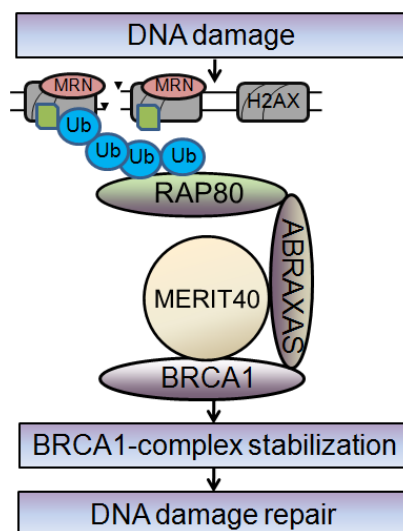


Figure 6.8: Anticipated model of MERIT40 mediated BRCA1-complex stabilization [199].

[13, 19]. Stabilization of ABRAXAS further helps in maintaining the structural integrity of BRCA1-complex. This study will provide insights into the diverse interactions involved among various members which are essential for DNA repair function of BRCA1 complex. These findings will be further helpful in understanding functional interplay of DNA damage repair proteins and mechanisms that abrogates the synergy of the protein complex, consequently, its DNA damage repair function.

Chapter 7

Structural Studies of BRCA1-KIF16 Complex

7.1 Introduction

Most of the intracellular trafficking occurs through the involvement of motor proteins such as kinesins. The kinesins belong to the super family of microtubule-stimulated ATPases, and are characterized by a conserved motor domain of ~350 amino acids. The Kinesin comprises different functional domains including microtubules interacting domain at C-terminal and adenine nucleotides interacting domain at N-terminal [204, 205]. Most kinesins utilize ATP-dependent motors to perform transport function of intracellular cargo using microtubules. These mechanochemical ATPases generate force to transport membrane bound vesicles, organelles or chromosomes towards the plus end of microtubules [206-209]. On the basis of sequence variability in the motor domain, kinesins can be classified into different families. Members of the same subfamily have highly conserved motor domains, particularly, at their extreme ends although the tail domain tends to be more different (**Figure 7.1**). The variable tail domain may specify the cargo molecule or organelle that is transported by a particular Kinesin Family member (KIF) [210]. KIF1 is one of the family and its most important member, KIF1B, was originally anticipated to be a motor for moving mitochondria toward microtubule positive end [211]. The *kif1b* gene was found to be located within a distal region of chromosome 1p, where loss of heterozygosity (LOH) is frequently observed in case of human neuroblastomas [212]. *kif1b* encodes as many as eight different isoforms that fall into two general classes of molecular weights: ~130,000 (KIF1Bp130) and ~204,000 (KIF1Bp204). The KIF1Bp130 class corresponds to the originally described KIF1B which binds to mitochondria through its unique C-terminal domain [213]. Mitochondria are the main cargo for KIF1b in addition to some other entities which remain to be

discovered [211]. Furthermore, a number of KIFs which are primarily found in neurons, are believed to participate in anterograde transport by carrying synaptic proteins and organelles from the neuronal soma [214]. Interactions with cargoes may also be regulated either by other kinesin-associated proteins or by phosphorylation of the polypeptide chain [215, 216]. The phosphorylation dependent interaction makes KIF1b a suitable candidate to study its role in DNA damage repair and cell -cycle regulation. It contains a pS-X-X-F motif which is commonly found in the BRCT domain interacting proteins.

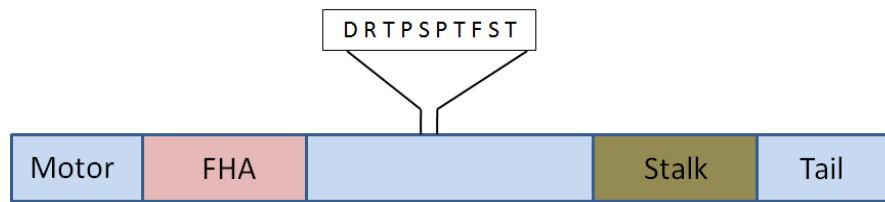


Figure 7.1: Schematic representation of the structure of KIF1b.

The BRCT domain is found in different proteins, including RAD9 [217], 53BP1[218], *Crb2* [219] and RAP1[220], and it possesses phosphopeptide-binding modules through which it functions in cell -cycle regulation , DNA-damage checkpoint control and DNA repair [10, 21]. Association of the sixth BRCT domain of TOPBP1 with E2F1 was shown to be regulated by phosphorylation [221]. BRCA1 BRCT phosphopeptide binding domains preferred aromatic amino acids at the position (+1) and hydrophobic or aromatic residues at the position (+2) [21]. However, towards the C-terminal to the phosphopeptide, phenylalanine at position (+3) is strongly preferred, hence this consensus sequences containing proteins are potential reported binding partners of the BRCA1 BRCT. This domain preferred a [YE]-E-[TV]-pS-[YFH]-[VTYF]-[FY] motif [21]. Different structures have already been reported in the Protein Data Bank in which BRCA1-BRCT showed its specificity towards (pS)-X-X-F motif containing protein such

as CTIP, BACH1 [222, 223]. Moreover, other BRCT such as MDC1 BRCT also showed its specificity to the same consensus sequence [55]. Several structures of different BRCT domains in complex with phosphopeptide have been reported in the Protein Data Bank. For example, BRCA1 BRCT – BRCA1-Binding Helicase-Like Protein (BACH1; 1T29) complex structure, complex structure of BRCA1 BRCT and CtBP interacting protein (CtIP; 1Y98), structure of BRCA1 BRCT in complex with acetyl-CoA carboxylase 1 (ACC; 3COJ) and BRCT –tetrapeptide complex structure (3K0K) [91] [55] [224] [225]. Presence of similar domain in KIF1b prompted us to explore the binding mechanism with BRCT domain.

BRCA1 has two BRCT repeats at its C-terminal through which it regulates different functions such as DNA damage repair, cell cycle regulation, gene transcription and apoptosis [9, 63, 73, 226]. We were particularly intrigued by the possibility of BRCT domains mediating phosphorylation-dependent interactions, because protein phosphorylation by kinases is known to trigger the assembly of phosphorylation-dependent signaling pathways critical for cell growth and survival [227]. Different proteins that bind to phosphopeptides have been identified, including the BARD1, MDC1, XRCC1 [225, 227-230]. Each of these domains recognizes a subset of phosphorylated proteins which contain SPTF motif. From the synthetic peptide library search, it has been observed that KIF1b has pS-X-X-F containing motif [21]. An attempt was made to crystallize and determine the complex structure of BRCA1 BRCT domains in association with (pS)PTF containing phosphopeptide motif of KIF1b. A synthetic peptide having the sequence [NH₂-DRTP(pS)PTFST-COOH] was used to determine the

binding affinity and mechanism of interaction between the KIF1b and BRCA1 BRCT (Figure 7.2).

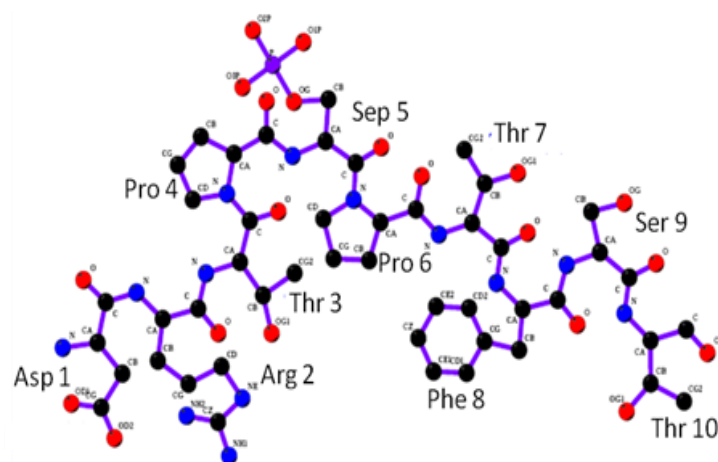


Figure 7.2: Schematic representation of Kif1b peptide used in the binding and structure determination study. SEP represents the phosphorylated serine.

7.2 Material and methods

7.2.1 Expression and Purification of BRCA1 BRCT domain: The pGEX-kT BRCA1 BRCT (1646-1859) construct was transformed into the bacterial expression host such as *E. coli* BL21 (DE3) strain, and the fusion protein was expressed using the methodology described in chapter 2. Expressed fusion protein was purified using GST affinity chromatography, and the GST tag was removed by treating the fusion protein with TEV protease. The purified protein was further subjected to size exclusion chromatography (AKTA- FPLC) Superdex- 200 column to get homogenous population. The purified BRCT domain was analysed by mass spectrometry to determine its exact molecular weight and identity.

7.2.2 Interaction analysis using ITC: The isothermal titration calorimetry reaction was performed by using *Microcal ITC 200* (GE, Sweden). The FPLC purified BRCA1 BRCT protein (0.02 mM) was titrated against the KIF1b phosphopeptide (0.200 mM). Both the protein and peptide were in same buffer containing 300 mM NaCl, 50 mM Tris, and pH-

7.5. BRCA1 BRCT protein was in sample cell while peptide solution from syringe was injected (2 μ l each, total 16 injection) into the sample cell which was maintained at 25°C temperature by constant stirring at 1000 rpm (**Table 7.1**). The dissociation constant was calculated by fitting the sigmoidal curve using Origin software (Version 7.2). The heat of dilution was determined by titrating the peptide with the buffer and this was used for blank correction.

Table 7.1: ITC reaction parameters

Experimental parameters	
Total injections	16
Cell Temperature (°C)	25
Ref power (μ cal/sec.)	5
Initial Delay (sec.)	250
Syringe conc. (mM)	0.2
Cell conc. (mM)	0.02
Stirring speed (RPM)	1000
Injection parameters	
Injection volume (μ l)	2
Duration (sec.)	40
Spacing (sec.)	210
Pipette volume (μ l)	~38
Filter period (sec.)	5

7.2.3 Protein crystallization: FPLC purified BRCA1 BRCT protein at a concentration of 25 mg/ml was used for setting crystallization trials as described in chapter 2. The concentrated protein was mixed with the synthetic KIF1b peptide (10 mg/ml) to obtain a 1:1.5 molar ratio concentration. This mixture was incubated overnight at 4°C and centrifuged before setting up for crystallization trial. Crystallization trials were set using commercially available crystallization kits (Hampton Research, Inc) using sitting and hanging drop methods. Initial crystallization drop volume of 2 μ l containing 1 μ l protein

and 1 μ l mother liquor mixture was equilibrated through vapour diffusion against 0.5 ml of precipitant solution in the reservoir at 22°C. This crystallization conditions were observed at different time points, and further optimised by varying the concentration of precipitant and salt. The crystals were confirmed as protein crystals by staining with IZIT dye from Hampton Research.

7.2.4 Diffraction data collection and processing: The BRCA1-KIF1b complex crystal was cryo-protected using 30% (v/v) glycerol as cryo-protectant prepared in mother liquor solution, and thereafter crystals were flash frozen in liquid nitrogen before exposure to X-rays. Diffraction data were collected using rotating anode X-Ray Generator (Bruker) and Image plate (MAR Research), operated at 50 kV and 100 mA (Macromolecular X- diffraction facility at ACTREC). The crystal diffracted to about 3.5 Å resolution. A total of 237 oscillation frames were collected at 100 K with the detector set at a distance of 250 mm. The diffraction data were processed using iMOSFLM [231] software, and scaled using SCALA program from CCP4 suite [232]. V_M and filling map was calculated and solvent content was estimated [233].

7.2.5 Structure Solution of BRCT-KIF1b structure: Crystal structure of the BRCT-KIF1b complex was solved by Molecular Replacement (MR) using Phaser program [234]. The BRCA1 BRCT-CtIP complex structure (PDB ID-1Y98) was acquired from the Protein Data Bank, and was processed for removal of water and ligand molecules from the coordinate file, and was thereafter used as the search model [235]. The rotational and translation searches were carried out using PHASER software [236] and the MR solution was further refined by REFMAC 5 of CCP4 suite [232, 237]. Initially the model was refined as rigid body, and later rounds were of restrained refinement [237].

Table 7.2: Summary of data collection, processing and refinement statistics

Parameter	Overall Value (highest resolution shell)
No. of crystals used	1
X-ray source	Home source (ACTREC)
Wavelength (Å)	1.54179
Crystal to detector distance (mm)	250
Space group	P2 ₁ 2 ₁ 2 ₁
Unit cell parameters (Å)	a=84, b= 180.5, c=194 $\alpha=\beta=\gamma=90^\circ$
V _m (Å ³ /Daltons) value	3.8
Total no. of residues	207
Mosaicity (°)	1.14
No. of molecule in asymmetric unit	8
Resolution limit (Å)	76.0-3.5 (3.69-3.50)
Total no. of reflections measured	218448
Unique reflections	32513
<Multiplicity>	6.7 (6.3)
<I/ σ (I)>	3.7 (1.1)
Completeness (%)	86
R _{merge} (%)	38.0 (147.7)
Refinement statistics	
R _{work} (%)	27.5
R _{free} (%)	31.8
Total no. of atoms used in refinement	14128
RMSD bond length (Å)	0.0089
RMSD bond angle (°)	1.3054
Ramachandran plot	
Most favoured (%)	92.7
Additionally allowed (%)	6.1
Disallowed region (%)	0.8

A very tight Non Crystallography Symmetry (NCS) - restraints was imposed throughout the refinement (local NCS parameters: sigx 0.05, Maximum distance 4.2 Å, maximum difference of distances 1.0 Å). R_{free} was calculated as for R_{work} but only 5% data left out of refinement procedure has been used in the calculations. In each cycle of refinement figure of merit (FOM), R_{free} and R_{work} was observed as the criterion for the selection of improved structure. Furthermore, electron densities of peptide and protein were monitored during each refinement to evaluate the improvement in the refined atomic model.

7.3 Results and discussion

7.3.1 Purification and Characterization of BRCA1 BRCT domain: The BRCA1 BRCT domain was successfully cloned into pGEX-kT vector and expressed in *E. coli* BL21 (DE3) strain. The protein, purified by affinity chromatography using GST sepharose 4B (Amersham) resins, showed a single band on SDS-PAGE (**Figure 7.3**). On AKTA-FPLC, size exclusion column (Superdex -75), BRCA1 BRCT protein eluted at the expected elution volume and has shown the secondary peak corresponding to higher oligomer of BRCA1 BRCT. FPLC purified protein has shown a single band around 24 kDa on SDS PAGE (**Figure 7.3**). The mass profile showed a single peak which confirms homogeneity of the sample, and experimental molecular weight matched the expected value of 24660 Daltons.

7.3.2 Binding isotherm of BRCA1 BRCT domain and KIF1b peptide: The enthalpy change on interaction of BRCA1 BRCT with phosphopeptide of KIF1b was found to be sigmoidal in nature (**Figure 7.4**). It has been observed that KIF1b phosphopeptide exothermically interacts with BRCA1 BRCT domain, and the binding affinity (K_d) of this

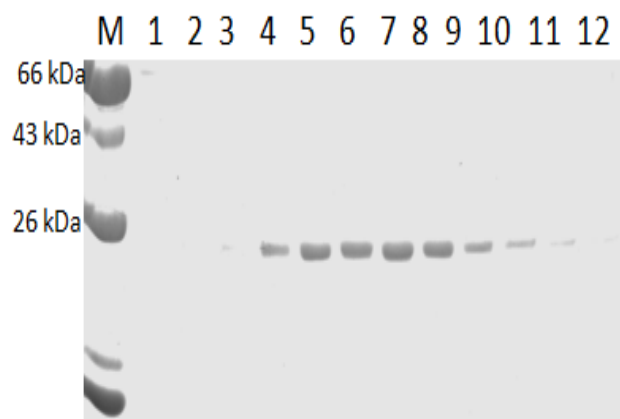


Figure 7.3: SDS-PAGE showing the FPLC purified protein BRCA1 BRCT interaction was calculated to be 2.83 μM . So far the highest binding affinity is reported for BACH1 peptides having a K_d value of 0.9 μM . The weakest affinity is shown by the Acetyl CO-A peptide having a dissociation constant K_d 5.2 μM (**Table 7.3**). This suggests that KIF1b is one of the potential binding partner of BRCA1 BRCT.

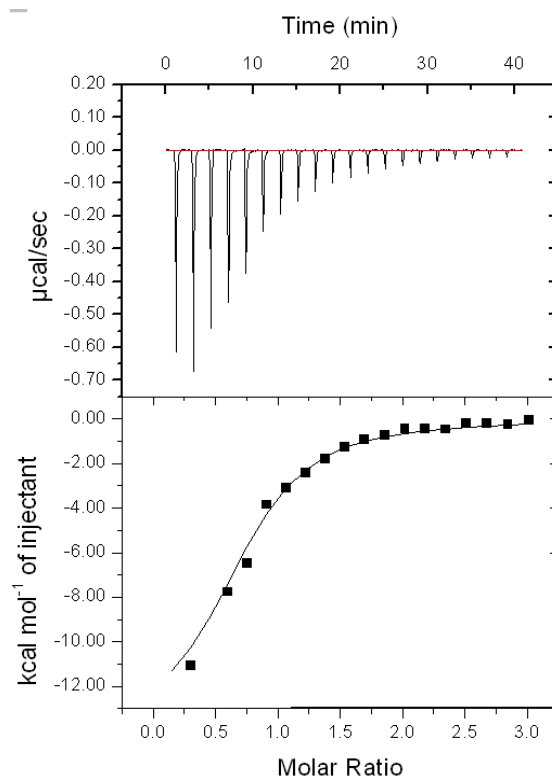


Figure 7.4: Binding isotherm of BRCA1 BRCT-KIF1b

Table 7.3: Binding comparison of various BRCA1 BRCT binding peptides

Complex	Peptide sequence	Dissociation constant (K_d)
BRCA1 BRCT-KIF1b	NH2-DRTP(pS)PTFST-COOH	2.83 μ M
BRCA1 BRCT-ACC1	NH2-DSPPQ-pS-PTFPEAGH- COOH	5.2 μ M
BRCA1 BRCT-BACH1	NH2-ISRSTpSPTFNKQ-COOH	0.9 μ M
BRCA1 BRCT-CtIP	NH2-PTRVSpSPVFGAT-COOH	3.7 μ M
BRCA1 BRCT-Tetrapeptide	NH2-pSPTF-COOH	3.7 μ M

7.3.3 Crystallization and X-ray diffraction of BRCA1 BRCT domain: Crystals of the complex molecule were obtained by co-crystallization under following conditions: 0.2 M ammonium sulphate, 30% PEG 5000 MME, 0.1 M MES pH 7.5 using Hampton crystal screen. Purified protein produced hexagonal shape crystal during crystallization.

7.3.4 Structure solution of BRCA1 BRCT and KIF1b complex: Based on intensity equivalents, systematic absences and MR calculations, it has been observed that the crystal belongs to the space group $P2_12_12_1$ with unit cell parameters: $a = 84.01 \text{ \AA}$, $b = 180.47 \text{ \AA}$, $c = 194.22 \text{ \AA}$, $\alpha = \beta = \gamma = 90^\circ$. The data processing statistics is given in **Table 7.2**. The useful data extend to 3.5 \AA resolution, with a completeness of 86%. Calculated V_M value suggested presence of 8 molecules in the asymmetric unit. Molecular Replacement generated a single solution characterized by a TFZ score of around 10 and a log likelihood gain (LLG) score of about 105 [234]. The solution could be refined to obtain reasonably low value for crystallographic R-factor of 0.288 (**Table 7.2**). R_{free} value of 0.315 indicates good model fitting in the electron density at this resolution. The electron density map showed prominent electron density near Gly1656 and Ser1655

residue in the **Fo-Fc** difference map, confirming that the structure is of a molecular complex between BRCA1-BRCT and KIF1b peptide. Also, in the crystal structure, there is clear difference electron density for the KIF1b oligopeptides bound to all the eight subunits. In the complex, eight of the ten residues of the ligand could be built successfully in the electron density map and refined satisfactorily. However, the terminal residue side chain showed weak electron density which may be due to disordered regions (**Figure 7.5**). The complete refinement statistics are given in **Table 7.2**. The major interactions between the KIF1b peptide and the BRCA1 BRCT domain are hydrogen bonds between phosphoserine 5 from the peptide and residues Ser1655 and Gly1656 from the BRCA1 BRCT domain (**Figure 7.5**). The nitrogen atom of Gly1656 and O γ of Ser1655 form the hydrogen bonds with second oxygen of phosphate group of peptide Ser 5. Third oxygen of the phosphate group of peptide is involved in hydrogen bonding with side chain nitrogen atom of Lys1702. Similarly, carbonyl group of the Arg 1699 extend its hydrogen bonding through side chain carboxylic group of peptide Phe 8. The protein

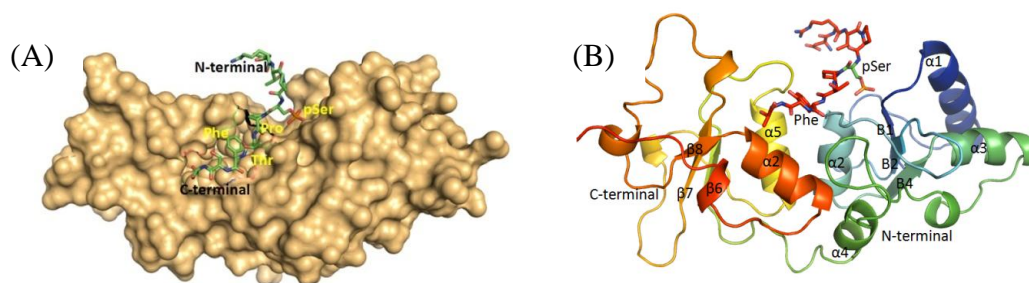


Figure 7.5: Structure solution of BRCA1 BRCT-KIF1b complex. (A) surface view of BRCT molecule along with KIF peptide. (B) Single chain BRCT with peptide.

peptide complex is stabilized by different hydrophobic interactions Val 1741, Met 1775, Phe 1704, Arg 1835 of protein with Phe 8 of peptide. There is also the hydrophobic

bonding contribution of Asn 1774 of protein with Pro 6 and Thr 7 of peptide (**Figure 7.6**) which indicate that the structure has been determined to good accuracy at this resolution.

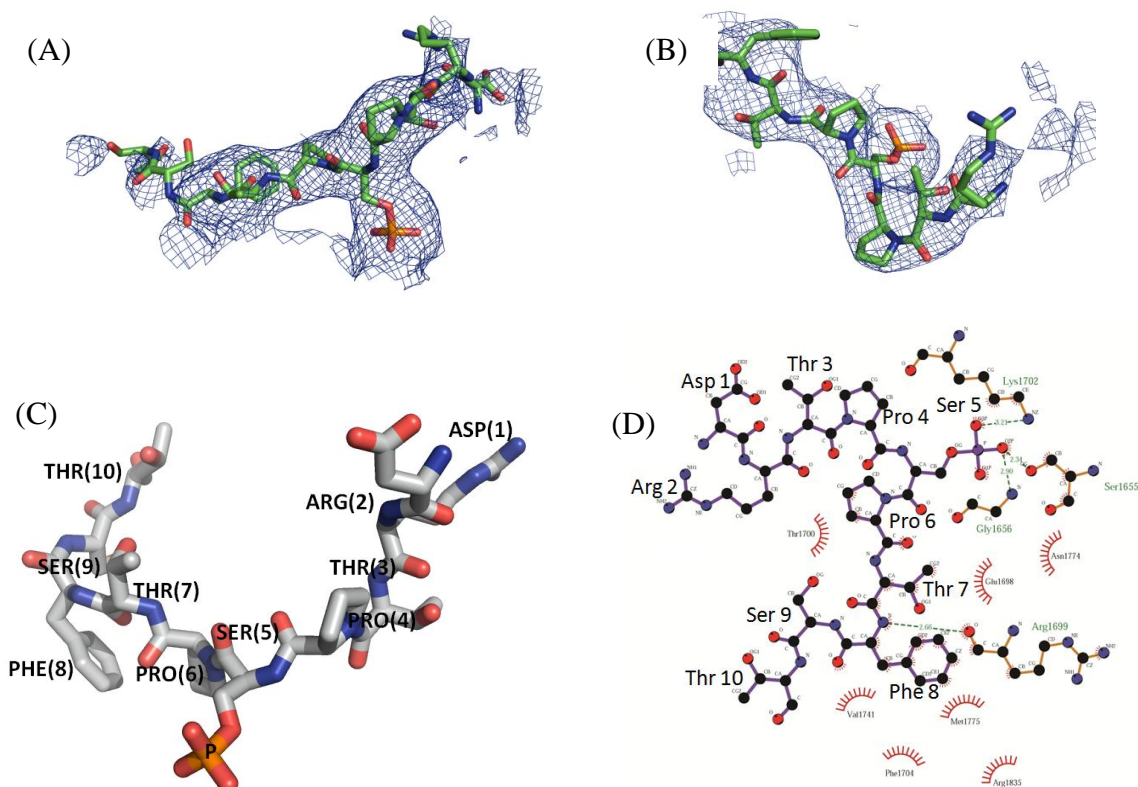


Figure 7.6: (A) and (B): Peptide showing electron density map ($F_o - F_c$, sigma cut off 1.0). (C) Stick representation of conformation of peptide. (D) Interaction analysis of BRCA1-BRCT with peptide KIF1b using LigPLOT. hydrogen bonds are shown as dashed lines, and hydrophobic interactions as arcs with radial spokes. This figure was made using LIGPLOT [238]

7.3.5 Structure analysis and validation: The overall quality and reliability of the structure was assessed with respect to energy and stereochemical geometry (overall quality factor (ERRAT): 93.86, Verify_3D (of 93.86% residues) score>0.2) [239]. The Ramachandran plot showed 99.0% of the residues were present in the allowed region [169] (**Figure 7.7**) [170, 171]. Verify_3D [239] profile revealed the average score 0.2 which is significant. Furthermore, structural geometry was evaluated using Mol Probiy web server (<http://molprobiy.biochem.duke.edu/>) which showed desired parameters in required range. There was no C_β deviation $>0.25\text{\AA}$ found in the structure. Similarly,

other structural pattern such as bad contact was around negligible. Values of some structural parameters are as follow, Wilson B-factor=64.3, Overall B-factor=47.47, Overall CC 1/2 value =0.966. Real space co-relation coefficient was 0.736463 for the peptide KIF1b. This indicates that the atomic parameters are in very good agreement with expected stereochemistry.

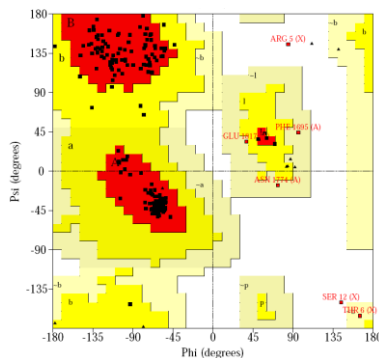


Figure 7.7: BRCA1 BRCT-KIF1b Ramachandran plot (one monomer of asymmetric subunit)

7.3.6 Secondary and tertiary structure analysis of BRCA1 BRCT - KIF1b

complex: The BRCA1 BRCT has two tandem repeats and consists of six α -helices and eight β - strands. The sequence of secondary structure followed in the BRCTs as $\beta 1$ - $\alpha 1$ - $\beta 2$ - $\beta 3$ - $\alpha 2$ - $\beta 4$ - $\alpha 3$ - $\beta 5$ - $\alpha 4$ - $\beta 6$ - $\alpha 5$ - $\beta 7$ - $\beta 8$ - $\alpha 6$. Each BRCT is made up of three parallel β -strands and a single anti-parallel β -strand. The KIF1b peptide is present in the groove of BRCT domain, in which the phosphorylated Serine residue of KIF1b interacts with the N-terminal BRCT whereas C-terminal BRCT is involved in interaction with phenylalanine residue of the peptide (**Figure 7.8**).

7.3.7 Structural alignment of different chains of BRCA1 BRCT and peptide of asymmetric unit:

There are eight molecules present in the asymmetric unit of BRCA1 BRCT-KIF1b complex structure. The average of the root mean square deviation (RMSD) between aligned structure of different chains of BRCTs present in the asymmetric unit is

0.09 Å (**Figure 7.9**) [240] over 203 Cα atom pairs, suggesting BRCT protein conformation remains same in different chains. Structural superposition of the eight of KIF1b ligands of asymmetric unit show greater variation (RMSD= 2.13 Å) towards its

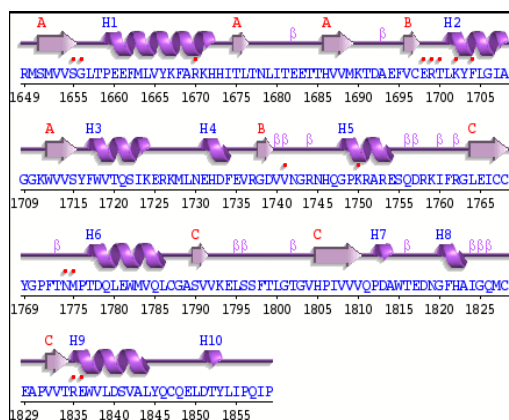


Figure 7.8: schematics representation of secondary structural component of BRCA1 BRCT-KIF1b.

terminal. However, the central tetrapeptide region, SPTF, superpose with the low RMSD of 0.274 Å (**Figure 7.9**) [31]. The peptide region which is involved in binding with the BRCT domain, particularly SPTF motif, thus maintains similar structural conformation.

7.3.8 Structural alignment of BRCA1 BRCT and KIF1b with reported structures:

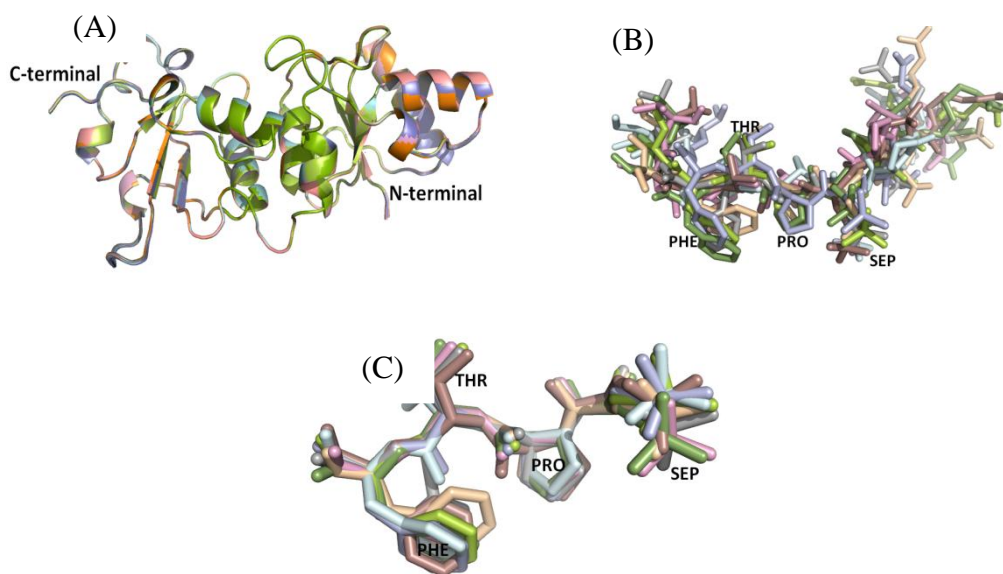


Figure 7.9: (A): Structural superposition of different chains of BRCA1-BRCT in asymmetric unit, without peptides. (B): Structural superposition of different chains of KIF1b in asymmetric unit, without BRCT. (C): Structural superposition of different chains of KIF1b (SPTF motif only) in asymmetric unit, without BRCT.

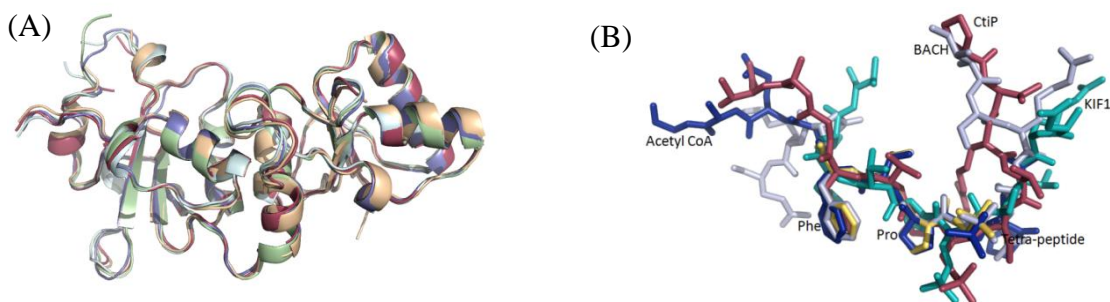


Figure 7.10: (A): Structural overlay of BRCA1-BRCT of the various reported complex with solve structure of BRCA1-BRCT, without peptides. Reddish brown, BRCA1 BRCT (PDB: 1JNX); Light Green, BRCT-BACH1 (PDB:1T29); Blue, BRCT-CtIP (PDB:1Y98); Golden, BRCT-ACOA (PDB: 3COJ), Light blue BRCT-tetrapeptide (PDB: 3K0K). (B): Structural overlay of reported BRCA1-BRCT bound peptides and KIF1b peptide

The RMSD between aligned structure of BRCA1 BRCT (PDB: 1JNX); BRCT-BACH1 (PDB:1T29); BRCT-CtIP (PDB:1Y98); BRCT-ACOA (PDB: 3COJ), BRCT-tetrapeptide (PDB: 3K0K) with solved structure showed a value of 0.34 Å (**Figure 7.10**) [240] over 203 Cα atom pairs, suggesting BRCT protein conformation largely remains unaltered upon ligand binding. Interestingly, central region of superpose peptides revealed a similar structural architecture towards the C-terminus of peptide and particularly between the residues proline and phenylalanine (**Figure 7.10**). The peptide region which is involved in binding with the BRCT domain remains structurally conserved, indicting a common mechanism of binding which might be critical for binding of phosphoserine specific ligands.

7.4 Conclusion

There are reports that numerous residues in the BRCA1 BRCT repeats are mutated in breast cancer cells [241, 242], However only reported pathogenic mutation could disrupt the interaction of BRCT domains with phosphorylated protein target thereby causing disease phenotypes. The KIF1b protein contains the signature sequence pS-X-X-F which is considered as a potential binding region for BRCA1-BRCT domain. There are

different BRCT domains containing proteins, MDC1[225], BARD1 [243], PARP1 [244] and XRCC1 [245], etc. which recognize phosphospecific binding partner [9]. BRCT domains, as phosphopeptide-interacting motifs, assist the congregation of proteins in complexes involved in signaling pathways that sense the activation of DNA damage and checkpoint kinases [9]. In the present studies, we found that KIF1b peptide binds to inter-BRCT domain with an affinity of 2.83 μ M. The stoichiometry of binding to BRCT domain with peptides was found to be 1:1. The crystals of BRCT-KIF1b complex belong to the space group $P2_12_12_1$, and unit cell contains eight molecules per asymmetric unit. The complex structure of BRCA1 BRCT and KIF1b peptide was determined at 3.5 Å resolution. The pSer (0) and Phe (+3) in the bound peptide form the main centers of interaction with the BRCT domain. pSer interacts with N-terminal of BRCT through number of hydrogen bonds, while the Phe interacts with interface region and the C-terminal BRCT through hydrophobic interactions. pSer residue forms the hydrogen bond with Ser 1655 and Gly1656, and Phe is also involved in the hydrogen bonding with Arg1699, thereby stabilizing the complex. Phosphopeptide superposition reveals a common structural scaffold for the SPTF motif of peptide bound to the BRCT domain. Furthermore, pSer and Phe are the most common conserved residues involved in binding with BRCT. This provides a common mechanism of BRCT binding to its peptide and highlights the importance of mutations. The crystal structure of BRCA1 BRCT-KIF1b complex provides new insight into the molecular interactions of BRCA1 BRCT with various phospho-peptide and their binding specificities.

Chapter 8

Summary and future prospects

BRCA1 interacts with different molecules through its functional domains and is responsible for multiple functions like cell-cycle regulation, DNA damage repair and transcription [89]. BRCA1 BRCT interacts with different proteins via their (pS)-X-X-F motifs and forms three distinct complexes-the BRCA1 A, BRCA1 B and BRCA1 C complexes [90]. BRCA1-BRCA2-containing complex (BRCC36) is a member of the four-subunit of BRISC (BRCC36 -isopeptidase complex) that contain ABRAXAS paralog, KIAA0157/Abro1 [92]. The BRISC complex contains three additional protein components RAP80, BRCC45 and MERIT40. In addition, the BRCA1-complex contains a five-member stoichiometric complex consisting of RAP80, ABRAXAS, BRCC36, BRCC45 and MERIT40 [13]. However, the structural organization of the complex is not known.

8.1 RAP80

The Receptor Associated Protein 80 (RAP80) identified as a retinoid-related protein comprises different functionally important domains that interact with BRCA1. RAP80 is responsible for recruitment of BRCA1 to the site of DNA repair. Alteration in BRCA1 and its associated partner may affect the recruitment of BRCA1 complex to DNA damage site, and hence disturb the DNA repair process also. RAP80 acts upstream of CCDC98 and BRCA1 in DDR [6, 16, 112]. RAP80 knockdown cells have shown hypersensitivity to IR and ultraviolet (UV) light, cell- cycle dysfunction and defective homologous recombination (HR) repair [6, 7, 16, 87]. RAP80 and p53 auto-regulate each other and influence apoptosis [154]. Loss of RAP80 alleles (RAP80^{-/-}) increase the susceptibility to lymphoma, and promote tumor development in both p53^{-/-} and p53^{+/-} mice [155].

RAP80 (1-130 amino acids) wild-type and $\Delta E81$ proteins are soluble in bacterial system. Thermal and proteolytic stability of wild -type was found to be significantly higher as compared to $\Delta E81$, albeit both unfold likely with two-state transition. RAP80 UIMs are found in equilibrium between random-coil and helical states. This fact is supported by low T_m values of both wild- type and $\Delta E81$. The reason behind dynamic nature of UIMs is to provide immense flexibility of dissociation and association of ubiquitin molecules during the protein trafficking process. Perhaps UIMs also use this mechanism for multiple mode of binding (monovalent and multivalent) so as to achieve cooperativity in binding interactions. This dynamic nature is essential for a flexible and transient initiation mechanism of the DNA damage repair process. Deletion of 81E residue alters the helical state conformation, thus shifting equilibrium towards a random structure. Helical to random structure transition results in loss of several weak intermolecular hydrogen bonds and hydrophobic interactions between the UIMs and Di-Ub (K-63 linked), thereby making the binding interactions unfavorable for ubiquitin. Since binding affinity of individual UIM for mono-ubiquitin is low [161], an avidity-based mechanism probably makes the interaction between RAP80 and Lys 63-linked polyubiquitin highly robust. Co-operative binding between multiple UIMs and ubiquitin chains likely occurs, which favors the interaction of second UIM with ubiquitin after positioning of the first [159]. It has been reported [5] that expression of RAP80 $\Delta E81$ allele abates recruitment of BRCA1 complex at DSB site, which further augment chromosomal aberration (chromatic breaks). The results presented in this study also suggest that deletion of 81 Glutamic acid residue significantly obliterates RAP80 structure and impairs it's binding with polyubiquitin chain. Unstable nature of mutant and di-ubiquitin complex may be

responsible for defective recruitment of RAP80-BRCA1 complex to the DNA damage sites. Defective DNA damage repair perhaps leads to chromosomal aberration. Prolific comparison of RAP80 Δ E81 with wild-type will help in understanding its role in various diseases and repair defects. It will further explore the possibility of structure based inhibitor design for therapeutic application that can compensate the effect of such mutation.

8.2 MERIT40

MERIT40 is a key molecule in the BRCA1 complex and plays a decisive role in complex stabilization. It facilitates recruitment of BRCA1 to the site of DNA damage and favors homologous recombination repair. MERIT40 can be classified as an intrinsically disordered protein due to the presence of N- and C-terminal disorder region. However, its middle region showed well defined compact structure which indicates a possibility of crystallization of middle region for X-ray diffraction study. Structural homologous of MERIT40 suggests its plausible role in complement activation pathway. It perhaps sets up an interactions network in BRCA1 complex which is being utilized for stabilization of ABRAXAS since knockdown of MERIT40 significantly reduces the ABRAXAS and RAP80 levels [13, 19]. Stabilization of ABRAXAS probably further helps in maintaining the integrity of BRCA1-complex. This study will provide insights into the diverse interactions involved among various members which are essential for DNA repair function of BRCA1 complex. MERIT40 could be a multifunction molecule having role in DNA damage repair and complement activation.

8.3 ABRAXAS

ABRAXAS is the key member of BRCA1 complex and acts as a bridging molecule among various members. ABRAXAS expression was significantly correlated with lower chance of tumor response in patients with advanced non small-cell lung cancer receiving first-line platinum–gemcitabine chemotherapy[246]. Knockout studies of ABRAXAS showed defective recruitment of BRCA1 complex and hence the defective DNA repair[6, 7, 247]. Thus, it is a multifaceted molecule which plays a dispersive role in cancer progression, and BRCA1 mediated homologous recombination repair.

Multiple sequence alignment of ABRAXAS considering various species in phylogenic order has shown highly conserved nature of Arg361 residue. The modeled structure of ABRAXAS wild- type and R361Q mutant showed structural alteration. However, the observed structural changes were not contributing in oligomeric properties. Wild- type and mutant have shown similar secondary structural composition while the relative orientation of Trp and Tyr was slightly disturbed. This indicates that R361Q mutation bringing several localized changes in structure pattern of ABRAXAS which altogether furnish a different conformational stability of structure in a cumulative manner. These conformational changes are very minor and hence could not be detected at secondary structural level while their relative positions were traced during three dimension unfolding pathway. The relative redundancy of intermediate species in case of wild -type suggest the existence of different unfolding pathway which is partially followed by mutant. Altogether, the localized changes in the mutant structure brings down its thermal and chemical stability which further perturb interaction with RAP80. The cumulative global changes in mutant structure was sufficient to disturb critical interaction

necessary for BRCA1 complex integrity and localization. Therefore, in the presence of R361Q mutation, ABRAXAS could not extend its bridging interaction through RAP80 which perhaps reduce the recruitment of BRCA1 complex to the DNA damage site. Consequently, the nuclear retention of BRCA1 is adversely affected which further agitates G2/M checkpoint and homology-directed DNA repair [11]. These finding would substantially list ABRAXAS as a new susceptibility gene to cancer predisposition. It also opens the vast perspective of considering R361Q mutation role in disease progression such as cancer. It will further explore the opportunity of inhibitor design for therapeutic application that can recompense the effect of such adverse mutation.

8.4 BRCA1-complex

Protein-Protein Interactions in BRCA1 complex plays very significant role and alterations in interaction profile leads to stabilization/destabilization of whole complex. MERIT40 is an essential component of BRCA1 complex. MERIT40 showed direct interaction with ABRAXAS and BRCA1-BRCT as indicated with significant heat change during reaction. The binding association between MERIT40 and ABRAXAS was found $K_d = 2.56 \pm 0.24 \mu\text{M}$. Interactions between BRCA1-BRCT domain and MERIT40 with a affinity constant $K = 521 \pm 7 \mu\text{M}$ was established. MERIT40 showed a direct interaction with BRCA1-BRCT. In order to validate this finding, pull-down assay was performed, which confirm the interaction between MERIT40 and ABRAXAS. It is reported that ABRAXAS form phosphodependent interaction with BRCA1 BRCT and phospho-independent interaction with RAP80, thus acting as a bridging molecule in BRCA1 complex. Binding interaction between MERIT40 and ABRAXAS probably help

in extension of bridging interaction among various members of BRCA1 complex and thereby maintaining its integrity.

8.5 Future perspective

The structural and functional analysis of BRCA1 complex was studied to understand its role in DNA damage repair. Two important mutations and their disease correlation have been analyzed in addition to dissecting the stabilization mechanism of BRCA1 complex through the MERIT40. The next approach could be the expression of MERIT40 stable domain (72-300 amino acids) in higher system such as yeast or insect cells lines in order to get a well folded protein with all desired post translation modification. The main issue in bacterial system was oligomerization which highly reduce the propensity of MERIT40 crystallization. However, a rational approach of cloning into higher expression system can result a well folded homogenous protein which would most likely form the crystal and good diffraction pattern. The next approach could be co-crystallization of BRCA1 members such as RAP80-ABRAXAS and MERIT40-BRCA1 BRCT. In this regard, the major problem was the purification of functional and homogenous ABRAXAS protein. The issue can be significantly resolve by its co-expression with other members either in bacterial or other expression system. In addition, role of MERIT40 can be explored in complement activation pathway due to its structural resemblance with complement factor B.

References

References:

1. Alberg, A.J., A.P. Lam, and K.J. Helzlsouer, *Epidemiology, prevention, and early detection of breast cancer*. Curr Opin Oncol, 1999. **11**(6): p. 435-41.
2. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1*. Science, 1994. **266**(5182): p. 66-71.
3. Venkitaraman, A.R., *Cancer susceptibility and the functions of BRCA1 and BRCA2*. Cell, 2002. **108**(2): p. 171-82.
4. Pylkas, K., et al., *Analysis of large deletions in BRCA1, BRCA2 and PALB2 genes in Finnish breast and ovarian cancer families*. BMC Cancer, 2008. **8**: p. 146.
5. Nikkila, J., et al., *Familial breast cancer screening reveals an alteration in the RAP80 UIM domain that impairs DNA damage response function*. Oncogene, 2009. **28**(16): p. 1843-52.
6. Wang, B., et al., *Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response*. Science, 2007. **316**(5828): p. 1194-8.
7. Kim, H., J. Huang, and J. Chen, *CCDC98 is a BRCA1-BRCT domain-binding protein involved in the DNA damage response*. Nat Struct Mol Biol, 2007. **14**(8): p. 710-5.
8. Monteiro, A.N., A. August, and H. Hanafusa, *Evidence for a transcriptional activation function of BRCA1 C-terminal region*. Proc Natl Acad Sci U S A, 1996. **93**(24): p. 13595-9.
9. Yu, X., et al., *The BRCT domain is a phospho-protein binding domain*. Science, 2003. **302**(5645): p. 639-42.
10. Manke, I.A., et al., *BRCT repeats as phosphopeptide-binding modules involved in protein targeting*. Science, 2003. **302**(5645): p. 636-9.
11. Solyom, S., et al., *Breast cancer-associated Abraxas mutation disrupts nuclear localization and DNA damage response functions*. Sci Transl Med, 2012. **4**(122): p. 122ra23.
12. Arnold, K., et al., *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling*. Bioinformatics, 2006. **22**(2): p. 195-201.
13. Shao, G., et al., *MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks*. Genes Dev, 2009. **23**(6): p. 740-54.
14. Holm, L. and P. Rosenstrom, *Dali server: conservation mapping in 3D*. Nucleic Acids Res, 2010. **38**(Web Server issue): p. W545-9.
15. Celeste, A., et al., *Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks*. Nat Cell Biol, 2003. **5**(7): p. 675-9.
16. Sobhian, B., et al., *RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites*. Science, 2007. **316**(5828): p. 1198-202.
17. Mailand, N., et al., *RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins*. Cell, 2007. **131**(5): p. 887-900.
18. Wang, Y., et al., *BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures*. Genes Dev, 2000. **14**(8): p. 927-39.
19. Feng, L., J. Huang, and J. Chen, *MERIT40 facilitates BRCA1 localization and DNA damage repair*. Genes Dev, 2009. **23**(6): p. 719-28.
20. Wang, B., et al., *NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control*. Genes Dev, 2009. **23**(6): p. 729-39.
21. Rodriguez, M., et al., *Phosphopeptide binding specificities of BRCA1 COOH-terminal (BRCT) domains*. J Biol Chem, 2003. **278**(52): p. 52914-8.

22. Bloom, H.J. and W.W. Richardson, *Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years*. Br J Cancer, 1957. **11**(3): p. 359-77.
23. Levy-Lahad, E. and E. Friedman, *Cancer risks among BRCA1 and BRCA2 mutation carriers*. Br J Cancer, 2007. **96**(1): p. 11-5.
24. Hall, J.M., et al., *Linkage of early-onset familial breast cancer to chromosome 17q21*. Science, 1990. **250**(4988): p. 1684-9.
25. Narod, S.A., et al., *Familial breast-ovarian cancer locus on chromosome 17q12-q23*. Lancet, 1991. **338**(8759): p. 82-3.
26. Wooster, R., et al., *Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13*. Science, 1994. **265**(5181): p. 2088-90.
27. Petrucelli, N., M.B. Daly, and G.L. Feldman, *BRCA1 and BRCA2 Hereditary Breast and Ovarian Cancer*. 1993.
28. Kelsell, D.P., et al., *Genetic analysis of the BRCA1 region in a large breast/ovarian family: refinement of the minimal region containing BRCA1*. Hum Mol Genet, 1993. **2**(11): p. 1823-8.
29. Neuhausen, S.L. and C.J. Marshall, *Loss of heterozygosity in familial tumors from three BRCA1-linked kindreds*. Cancer Res, 1994. **54**(23): p. 6069-72.
30. De Bont, R. and N. van Larebeke, *Endogenous DNA damage in humans: a review of quantitative data*. Mutagenesis, 2004. **19**(3): p. 169-85.
31. Rodriguez-Vargas, J.M., et al., *ROS-induced DNA damage and PARP-1 are required for optimal induction of starvation-induced autophagy*. Cell Res, 2012. **22**(7): p. 1181-98.
32. Huen, M.S., S.M. Sy, and J. Chen, *BRCA1 and its toolbox for the maintenance of genome integrity*. Nat Rev Mol Cell Biol, 2010. **11**(2): p. 138-48.
33. Zhong, Q., et al., *Deficient nonhomologous end-joining activity in cell-free extracts from Brca1-null fibroblasts*. Cancer Res, 2002. **62**(14): p. 3966-70.
34. Zhong, Q., et al., *BRCA1 facilitates microhomology-mediated end joining of DNA double strand breaks*. J Biol Chem, 2002. **277**(32): p. 28641-7.
35. Harper, J.W. and S.J. Elledge, *The DNA damage response: ten years after*. Mol Cell, 2007. **28**(5): p. 739-45.
36. Cimprich, K.A. and D. Cortez, *ATR: an essential regulator of genome integrity*. Nat Rev Mol Cell Biol, 2008. **9**(8): p. 616-27.
37. Jazayeri, A., et al., *ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks*. Nat Cell Biol, 2006. **8**(1): p. 37-45.
38. Myers, J.S. and D. Cortez, *Rapid activation of ATR by ionizing radiation requires ATM and Mre11*. J Biol Chem, 2006. **281**(14): p. 9346-50.
39. Lee, J.H. and T.T. Paull, *ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex*. Science, 2005. **308**(5721): p. 551-4.
40. Berkovich, E., R.J. Monnat, Jr., and M.B. Kastan, *Roles of ATM and NBS1 in chromatin structure modulation and DNA double-strand break repair*. Nat Cell Biol, 2007. **9**(6): p. 683-90.
41. Falck, J., J. Coates, and S.P. Jackson, *Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage*. Nature, 2005. **434**(7033): p. 605-11.
42. Kitagawa, R., et al., *Phosphorylation of SMC1 is a critical downstream event in the ATM-NBS1-BRCA1 pathway*. Genes Dev, 2004. **18**(12): p. 1423-38.
43. Buis, J., et al., *Mre11 nuclease activity has essential roles in DNA repair and genomic stability distinct from ATM activation*. Cell, 2008. **135**(1): p. 85-96.

44. Schaetzlein, S., et al., *Exonuclease-1 deletion impairs DNA damage signaling and prolongs lifespan of telomere-dysfunctional mice*. *Cell*, 2007. **130**(5): p. 863-77.
45. Bartek, J. and J. Lukas, *Chk1 and Chk2 kinases in checkpoint control and cancer*. *Cancer Cell*, 2003. **3**(5): p. 421-9.
46. Osborn, A.J. and S.J. Elledge, *Mrc1 is a replication fork component whose phosphorylation in response to DNA replication stress activates Rad53*. *Genes Dev*, 2003. **17**(14): p. 1755-67.
47. Chini, C.C., J. Wood, and J. Chen, *Chk1 is required to maintain claspin stability*. *Oncogene*, 2006. **25**(30): p. 4165-71.
48. Yang, X.H., et al., *Chk1 and Claspin potentiate PCNA ubiquitination*. *Genes Dev*, 2008. **22**(9): p. 1147-52.
49. Kumagai, A. and W.G. Dunphy, *Repeated phosphopeptide motifs in Claspin mediate the regulated binding of Chk1*. *Nat Cell Biol*, 2003. **5**(2): p. 161-5.
50. Kumagai, A., S.M. Kim, and W.G. Dunphy, *Claspin and the activated form of ATR-ATRIP collaborate in the activation of Chk1*. *J Biol Chem*, 2004. **279**(48): p. 49599-608.
51. Mailand, N., et al., *Destruction of Claspin by SCFbetaTrCP restrains Chk1 activation and facilitates recovery from genotoxic stress*. *Mol Cell*, 2006. **23**(3): p. 307-18.
52. Peschiaroli, A., et al., *SCFbetaTrCP-mediated degradation of Claspin regulates recovery from the DNA replication checkpoint response*. *Mol Cell*, 2006. **23**(3): p. 319-29.
53. Bassermann, F., et al., *The Cdc14B-Cdh1-Plk1 axis controls the G2 DNA-damage-response checkpoint*. *Cell*, 2008. **134**(2): p. 256-67.
54. Hu, Y.F. and R. Li, *JunB potentiates function of BRCA1 activation domain 1 (AD1) through a coiled-coil-mediated interaction*. *Genes Dev*, 2002. **16**(12): p. 1509-17.
55. Varma, A.K., et al., *Structural basis for cell cycle checkpoint control by the BRCA1-CtIP complex*. *Biochemistry*, 2005. **44**(33): p. 10941-6.
56. Stucki, M., et al., *MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks*. *Cell*, 2005. **123**(7): p. 1213-26.
57. Krum, S.A., et al., *BRCA1 associates with processive RNA polymerase II*. *J Biol Chem*, 2003. **278**(52): p. 52012-20.
58. Xu, B., S. Kim, and M.B. Kastan, *Involvement of Brca1 in S-phase and G(2)-phase checkpoints after ionizing irradiation*. *Mol Cell Biol*, 2001. **21**(10): p. 3445-50.
59. Zhang, J., et al., *Chk2 phosphorylation of BRCA1 regulates DNA double-strand break repair*. *Mol Cell Biol*, 2004. **24**(2): p. 708-18.
60. Kim, O.H., et al., *Influence of p53 and p21Waf1 expression on G2/M phase arrest of colorectal carcinoma HCT116 cells to proteasome inhibitors*. *Int J Oncol*, 2004. **24**(4): p. 935-41.
61. O'Driscoll, M. and P.A. Jeggo, *The role of double-strand break repair - insights from human genetics*. *Nat Rev Genet*, 2006. **7**(1): p. 45-54.
62. Lane, T.F., *BRCA1 and transcription*. *Cancer Biol Ther*, 2004. **3**(6): p. 528-33.
63. Chai, Y.L., et al., *The second BRCT domain of BRCA1 proteins interacts with p53 and stimulates transcription from the p21WAF1/CIP1 promoter*. *Oncogene*, 1999. **18**(1): p. 263-8.
64. Ouchi, T., et al., *BRCA1 regulates p53-dependent gene expression*. *Proc Natl Acad Sci U S A*, 1998. **95**(5): p. 2302-6.
65. MacLachlan, T.K., et al., *Repression of BRCA1 through a feedback loop involving p53*. *J Biol Chem*, 2000. **275**(41): p. 31869-75.

66. Chen, C.F., et al., *The nuclear localization sequences of the BRCA1 protein interact with the importin-alpha subunit of the nuclear transport signal receptor*. J Biol Chem, 1996. **271**(51): p. 32863-8.
67. Thakur, S., et al., *Localization of BRCA1 and a splice variant identifies the nuclear localization signal*. Mol Cell Biol, 1997. **17**(1): p. 444-52.
68. Rodriguez, J.A., et al., *Nuclear-cytoplasmic shuttling of BARD1 contributes to its proapoptotic activity and is regulated by dimerization with BRCA1*. Oncogene, 2004. **23**(10): p. 1809-20.
69. Gu, J., et al., *Identification of p53 sequence elements that are required for MDM2-mediated nuclear export*. Mol Cell Biol, 2001. **21**(24): p. 8533-46.
70. Fabbro, M., et al., *BARD1 induces BRCA1 intranuclear foci formation by increasing RING-dependent BRCA1 nuclear import and inhibiting BRCA1 nuclear export*. J Biol Chem, 2002. **277**(24): p. 21315-24.
71. Morris, J.R. and E. Solomon, *BRCA1 : BARD1 induces the formation of conjugated ubiquitin structures, dependent on K6 of ubiquitin, in cells during DNA replication and repair*. Hum Mol Genet, 2004. **13**(8): p. 807-17.
72. Ye, Q., et al., *BRCA1-induced large-scale chromatin unfolding and allele-specific effects of cancer-predisposing mutations*. J Cell Biol, 2001. **155**(6): p. 911-21.
73. Aiyar, S.E., et al., *Attenuation of estrogen receptor alpha-mediated transcription through estrogen-stimulated recruitment of a negative elongation factor*. Genes Dev, 2004. **18**(17): p. 2134-46.
74. Yun, J. and W.H. Lee, *Degradation of transcription repressor ZBRK1 through the ubiquitin-proteasome pathway relieves repression of Gadd45a upon DNA damage*. Mol Cell Biol, 2003. **23**(20): p. 7305-14.
75. Tan, W., S. Kim, and T.G. Boyer, *Tetrameric oligomerization mediates transcriptional repression by the BRCA1-dependent Kruppel-associated box-zinc finger protein ZBRK1*. J Biol Chem, 2004. **279**(53): p. 55153-60.
76. Tan, W., et al., *Functional dissection of transcription factor ZBRK1 reveals zinc fingers with dual roles in DNA-binding and BRCA1-dependent transcriptional repression*. J Biol Chem, 2004. **279**(8): p. 6576-87.
77. Zheng, L., et al., *Sequence-specific transcriptional corepressor function for BRCA1 through a novel zinc finger protein, ZBRK1*. Mol Cell, 2000. **6**(4): p. 757-68.
78. Wang, Q., et al., *BRCA1 binds c-Myc and inhibits its transcriptional and transforming activity in cells*. Oncogene, 1998. **17**(15): p. 1939-48.
79. Xiong, J., et al., *BRCA1 inhibition of telomerase activity in cultured cells*. Mol Cell Biol, 2003. **23**(23): p. 8668-90.
80. Yarden, R.I., et al., *BRCA1 regulates the G2/M checkpoint by activating Chk1 kinase upon DNA damage*. Nat Genet, 2002. **30**(3): p. 285-9.
81. Aprelikova, O.N., et al., *BRCA1-associated growth arrest is RB-dependent*. Proc Natl Acad Sci U S A, 1999. **96**(21): p. 11866-71.
82. Baer, R. and T. Ludwig, *The BRCA1/BARD1 heterodimer, a tumor suppressor complex with ubiquitin E3 ligase activity*. Curr Opin Genet Dev, 2002. **12**(1): p. 86-91.
83. Sankaran, S., et al., *Centrosomal microtubule nucleation activity is inhibited by BRCA1-dependent ubiquitination*. Mol Cell Biol, 2005. **25**(19): p. 8656-68.
84. Wang, R.H., H. Yu, and C.X. Deng, *A requirement for breast-cancer-associated gene 1 (BRCA1) in the spindle checkpoint*. Proc Natl Acad Sci U S A, 2004. **101**(49): p. 17108-13.
85. Rogakou, E.P., et al., *Megabase chromatin domains involved in DNA double-strand breaks in vivo*. J Cell Biol, 1999. **146**(5): p. 905-16.

86. Rogakou, E.P., et al., *DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139*. J Biol Chem, 1998. **273**(10): p. 5858-68.
87. Wang, B. and S.J. Elledge, *Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brca1/Brcc36 complex in response to DNA damage*. Proc Natl Acad Sci U S A, 2007. **104**(52): p. 20759-63.
88. Lou, Z., et al., *MDC1 maintains genomic stability by participating in the amplification of ATM-dependent DNA damage signals*. Mol Cell, 2006. **21**(2): p. 187-200.
89. Venkitaraman, A.R., *Breast cancer genes and DNA repair*. Science, 1999. **286**(5442): p. 1100-2.
90. Yu, X. and J. Chen, *DNA damage-induced cell cycle checkpoint control requires CtIP, a phosphorylation-dependent binding partner of BRCA1 C-terminal domains*. Mol Cell Biol, 2004. **24**(21): p. 9478-86.
91. Shiozaki, E.N., et al., *Structure of the BRCT repeats of BRCA1 bound to a BACH1 phosphopeptide: implications for signaling*. Mol Cell, 2004. **14**(3): p. 405-12.
92. Cooper, E.M., et al., *K63-specific deubiquitination by two JAMM/MPN+ complexes: BRISC-associated Brcc36 and proteasomal Poh1*. EMBO J, 2009. **28**(6): p. 621-31.
93. Rodriguez, J.A., W.W. Au, and B.R. Henderson, *Cytoplasmic mislocalization of BRCA1 caused by cancer-associated mutations in the BRCT domain*. Exp Cell Res, 2004. **293**(1): p. 14-21.
94. Williams, R.S., et al., *Detection of protein folding defects caused by BRCA1-BRCT truncation and missense mutations*. J Biol Chem, 2003. **278**(52): p. 53007-16.
95. Irminger-Finger, I. and C.E. Jefford, *Is there more to BARD1 than BRCA1?* Nat Rev Cancer, 2006. **6**(5): p. 382-91.
96. Narod, S.A. and W.D. Foulkes, *BRCA1 and BRCA2: 1994 and beyond*. Nat Rev Cancer, 2004. **4**(9): p. 665-76.
97. McCarthy, E.E., et al., *Loss of Bard1, the heterodimeric partner of the Brca1 tumor suppressor, results in early embryonic lethality and chromosomal instability*. Mol Cell Biol, 2003. **23**(14): p. 5056-63.
98. Thai, T.H., et al., *Mutations in the BRCA1-associated RING domain (BARD1) gene in primary breast, ovarian and uterine cancers*. Hum Mol Genet, 1998. **7**(2): p. 195-202.
99. Lorick, K.L., et al., *RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination*. Proc Natl Acad Sci U S A, 1999. **96**(20): p. 11364-9.
100. Wu, L.C., et al., *Identification of a RING protein that can interact in vivo with the BRCA1 gene product*. Nat Genet, 1996. **14**(4): p. 430-40.
101. Brzovic, P.S., et al., *Structure of a BRCA1-BARD1 heterodimeric RING-RING complex*. Nat Struct Biol, 2001. **8**(10): p. 833-7.
102. Kleiman, F.E., et al., *BRCA1/BARD1 inhibition of mRNA 3' processing involves targeted degradation of RNA polymerase II*. Genes Dev, 2005. **19**(10): p. 1227-37.
103. Irminger-Finger, I., et al., *Identification of BARD1 as mediator between proapoptotic stress and p53-dependent apoptosis*. Mol Cell, 2001. **8**(6): p. 1255-66.
104. Feki, A., et al., *BARD1 expression during spermatogenesis is associated with apoptosis and hormonally regulated*. Biol Reprod, 2004. **71**(5): p. 1614-24.
105. Irminger-Finger, I., et al., *In vitro repression of Brca1-associated RING domain gene, Bard1, induces phenotypic changes in mammary epithelial cells*. J Cell Biol, 1998. **143**(5): p. 1329-39.
106. Fisch, P., et al., *The chromosomal translocation t(X;14)(q28;q11) in T-cell pro-lymphocytic leukaemia breaks within one gene and activates another*. Oncogene, 1993. **8**(12): p. 3271-6.

107. Dong, Y., et al., *Regulation of BRCC, a holoenzyme complex containing BRCA1 and BRCA2, by a signalosome-like subunit and its role in DNA repair*. Mol Cell, 2003. **12**(5): p. 1087-99.
108. Yan, Z., Y.S. Kim, and A.M. Jetten, *RAP80, a novel nuclear protein that interacts with the retinoid-related testis-associated receptor*. J Biol Chem, 2002. **277**(35): p. 32379-88.
109. Lei, W., et al., *Cloning of the human orphan receptor germ cell nuclear factor/retinoid receptor-related testis-associated receptor and its differential regulation during embryonal carcinoma cell differentiation*. J Mol Endocrinol, 1997. **18**(2): p. 167-76.
110. Yan, J., et al., *Ubiquitin-interaction motifs of RAP80 are critical in its regulation of estrogen receptor alpha*. Nucleic Acids Res, 2007. **35**(5): p. 1673-86.
111. Yan, J. and A.M. Jetten, *RAP80 and RNF8, key players in the recruitment of repair proteins to DNA damage sites*. Cancer Lett, 2008. **271**(2): p. 179-90.
112. Kim, H., J. Chen, and X. Yu, *Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response*. Science, 2007. **316**(5828): p. 1202-5.
113. Wu, J., et al., *RAP80 protein is important for genomic stability and is required for stabilizing BRCA1-A complex at DNA damage sites in vivo*. J Biol Chem, 2012. **287**(27): p. 22919-26.
114. Sato, Y., et al., *Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by tandem UIMs of RAP80*. EMBO J, 2009. **28**(16): p. 2461-8.
115. Yu, X., et al., *The C-terminal (BRCT) domains of BRCA1 interact in vivo with CtIP, a protein implicated in the CtBP pathway of transcriptional repression*. J Biol Chem, 1998. **273**(39): p. 25388-92.
116. Scully, R., et al., *Genetic analysis of BRCA1 function in a defined tumor cell line*. Mol Cell, 1999. **4**(6): p. 1093-9.
117. Pickart, C.M., *Mechanisms underlying ubiquitination*. Annu Rev Biochem, 2001. **70**: p. 503-33.
118. Weissman, A.M., *Themes and variations on ubiquitylation*. Nat Rev Mol Cell Biol, 2001. **2**(3): p. 169-78.
119. Chau, V., et al., *A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein*. Science, 1989. **243**(4898): p. 1576-83.
120. Gregori, L., et al., *A uniform isopeptide-linked multiubiquitin chain is sufficient to target substrate for degradation in ubiquitin-mediated proteolysis*. J Biol Chem, 1990. **265**(15): p. 8354-7.
121. Pickart, C.M., *Ubiquitin in chains*. Trends Biochem Sci, 2000. **25**(11): p. 544-8.
122. Pickart, C.M. and D. Fushman, *Polyubiquitin chains: polymeric protein signals*. Curr Opin Chem Biol, 2004. **8**(6): p. 610-6.
123. Kerscher, O., R. Felberbaum, and M. Hochstrasser, *Modification of proteins by ubiquitin and ubiquitin-like proteins*. Annu Rev Cell Dev Biol, 2006. **22**: p. 159-80.
124. Hoege, C., et al., *RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO*. Nature, 2002. **419**(6903): p. 135-41.
125. Zheng, N., et al., *Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases*. Cell, 2000. **102**(4): p. 533-9.
126. Choudhury, A.D., H. Xu, and R. Baer, *Ubiquitination and proteasomal degradation of the BRCA1 tumor suppressor is regulated during cell cycle progression*. J Biol Chem, 2004. **279**(32): p. 33909-18.
127. Nishikawa, H., et al., *Mass spectrometric and mutational analyses reveal Lys-6-linked polyubiquitin chains catalyzed by BRCA1-BARD1 ubiquitin ligase*. J Biol Chem, 2004. **279**(6): p. 3916-24.

128. Dikic, I., S. Wakatsuki, and K.J. Walters, *Ubiquitin-binding domains - from structures to functions*. Nat Rev Mol Cell Biol, 2009. **10**(10): p. 659-71.
129. Verma, R., et al., *Multiubiquitin chain receptors define a layer of substrate selectivity in the ubiquitin-proteasome system*. Cell, 2004. **118**(1): p. 99-110.
130. Zhu, H. and R.A. Dean, *A novel method for increasing the transformation efficiency of Escherichia coli-application for bacterial artificial chromosome library construction*. Nucleic Acids Res, 1999. **27**(3): p. 910-1.
131. Pace, C.N. and K.L. Shaw, *Linear extrapolation method of analyzing solvent denaturation curves*. Proteins, 2000. **Suppl 4**: p. 1-7.
132. Pace, C.N., *Determination and analysis of urea and guanidine hydrochloride denaturation curves*. Methods Enzymol, 1986. **131**: p. 266-80.
133. Yadav, S. and F. Ahmad, *A new method for the determination of stability parameters of proteins from their heat-induced denaturation curves*. Anal Biochem, 2000. **283**(2): p. 207-13.
134. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data*. Electrophoresis, 1999. **20**(18): p. 3551-67.
135. Sael, L., M. Chitale, and D. Kihara, *Structure- and sequence-based function prediction for non-homologous proteins*. J Struct Funct Genomics, 2012. **13**(2): p. 111-23.
136. Kim, D.E., D. Chivian, and D. Baker, *Protein structure prediction and analysis using the Robetta server*. Nucleic Acids Res, 2004. **32**(Web Server issue): p. W526-31.
137. Schneidman-Duhovny, D., et al., *PatchDock and SymmDock: servers for rigid and symmetric docking*. Nucleic Acids Res, 2005. **33**(Web Server issue): p. W363-7.
138. Zerovnik, E., *Oligomerization preceding amyloid fibril formation: a process in common to intrinsically disordered and globular proteins*. Network, 2011. **22**(1-4): p. 154-61.
139. Wang, X., et al., *Structural NMR of protein oligomers using hybrid methods*. J Struct Biol, 2011. **173**(3): p. 515-29.
140. Cizas, P., et al., *Size-dependent neurotoxicity of beta-amyloid oligomers*. Arch Biochem Biophys, 2010. **496**(2): p. 84-92.
141. Melcher, K., *New chemical crosslinking methods for the identification of transient protein-protein interactions with multiprotein complexes*. Curr Protein Pept Sci, 2004. **5**(4): p. 287-96.
142. Feeney, R.E., *Chemical modification of proteins: comments and perspectives*. Int J Pept Protein Res, 1987. **29**(2): p. 145-61.
143. Molgaard, A. and S. Larsen, *Crystal packing in two pH-dependent crystal forms of rhamnogalacturonan acetyltransferase*. Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 3): p. 472-8.
144. D'arcy, A., et al., *A novel approach to crystallising proteins under oil*. Journal of crystal growth, 1996. **168**(1): p. 175-180.
145. Chayen, N.E., et al., *An automated system for micro-batch protein crystallization and screening*. Journal of applied crystallography, 1990. **23**(4): p. 297-302.
146. Zeppezauer, M., H. Eklund, and E.S. Zeppezauer, *Micro diffusion cells for the growth of single protein crystals by means of equilibrium dialysis*. Archives of biochemistry and biophysics, 1968. **126**(2): p. 564-573.
147. Paull, T.T., et al., *A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage*. Curr Biol, 2000. **10**(15): p. 886-95.
148. Petrini, J.H. and T.H. Stracker, *The cellular response to DNA double-strand breaks: defining the sensors and mediators*. Trends Cell Biol, 2003. **13**(9): p. 458-62.

149. Huen, M.S., et al., *RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly*. Cell, 2007. **131**(5): p. 901-14.
150. Kolas, N.K., et al., *Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase*. Science, 2007. **318**(5856): p. 1637-40.
151. Bassing, C.H. and F.W. Alt, *H2AX may function as an anchor to hold broken chromosomal DNA ends in close proximity*. Cell Cycle, 2004. **3**(2): p. 149-53.
152. Bassing, C.H., et al., *Increased ionizing radiation sensitivity and genomic instability in the absence of histone H2AX*. Proc Natl Acad Sci U S A, 2002. **99**(12): p. 8173-8.
153. Stucki, M. and S.P. Jackson, *gammaH2AX and MDC1: anchoring the DNA-damage-response machinery to broken chromosomes*. DNA Repair (Amst), 2006. **5**(5): p. 534-43.
154. Yan, J., et al., *A regulatory loop composed of RAP80-HDM2-p53 provides RAP80-enhanced p53 degradation by HDM2 in response to DNA damage*. J Biol Chem, 2009. **284**(29): p. 19280-9.
155. Yin, Z., et al., *RAP80 Is Critical in Maintaining Genomic Stability and Suppressing Tumor Development*. Cancer Res, 2012. **72**(19): p. 5080-90.
156. Vikrant, et al., *Structural and Functional Implication of RAP80 DeltaGlu81 Mutation*. PLoS One, 2013. **8**(9): p. e72707.
157. Laskowski, R.A., et al., *PDBsum: a Web-based database of summaries and analyses of all PDB structures*. Trends Biochem Sci, 1997. **22**(12): p. 488-90.
158. Markin, C.J., W. Xiao, and L. Spyropoulos, *Mechanism for recognition of polyubiquitin chains: balancing affinity through interplay between multivalent binding and dynamics*. J Am Chem Soc, 2010. **132**(32): p. 11247-58.
159. Sato, Y., et al., *Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by NZF domains of TAB2 and TAB3*. EMBO J, 2009. **28**(24): p. 3903-9.
160. Sekiyama, N., et al., *NMR analysis of Lys63-linked polyubiquitin recognition by the tandem ubiquitin-interacting motifs of Rap80*. J Biomol NMR, 2012. **52**(4): p. 339-50.
161. Hurley, J.H., S. Lee, and G. Prag, *Ubiquitin-binding domains*. Biochem J, 2006. **399**(3): p. 361-72.
162. Haile, D.T. and J.D. Parvin, *Activation of transcription in vitro by the BRCA1 carboxyl-terminal domain*. J Biol Chem, 1999. **274**(4): p. 2113-7.
163. Xu, X., et al., *Centrosome amplification and a defective G2-M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform-deficient cells*. Mol Cell, 1999. **3**(3): p. 389-95.
164. Foulkes, W.D., *BRCA1 functions as a breast stem cell regulator*. J Med Genet, 2004. **41**(1): p. 1-5.
165. Scully, R., et al., *Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage*. Cell, 1997. **90**(3): p. 425-35.
166. Hegyi, H. and M. Gerstein, *The relationship between protein structure and function: a comprehensive survey with application to the yeast genome*. J Mol Biol, 1999. **288**(1): p. 147-64.
167. Wrzeszczynski, K.O. and B. Rost, *Cell cycle kinases predicted from conserved biophysical properties*. Proteins, 2009. **74**(3): p. 655-68.
168. Vikrant, et al., *Structural and functional characterization of the MERIT40 to understand its role in DNA repair*. Journal of Biomolecular Structure and Dynamics, 2013(ahead-of-print): p. 1-16.
169. Morris, A.L., et al., *Stereochemical quality of protein structure coordinates*. Proteins, 1992. **12**(4): p. 345-64.

170. Ramachandran, G.N., C. Ramakrishnan, and V. Sasisekharan, *Stereochemistry of polypeptide chain configurations*. J Mol Biol, 1963. **7**: p. 95-9.
171. Ramachandran, G.N. and V. Sasisekharan, *Conformation of polypeptides and proteins*. Adv Protein Chem, 1968. **23**: p. 283-438.
172. Metfessel, B.A., et al., *Cross-validation of protein structural class prediction using statistical clustering and neural networks*. Protein Sci, 1993. **2**(7): p. 1171-82.
173. Bagaria, A., et al., *Protein structure validation by generalized linear model root-mean-square deviation prediction*. Protein Sci, 2012. **21**(2): p. 229-38.
174. Rajagopalan, K., et al., *A majority of the cancer/testis antigens are intrinsically disordered proteins*. J Cell Biochem, 2011. **112**(11): p. 3256-67.
175. Uversky, V.N. and A.K. Dunker, *Understanding protein non-folding*. Biochim Biophys Acta, 2010. **1804**(6): p. 1231-64.
176. Uversky, V.N., C.J. Oldfield, and A.K. Dunker, *Intrinsically disordered proteins in human diseases: introducing the D2 concept*. Annu Rev Biophys, 2008. **37**: p. 215-46.
177. Fontana, A., et al., *Identifying disordered regions in proteins by limited proteolysis*. Methods Mol Biol, 2012. **896**: p. 297-318.
178. Fontana, A., et al., *Probing the conformational state of apomyoglobin by limited proteolysis*. J Mol Biol, 1997. **266**(2): p. 223-30.
179. Hofmann, K. and P. Bucher, *The PCI domain: a common theme in three multiprotein complexes*. Trends Biochem Sci, 1998. **23**(6): p. 204-5.
180. Vogel, C.W. and D.C. Fritzing, *Cobra venom factor: Structure, function, and humanization for therapeutic complement depletion*. Toxicon, 2010. **56**(7): p. 1198-222.
181. Vogel, C.W., C.A. Smith, and H.J. Muller-Eberhard, *Cobra venom factor: structural homology with the third component of human complement*. J Immunol, 1984. **133**(6): p. 3235-41.
182. Eggertsen, G., P. Lind, and J. Sjoquist, *Molecular characterization of the complement activating protein in the venom of the Indian cobra (Naja N. siamensis)*. Mol Immunol, 1981. **18**(2): p. 125-33.
183. Fritzing, D.C., R. Bredehorst, and C.W. Vogel, *Molecular cloning and derived primary structure of cobra venom factor*. Proc Natl Acad Sci U S A, 1994. **91**(26): p. 12775-9.
184. Wootton, S.K. and D. Yoo, *Homo-oligomerization of the porcine reproductive and respiratory syndrome virus nucleocapsid protein and the role of disulfide linkages*. J Virol, 2003. **77**(8): p. 4546-57.
185. Zhang, C., et al., *Determination of atomic desolvation energies from the structures of crystallized proteins*. J Mol Biol, 1997. **267**(3): p. 707-26.
186. Lobley, A., L. Whitmore, and B.A. Wallace, *DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra*. Bioinformatics, 2002. **18**(1): p. 211-2.
187. Muniz, V.A., et al., *The role of the local environment of engineered Tyr to Trp substitutions for probing the denaturation mechanism of FIS*. Protein Sci, 2011. **20**(2): p. 302-12.
188. Pace, C.N., et al., *Forces contributing to the conformational stability of proteins*. FASEB J, 1996. **10**(1): p. 75-83.
189. Walters, J., S.L. Milam, and A.C. Clark, *Practical approaches to protein folding and assembly: spectroscopic strategies in thermodynamics and kinetics*. Methods Enzymol, 2009. **455**: p. 1-39.

190. Solyom, S., et al., *Mutation screening of the MERIT40 gene encoding a novel BRCA1 and RAP80 interacting protein in breast cancer families*. Breast Cancer Res Treat, 2010. **120**(1): p. 165-8.
191. Williams, R.S. and J.N. Glover, *Structural consequences of a cancer-causing BRCA1-BRCT missense mutation*. J Biol Chem, 2003. **278**(4): p. 2630-5.
192. Liu, Z., J. Wu, and X. Yu, *CCDC98 targets BRCA1 to DNA damage sites*. Nat Struct Mol Biol, 2007. **14**(8): p. 716-20.
193. Osorio, A., et al., *Evaluation of the BRCA1 interacting genes RAP80 and CCDC98 in familial breast cancer susceptibility*. Breast Cancer Res Treat, 2009. **113**(2): p. 371-6.
194. Vikrant, et al., *Mislocalization of BRCA1-complex due to ABRAXAS Arg361Gln mutation*. J Biomol Struct Dyn, 2014: p. 1-11.
195. Zhukov, I., L. Jaroszewski, and A. Bierzynski, *Conservative mutation Met8 --> Leu affects the folding process and structural stability of squash trypsin inhibitor CMTI-I*. Protein Sci, 2000. **9**(2): p. 273-9.
196. Kolchanov, N.A., V.V. Soloviov, and A.A. Zharkikh, *The effects of mutations, deletions and insertions of single amino acids on the three-dimensional structure of globins*. FEBS Lett, 1983. **161**(1): p. 65-70.
197. Havlis, J., et al., *Fast-response proteomics by accelerated in-gel digestion of proteins*. Anal Chem, 2003. **75**(6): p. 1300-6.
198. Jimenez, C.R., et al., *In-gel digestion of proteins for MALDI-MS fingerprint mapping*. Curr Protoc Protein Sci, 2001. **Chapter 16**: p. Unit 16 4.
199. Vikrant, U.U. Sawant, and A.K. Varma, *Role of MERIT40 in stabilization of BRCA1 complex: A protein-protein interaction study*. Biochem Biophys Res Commun, 2014.
200. Whitmore, L. and B.A. Wallace, *Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases*. Biopolymers, 2008. **89**(5): p. 392-400.
201. Whitmore, L. and B.A. Wallace, *DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data*. Nucleic Acids Res, 2004. **32**(Web Server issue): p. W668-73.
202. Stephens, P.J., et al., *Circular dichroism and magnetic circular dichroism of reduced molybdenum-iron protein of Azotobacter vinelandii nitrogenase*. Biochemistry, 1981. **20**(10): p. 2857-64.
203. Sreerama, N., S.Y. Venyaminov, and R.W. Woody, *Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis*. Anal Biochem, 2000. **287**(2): p. 243-51.
204. Bloom, G.S., et al., *Native structure and physical properties of bovine brain kinesin and identification of the ATP-binding subunit polypeptide*. Biochemistry, 1988. **27**(9): p. 3409-16.
205. Kuznetsov, S.A., et al., *The quaternary structure of bovine brain kinesin*. EMBO J, 1988. **7**(2): p. 353-6.
206. Bloom, G.S. and S.A. Endow, *Motor proteins 1: kinesins*. Protein Profile, 1995. **2**(10): p. 1105-71.
207. Hirokawa, N., Y. Noda, and Y. Okada, *Kinesin and dynein superfamily proteins in organelle transport and cell division*. Curr Opin Cell Biol, 1998. **10**(1): p. 60-73.
208. Hirokawa, N., *Kinesin and dynein superfamily proteins and the mechanism of organelle transport*. Science, 1998. **279**(5350): p. 519-26.
209. Goldstein, L.S. and A.V. Philp, *The road less traveled: emerging principles of kinesin motor utilization*. Annu Rev Cell Dev Biol, 1999. **15**: p. 141-83.

210. Goldstein, L.S., *The kinesin superfamily: tails of functional redundancy*. Trends Cell Biol, 1991. **1**(4): p. 93-8.
211. Nangaku, M., et al., *KIF1B, a novel microtubule plus end-directed monomeric motor protein for transport of mitochondria*. Cell, 1994. **79**(7): p. 1209-20.
212. Ohira, M., et al., *Identification and characterization of a 500-kb homozygously deleted region at 1p36.2-p36.3 in a neuroblastoma cell line*. Oncogene, 2000. **19**(37): p. 4302-7.
213. Conforti, L., et al., *The major brain isoform of kif1b lacks the putative mitochondria-binding domain*. Mamm Genome, 1999. **10**(6): p. 617-22.
214. Hirokawa, N., *Organelle transport along microtubules - the role of KIFs*. Trends Cell Biol, 1996. **6**(4): p. 135-41.
215. Goldstein, L.S., *With apologies to scheherazade: tails of 1001 kinesin motors*. Annu Rev Genet, 1993. **27**: p. 319-51.
216. Yamazaki, H., et al., *Cloning and characterization of KAP3: a novel kinesin superfamily-associated protein of KIF3A/3B*. Proc Natl Acad Sci U S A, 1996. **93**(16): p. 8443-8.
217. Hammet, A., et al., *Rad9 BRCT domain interaction with phosphorylated H2AX regulates the G1 checkpoint in budding yeast*. EMBO Rep, 2007. **8**(9): p. 851-7.
218. Ward, I., et al., *The tandem BRCT domain of 53BP1 is not required for its repair function*. J Biol Chem, 2006. **281**(50): p. 38472-7.
219. Sofueva, S., et al., *BRCT domain interactions with phospho-histone H2A target Crb2 to chromatin at double-strand breaks and maintain the DNA damage checkpoint*. Mol Cell Biol, 2010. **30**(19): p. 4732-43.
220. Zhang, W., et al., *Solution structure of Rap1 BRCT domain from Saccharomyces cerevisiae reveals a novel fold*. Biochem Biophys Res Commun, 2011. **404**(4): p. 1055-9.
221. Liu, K., et al., *Regulation of E2F1 by BRCT domain-containing protein TopBP1*. Mol Cell Biol, 2003. **23**(9): p. 3287-304.
222. Li, S., et al., *Binding of CtIP to the BRCT repeats of BRCA1 involved in the transcription regulation of p21 is disrupted upon DNA damage*. J Biol Chem, 1999. **274**(16): p. 11334-8.
223. Botuyan, M.V., et al., *Structural basis of BACH1 phosphopeptide recognition by BRCA1 tandem BRCT domains*. Structure, 2004. **12**(7): p. 1137-46.
224. Shen, Y. and L. Tong, *Structural evidence for direct interactions between the BRCT domains of human BRCA1 and a phospho-peptide from human ACC1*. Biochemistry, 2008. **47**(21): p. 5767-73.
225. Campbell, S.J., R.A. Edwards, and J.N. Glover, *Comparison of the structures and peptide binding specificities of the BRCT domains of MDC1 and BRCA1*. Structure, 2010. **18**(2): p. 167-76.
226. Feki, A., et al., *BARD1 induces apoptosis by catalysing phosphorylation of p53 by DNA-damage response kinase*. Oncogene, 2005. **24**(23): p. 3726-36.
227. Yaffe, M.B. and A.E. Elia, *Phosphoserine/threonine-binding domains*. Curr Opin Cell Biol, 2001. **13**(2): p. 131-8.
228. Clapperton, J.A., et al., *Structure and mechanism of BRCA1 BRCT domain recognition of phosphorylated BACH1 with implications for cancer*. Nat Struct Mol Biol, 2004. **11**(6): p. 512-8.
229. Elia, A.E., L.C. Cantley, and M.B. Yaffe, *Proteomic screen finds pSer/pThr-binding domain localizing Plk1 to mitotic substrates*. Science, 2003. **299**(5610): p. 1228-31.
230. Williams, R.S., et al., *Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1*. Nat Struct Mol Biol, 2004. **11**(6): p. 519-25.

231. Leslie, A., *Recent changes to the MOSFLM package for processing film and image plate data*. 1992.
232. Winn, M.D., et al., *Overview of the CCP4 suite and current developments*. Acta Crystallographica Section D: Biological Crystallography, 2011. **67**(4): p. 235-242.
233. Matthews, B.W., *Solvent content of protein crystals*. Journal of molecular biology, 1968. **33**(2): p. 491-497.
234. Rossmann, M.G., *The molecular replacement method*. Acta Crystallographica Section A: Foundations of Crystallography, 1990. **46**(2): p. 73-82.
235. Clapperton, J.A., et al., *Structure and mechanism of BRCA1 BRCT domain recognition of phosphorylated BACH1 with implications for cancer*. Nature structural & molecular biology, 2004. **11**(6): p. 512-518.
236. McCoy, A.J., et al., *Phaser crystallographic software*. Journal of applied crystallography, 2007. **40**(4): p. 658-674.
237. Schuttelkopf, A.W. and D.M. Van Aalten, *PRODRG: a tool for high-throughput crystallography of protein-ligand complexes*. Acta Crystallographica Section D: Biological Crystallography, 2004. **60**(8): p. 1355-1363.
238. Wallace, A.C., R.A. Laskowski, and J.M. Thornton, *LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions*. Protein engineering, 1995. **8**(2): p. 127-134.
239. Kleywegt, G.J., *Validation of protein crystal structures*. Acta Crystallogr D Biol Crystallogr, 2000. **56**(Pt 3): p. 249-65.
240. Williams, R.S., R. Green, and J.M. Glover, *Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1*. Nature Structural & Molecular Biology, 2001. **8**(10): p. 838-842.
241. Arver, B., et al., *Hereditary breast cancer: a review*. Semin Cancer Biol, 2000. **10**(4): p. 271-88.
242. Rosen, E.M., et al., *BRCA1 gene in breast cancer*. J Cell Physiol, 2003. **196**(1): p. 19-41.
243. Birrane, G., et al., *Crystal structure of the BARD1 BRCT domains*. Biochemistry, 2007. **46**(26): p. 7706-12.
244. Loeffler, P.A., et al., *Structural studies of the PARP-1 BRCT domain*. BMC Struct Biol, 2011. **11**: p. 37.
245. Taylor, R.M., A. Thistlethwaite, and K.W. Caldecott, *Central role for the XRCC1 BRCT I domain in mammalian DNA single-strand break repair*. Mol Cell Biol, 2002. **22**(8): p. 2556-63.
246. Joerger, M., et al., *Tubulin, BRCA1, ERCC1, Abraxas, RAP80 mRNA expression, p53/p21 immunohistochemistry and clinical outcome in patients with advanced non small-cell lung cancer receiving first-line platinum-gemcitabine chemotherapy*. Lung Cancer, 2011. **74**(2): p. 310-7.
247. Greenberg, R.A., *BRCA mutations and childhood cancer*. Cancer Biol Ther, 2006. **5**(9): p. 1103-4.

Structural and Functional Implication of RAP80 Δ Glu81 Mutation

Vikrant¹, Rajan Kumar¹, Lumbini R. Yadav¹, Pallavi Nakhwa¹, Sanjeev K. Waghmare¹, Peyush Goyal², Ashok K. Varma^{1*}

1 Advanced Centre for Treatment, Research and Education in Cancer, Kharghar, Navi Mumbai, Maharashtra, India, **2** Department of Biotechnology, CGO Complex, New Delhi, India

Abstract

Receptor Associated Protein 80 (RAP80) is a member of RAP80-BRCA1-CCDC98 complex family and helps in its recruitment to the DNA damage site for effective homologous recombination repair. It encompasses two tandem UIMs (UIM1 and UIM2) motif at its N-terminus, which interact with K-63 linked polyubiquitin chain(s) on H2AX and thereby assemble the RAP80-BRCA1 complex at the damage site. Nevertheless, how RAP80 helps in the structural integrity of BRCA1 complex is still elusive. Considering the role of RAP80 in the recruitment of BRCA1 complex at the DNA damage site, we attempted to explore the molecular mechanism associated with RAP80 and mutation that causes chromosomal aberrations due to its loss of function. There is a significant loss in structural characteristics of RAP80 Δ E81, which impairs its binding affinity with the polyubiquitin chain. This leads to the defective recruitment of RAP80 and BRCA1 complex at the DNA damage site. The results presented here are very useful in understanding the cause of various repair defects (chromosomal aberration) that arise due to this mutation. Comparative study of wild type and Δ E81 could be helpful in designing the small molecules that can potentially compensate the deleterious effect(s) of Δ E81 and hence useful for therapeutic application.

Citation: Vikrant, Kumar R, Yadav LR, Nakhwa P, Waghmare SK, et al. (2013) Structural and Functional Implication of RAP80 Δ Glu81 Mutation. PLoS ONE 8(9): e72707. doi:10.1371/journal.pone.0072707

Editor: Michael Massiah, George Washington University, United States of America

Received: January 23, 2013; **Accepted:** July 15, 2013; **Published:** September 9, 2013

Copyright: © 2013 Vikrant et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Funding for this study was supported by Department of Biotechnology, Government of India (No. BT/PR10765/BRB/664/2008 and BT/PR12565/BID/07/303/2009), Seed in Air grant from Tata Memorial Centre to AKV. Vikrant is supported by CSIR fellowship (09/513 (0072)/2008-EMR-I). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: avarma@actrec.gov.in

Introduction

Compromised genomic integrity leads to various genetic disorders and cancer. However, genomic stability is accomplished by the recital action of several cellular events, including DNA replication, DNA repair, senescence and cell death [1]. Cells have evolved a complex, dynamic and highly regulated network to achieve extreme fidelity, called DNA damage response (DDR). In genotoxic stress, DDR coordinates numerous cellular processes like cell cycle regulation, chromatin remodeling, DNA repair and transcription [2]. Sensing of DNA damage and promulgation of the DDR signaling cascade involve recruitment and assembly of many DDR mediators and effectors at the sites of damage [3] [4]. Double strand breaks elicit the activation of ATM and ATR kinases, which phosphorylate histone variant H2AX and MDC1 [5] [6] [7] [8] [9,10] [3]. This event endorses the assembly of DDR mediators, which in turn facilitate the recruitment of UBC13/RNF8 to the DNA damage sites [11] [12,13] [14]. In the signaling pathways, eventually this leads to the formation of polyubiquitin chains on H2AX, which are recognized by RAP80 [7,8,9] [10]. RAP80 has two tandem UIM (Ubiquitin-Interacting Motif) at its N-terminus, ABRAXAS (CCDC98) Interacting Region (AIR) at the central domain, and two zinc finger domains at its C-terminus [15]. It has been reported that RAP80 forms a stable complex with BRCA1 through an intermediate binding partner CCDC98 [16,17,18]. CCDC98 has a consensus sequence

SXXF motif at C-terminus, which involves in interaction with BRCA1-BRCT phosphospecific binding domain [16,18] [19,20]. BRCA1 acts as a tumor suppressor gene in hereditary breast and ovarian cancer, and plays a diverse role in cell cycle regulation, transcription control and DNA damage repair [21,22,23,24,25]. C-terminus of BRCA1 (BRCT) is essential for its co-localization with H2AX [26].

RAP80 acts upstream of CCDC98 and BRCA1 in DDR, and is required for the localization of the BRCA1 complex to ionizing radiation (IR)-induced foci (IRIFs) [17,18,27]. RAP80 Knock-down cells showed hypersensitivity to IR and ultraviolet (UV) light, cell cycle dysfunction and defective homologous recombination (HR) repair [10,16,17,18]. RAP80 and p53 auto-regulate each other and has influence on apoptosis [28]. Loss of RAP80 alleles (RAP80^{-/-}) increase the susceptibility to lymphoma, and promote tumor development in both p53^{-/-} and p53^{+/-} mice [29]. UIM1 and UIM2 motifs of RAP80 are very crucial since deletion of either or both significantly perturb the foci formation of RAP80-BRCA1 complex at the DNA damage site [30].

A novel alteration, c.241–243delGAA (Δ E81) that leads to an inframe deletion of glutamic acid residue has been identified at UIM1 motif of RAP80 [30]. The RAP80 Δ E81 variant was found in a patient diagnosed with breast cancer, and is highly conserved among all the vertebrates. This variant showed an observed frequency of 0.9% (1/112) in the familial cases compared to 0.3% (1/325) in the controls (P₁/4.0.45; OR₁/4.2.92; CI₁/4.0.18–47.1). One

RAP80 Δ E81 carrier was also diagnosed with bilateral breast cancer in a group of 503 breast cancer cases (0.2%, 1/503). RAP80 Δ E81 expressing cells showed abrogation of DSB localization of the RAP80–BRCA1 complex and exhibited genomic instability (chromosomal aberration) [30]. In this study, we have carried out a comparative structural, stability and binding analysis of RAP80 (1–130) wild type (referred as RAP80 wild type or wild type henceforth) and RAP80 (1–130) Δ E81 (referred as RAP80 Δ E81 or Δ E81 henceforth) to understand the functional implication(s) of this mutation. To our knowledge, this is the first multi model approach combining *in-silico* and *in-vitro* methods to study the functional implications of RAP80 wild type and the Δ E81. RAP80 Δ E81 relatively exhibited less thermal stability and significant secondary structure distortion, which impaired its binding affinity with di (poly)-ubiquitin. This further leads to defective recruitment of RAP80–BRCA1 complex to the DNA damage site and subsequently giving rise to genomic instability. Our study will be helpful in understanding the role of UIM motifs of RAP80 in RAP80–BRCA1 complex recruitment and hence their DNA damage repair function. It will further assist in elucidation of mechanism that alters the binding affinity of RAP80 UIMs for polyubiquitin chain due to Δ E81 mutation, and thereby its implication on damage repair.

Results and Discussion

RAP80 is 80 KDa nuclear protein that interacts with retinoid-related testis-associated receptor [15]. It is a member of BRCA1 complex and facilitates the recruitment of BRCA1 to the DNA damage site. Thus, it is a multifunctional molecule that plays a dispersive role in steroid hormone signaling, and BRCA1 mediated homologous recombination repair. SiRNA mediated silencing, and knockout studies of RAP80 showed defective recruitment of BRCA1 complex and hence the perturbed DNA repair [29,31,32,33]. *In-vitro* and *in-silico* findings from our study, will be helpful in understanding the mutational consequence of RAP80 Δ E81 in DNA damage and repair pathway. To our knowledge, this is the first report on a comparative functional characterization of RAP80 wild type and Δ E81.

Structural Organization of RAP80

Coomassie stained SDS-PAGE for RAP80 wild type and Δ E81 showed a single band corresponding to 14 KDa (**Figure 1A, B**). A single peak spectrum was observed in size exclusion chromatography (**Figure 1C**). Purified proteins were further subjected to MALDI-TOF (Matrix Assisted Laser Desorption Ionization -Time of Flight), and spectra corresponding to 14.958 KDa and 14.815 KDa for RAP80 wild type and Δ E81 respectively, were recorded with greater sensitivity. We found a close match between experimentally derived (wild type: 14.958 KDa, Δ E81 14.815 KDa) and theoretically predicted molecular weight (wild type: 14.898 KDa, Δ E81 14.751 KDa) (**Table 1**). The presence of single peak in mass spectroscopy and size exclusion chromatography indicates monomeric behavior of RAP80 wild type and Δ E81 (**Figure 1C**).

RAP80 (79–124) UIMs Δ E81 structure was successfully modeled using protein modeler [34,35] with a acceptable Ramachandran plot [36] [37]. UIM1 and UIM2 are connected with a linker in a head to tail manner. The three-dimensional structure of wild -type looks overall 59 Å long and α -helical in nature. However, in case of mutant, α -helix is partly distorted and shorten to 45 Å. UIM1 and UIM2 bind with their respective proximal and distal ubiquitin of Di-Ub (K-63 linked) in 1:1 affinity ratio [38] [39]. Glu residue at 81 position was found to be highly

conserved (**Figure 2C**) and forms ionic bond and hydrophobic interaction, with the Arg42 and Leu73 residue of proximal ubiquitin, respectively. It is widely reported that hydrogen bonding and hydrophobic interactions play an important role in protein stability and selection of the specific target [40]. There are changes in weak intermolecular interactions between RAP80 UIMs, RAP80 UIMs Δ E81 and Di-Ub (K-63 linked) (**Figure 2A, B**). The hydrogen bonds between Gln84, Ser92, Glu95, Ser117, Gln102 residues of RAP80 UIMs and the Leu8, Gly47, Thr66, His68, Arg72 of ubiquitin, and the hydrophobic interactions between Ser 92, Ser 117 of RAP80 UIMs and Ile44, Phe45, Ala46, Gly47, His68 of proximal ubiquitin are stabilizing the binding interface. However, a drastic conformational change in RAP80 UIMs Δ E81 was observed which significantly alter the weak intermolecular interactions with ubiquitin. Met 79, Glu 83 and Glu 93 of UIMs are involved in hydrogen bonding with His 68, Gly 47 of ubiquitin. Hydrophobic interactions between the Met 79, Arg122, residues of RAP80 UIMs Δ E81 with the Phe4, Leu43, Ile44, Phe45, Gly47, Lys48, Gln49, Leu50, Glu64, Ser65, Thr66, His68 residues of ubiquitin primarily holds the complex. Structural distortion in RAP80 UIMs Δ E81 probably renders its binding interaction unfavorable with Di-Ub (K-63 linked).

To understand structural integrity and determine the resistivity of RAP80 wild type and Δ E81 against the protease digestion, limited trypsin and chymotrypsin proteolysis was performed. RAP80 wild type and Δ E81 were treated with same concentration of proteases for limited time (**Figure 3A, 3B, 3C, 3D**). RAP80 wild type resistance against protease digestion gives the indication of having a relatively stable domain and well-formed structure. However, susceptibility of RAP80 Δ E81 towards protease digestion suggests that deletion of E81 is responsible for destabilizing the structural integrity of RAP80. Furthermore, we have compared the changes in secondary structure using far-UV circular Dichroism (**Figure 4A**). It was observed that RAP80 wild type has well-defined α/β characteristics whereas structure of Δ E81 showed deviation from typical α/β characteristic to random structure. Earlier report suggests that UIMs motif of RAP80 is found in equilibrium between α -helix and random structure [41]. Δ E81 mutation probably alters the α -helical conformation of RAP80 UIMs which leads to shift the equilibrium towards a random structure pattern.

Thermal stability

Stability profiles of RAP80 wild type and Δ E81 was compared at secondary (CD) and tertiary (Fluorescence) structure levels. The spectra obtained from Circular Dichroism corresponding to λ at 218 nm showed the maximum change in ellipticity and high signal to noise ratio (**Figure 4B**). Thermal stability of RAP80 Δ E81 (T_m 22°C, $\Delta G^\circ_{H_2O}$ 1.3 \pm 0.2 Kcal/mol, ΔH 1.0 \pm 0.5 Kcal/mol) was found significantly low compared to wild type (T_m 29°C, $\Delta G^\circ_{H_2O}$ 2.0 \pm 0.5 Kcal/mol, ΔH 5.0 \pm 2.0 Kcal/mol). ANS (8-Anilino-naphthalene-1-sulfonate) fluorescence spectroscopy has an agreement with CD data, and derived T_m value was 23°C for Δ E81 ($\Delta G^\circ_{H_2O}$ 1.4 \pm 0.3 Kcal/mol, ΔH 1.1 \pm 0.5 Kcal/mol) and 30°C for RAP80 wild type ($\Delta G^\circ_{H_2O}$ 2.4 \pm 0.5 Kcal/mol, ΔH 8.0 \pm 1.1 Kcal/mol) (**Figure 4C**). Both the methods showed that protein most likely unfolds without any intermediate species. These findings were further supported by Differential Scanning Calorimetry, which gave a T_m value of 28°C for RAP80 wild type (**Figure 4D**). However, we could not obtain a defined transition for Δ E81, due to lesser stability and saturation concentration (**Table 2**). These results suggest that three-dimensional folding of RAP80 Δ E81 is impaired in comparison to wild type. These findings also support the helix to random structure transition of

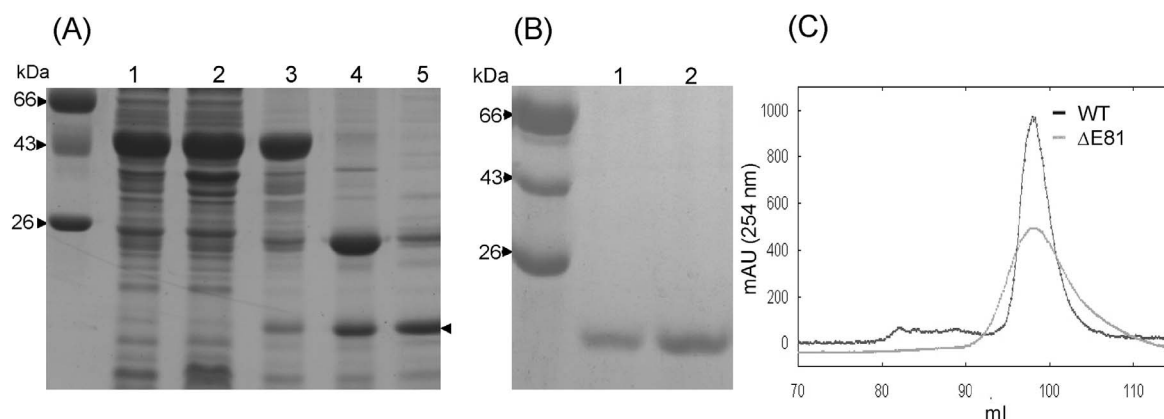


Figure 1. Expression and purification profile of RAP80 wild type and $\Delta E81$. (A) Whole-cell lysate, and supernatant obtained after sonication and centrifugation were heated with Laemmli buffer and loaded onto SDS-PAGE. Similarly, protein was eluted from beads by heating with Laemmli buffer and loaded on gel. Lane 1-Total protein, 2-soluble protein, 3-fusion protein bound on beads, 4- protein after on beads cleavage, 5-elution fraction of affinity purified proteins. Single arrow - RAP80 wild type protein (B) Purified protein after gel filtration chromatography on SDS-PAGE. Lane 1- RAP80 $\Delta E81$, 2- RAP80 wild type (C) Overlay of gel filtration spectra of RAP80 wild type and $\Delta E81$ (Superdex 200). Elution profiles of both the protein were similar and suggest their monomeric nature.
doi:10.1371/journal.pone.0072707.g001

UIMs motif. $\Delta E81$ mutation probably shifts this transition equilibrium towards the random structure.

Binding interaction of RAP80 wild type and $\Delta E81$ with di-Ub (K-63 linked)

It is well reported that RAP80 UIMs bind with K-63 linked polyubiquitin chain(s) on the H2AX and recruit the RAP80-BRCA1 complex to the DNA damage site [18] [27]. Considering structural distortion and stability of RAP80 $\Delta E81$, it can be suspected that it would further impair binding affinity for polyubiquitin chain. Binding analysis between RAP80 wild type and $\Delta E81$ with Di-Ub (K-63 linked) has been performed using Surface Plasma Resonance (SPR) and GST pull down assay. The observed binding affinity for RAP80 $\Delta E81$ (K_D : 0.459 μM) was several fold less as compared to wild type (K_D : 36.5 nM) in SPR (Figure 5 A, 5B). Association rate constant of RAP80 $\Delta E81$ was found significantly lower (K_a : $4.306 \times 10^1 M^{-1} s^{-1}$) than wild type (K_a : $3.06 \times 10^5 M^{-1} s^{-1}$). Besides this, RAP80 $\Delta E81$ showed high dissociation rate as compared to wild type. Furthermore, association constant of wild type is higher than $\Delta E81$ [K_A (Wild Type): $2.74 \times 10^7 M^{-1}$, K_A ($\Delta E81$): $2.18 \times 10^6 M^{-1}$]. GST pull down assay also supported the finding obtained using SPR (Figure 5C). It can be concluded that RAP80 wild type has higher binding affinity for the polyubiquitin chain, besides, it associates faster than $\Delta E81$. Mutant protein complex { $\Delta E81$ -Di (Ub)} was likely unstable due

to high dissociation rate and less binding affinity. Alteration in binding affinity of RAP80 $\Delta E81$ could be due to its structural deformation.

Conclusion

RAP80 wild type and $\Delta E81$ are moderately soluble. Thermal and proteolytic stability of wild type was found significantly higher as compared to $\Delta E81$, but both unfold likely with two state irreversible transition. RAP80 UIMs are found in equilibrium between random-coil and helical states. This fact is supported by low T_m values of both wild type and $\Delta E81$. The reason behind dynamic nature of UIMs is to provide immense flexibility of dissociation and association of ubiquitin molecules during the protein trafficking process. Perhaps UIMs also use this mechanism for multiple mode of binding (monovalent and multivalent) so as to achieve cooperativity in binding interactions. This dynamic nature is essential for a flexible and transient initiation mechanism of the DNA damage repair process. Deletion of 81E residue perhaps alters the helical state conformation, thus shifting equilibrium towards a random structure. Helical to random structure transition results in loss of several weak intermolecular hydrogen bonds and hydrophobic interactions between the UIMs and Di-Ub (K-63 linked), thereby making the binding interactions unfavorable for ubiquitin. Since binding affinity of individual UIM for mono-ubiquitin is low [42], an avidity-based mechanism

Table 1. Molecular weight estimation of purified protein.

	Theoretical Mol. Wt. (kDa) ^a	Ve/Vo ^b	Experimental Derived Mol. Wt. (kDa)	
			Gel Filtration Chromatography	Mass spectrometry (MALDI-TOF)
Wild type	14.897	2.107527	15.8	14.9
$\Delta E81$	14.750	2.107527	15.8	14.8

Ve/Vo: Elution volume/Void volume ratio in gel filtration chromatography (superdex 200 16/60).

^aDetermined from ProtParam, ExPASy.

^bDetermined from standard myoglobin, ovalbumin, albumin, IgG, Ferritin.

doi:10.1371/journal.pone.0072707.t001

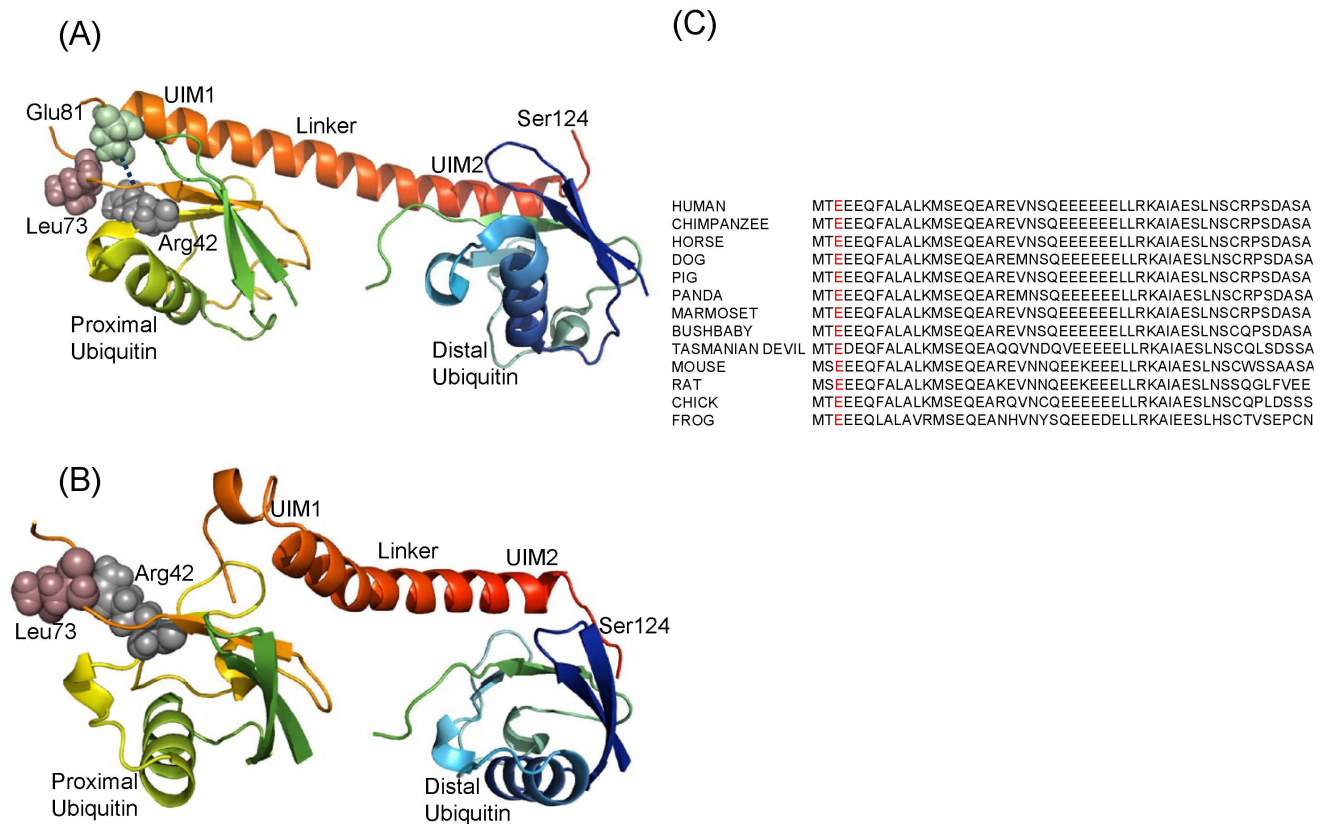


Figure 2. Binding interaction of RAP80 UIMs and Δ E81 with Di-Ub (K-63 linked). (A) Structure of Di-Ub (K-63 linked)-RAP80 UIMs (79–124) wild type (PDB ID: 2RR9), and (B) Di-Ub (K-63 linked)-RAP80 (79–124) UIMs Δ E81. Wild type and Di-Ub (K-63 linked) complex is stabilized by weak intermolecular interactions. α -helix of RAP80 (79–124) UIM Δ E81 was found to be distorted. (C) multiple sequence alignment of UIMs region showed it's highly conserved nature in various species. Glu 81 residue is highlighted in red color.

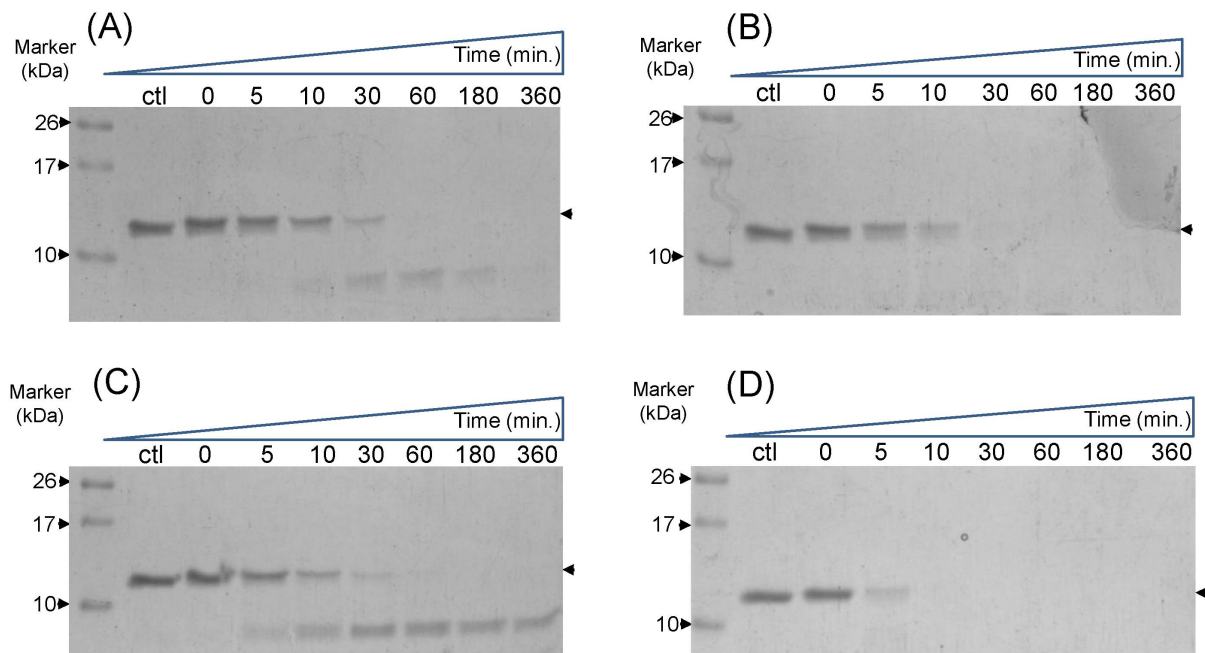


Figure 3. Resistivity profile of RAP80 wild type and Δ E81 towards Protease digestion. Limited proteolysis of RAP80 wild type (**A, C**) and Δ E81 (**B, D**) using trypsin (**A, B**) and Chymotrypsin (**C, D**) as proteases. Wild type showed relatively high resistance towards proteolysis as indicated by less rate of decrease of band intensity. This suggests a well-folded structure of wild type compared to Δ E81. Ctl- control was taken as untreated with proteases. doi:10.1371/journal.pone.0072707.g003

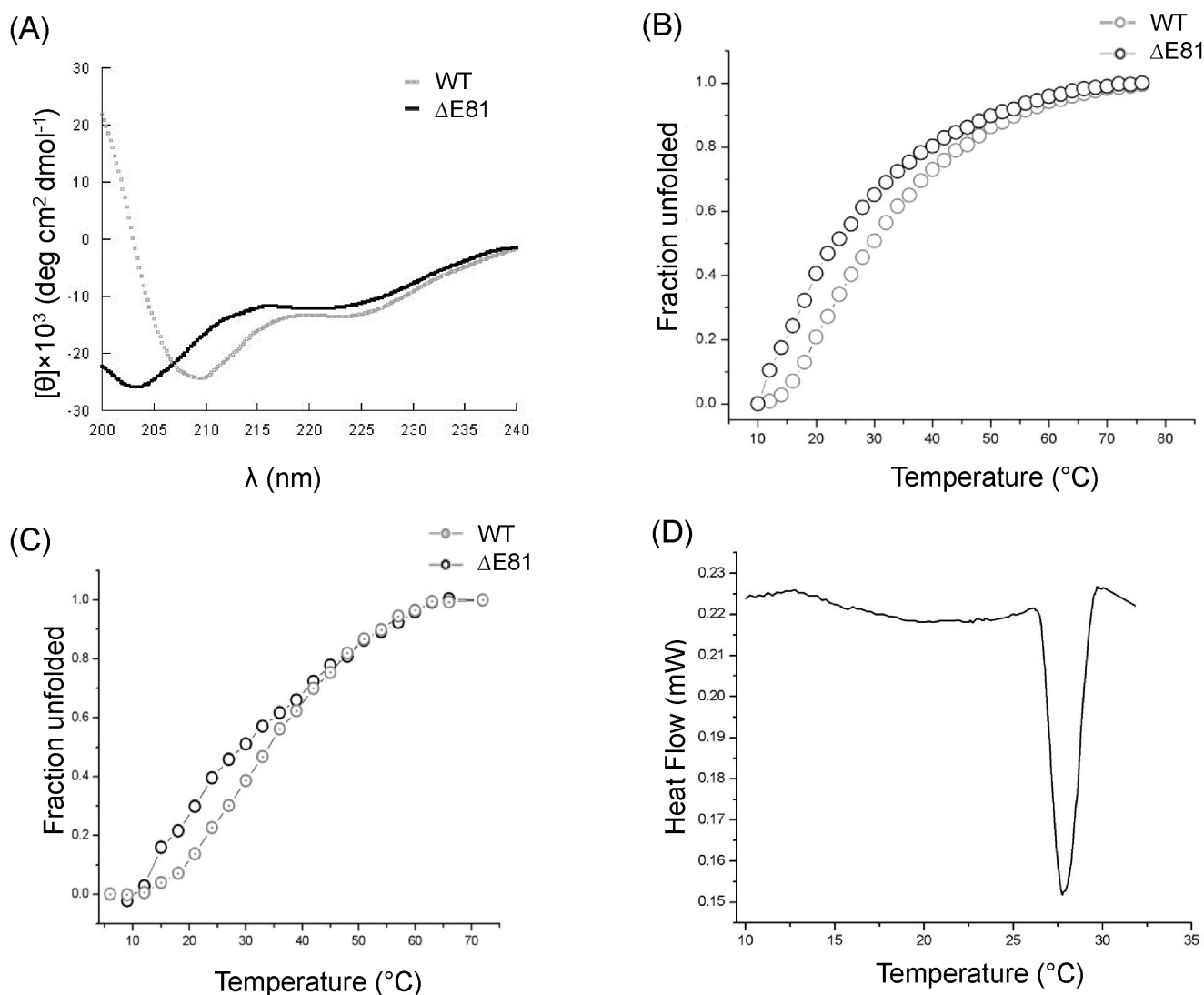


Figure 4. Structure and stability analysis of RAP80 wild type and $\Delta E81$. Secondary structural components and thermal stability of RAP80 wild type and $\Delta E81$. (A) Overlay of Far-UV Circular Dichroism spectrum of wild type and $\Delta E81$. Wild-type showed well-defined α/β characteristics compared to a random structure pattern of $\Delta E81$. Thermal stability of RAP80 wild type. (B) Thermal denaturation of RAP80 wild type and $\Delta E81$ using ANS as extrinsic fluorophore in Fluorescence. Unfolded fractions were calculated and plotted against different temperatures. (D) Differential Scanning Calorimetry profile of RAP80 wild type. Protein showed a well-defined transition around 28°C.
doi:10.1371/journal.pone.0072707.g004

Table 2. Thermal parameters of protein unfolding.

Method	Protein	T_m (°C)	$\Delta G^{\circ}_{H_2O}$ (Kcal/mol)	ΔH (Kcal/mol)
DSC	Wild type	28	-	8.7 ± 1.0
Fluorescence	Wild type	23	2.4 ± 0.5	8.0 ± 1.1
	$\Delta E81$	30	1.4 ± 0.3	1.1 ± 0.5
CD	Wild type	29	2.0 ± 0.5	5.0 ± 2.0
	$\Delta E81$	22	1.3 ± 0.2	1.0 ± 0.5

T_m Melting Temperature.
doi:10.1371/journal.pone.0072707.t002

probably makes the interaction between RAP80 and Lys 63-linked polyubiquitin highly robust. Co-operative binding between multiple UIMs and ubiquitin chains likely occurs, which favors the interaction of second UIM with ubiquitin after positioning of the first [39]. It has been reported [30] that expression of RAP80 $\Delta E81$ allele abates recruitment of BRCA1 complex at DSB site, which further augment chromosomal aberration (chromatic breaks). The results presented in this study also suggest that deletion of 81 Glutamic acid residue significantly obliterates RAP80 structure and impairs its binding with polyubiquitin chain. Unstable nature of mutant and di-ubiquitin complex may be responsible for defective recruitment of RAP80-BRCA1 complex to the DNA damage sites. Defective DNA damage repair perhaps leads to chromosomal aberration as shown in the model (Figure 6). Prolific comparison of RAP80 $\Delta E81$ with wild type will help in understanding its role in various diseases and repair defects. It will further explore the possibility of structure

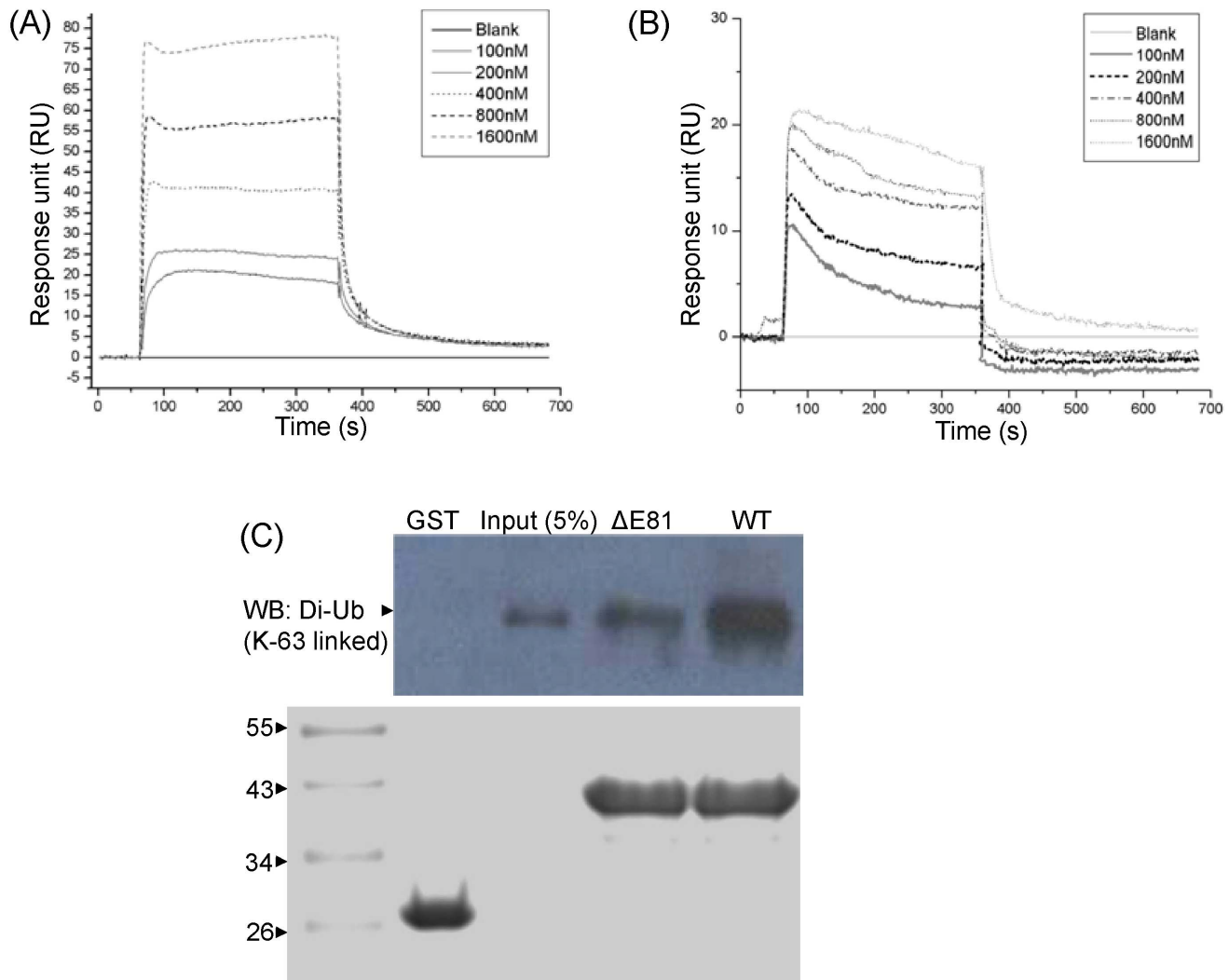


Figure 5. Binding analysis of RAP80 wild type and $\Delta E81$ with Di-Ub (K-63 linked). Sensogram of RAP80 wild type (A) and $\Delta E81$ (B) determined by Surface Plasma Resonance. 5 μ g of ligand (Di-Ub (K-63 linked)) was immobilized on CM5 sensor chip and different concentrations of analytes (wild type and $\Delta E81$) were passed. (C) GST pull down assay followed by western blotting. GST-RAP80 wild type and $\Delta E81$ were used as a bait and Di-Ub (K-63 linked) as prey. Di-Ub (K-63 linked) was probed with anti-ubiquitin antibody. Ponceau stained PVDF membrane showing the GST and GST fusion protein as bait(s). Wild type showed high binding affinity compare to $\Delta E81$. GST was taken as control.
doi:10.1371/journal.pone.0072707.g005

based inhibitor design for therapeutic application that can compensate the effect of such mutation.

Materials and Methods

Molecular biology or analytical grade chemicals were purchased from Sigma-Aldrich, unless otherwise specified with more than 99.99% purity. Restriction enzymes were purchased from Fermentas.

Gene cloning, protein expression and purification

Q96RL1 gene (1–390) in pGFP vector (Kind gift from J. Chen) was PCR amplified (Thermocycler, Biorad) followed by restriction digestion (BamH1/EcoR1), T4-ligation and cloned into pGEX-kT (Kind gift from John A. A. Ladias) vector. Primers (Sigma-Aldrich) having a TEV protease cleavage site (E-N-L-Y-F-Q/S) were used for PCR amplification. Positive clones were selected by restriction digestion followed by DNA sequencing. c.241–243delGAA mutation was incorporated into

wild type gene construct using site-directed mutagenesis. PCR amplified product was digested with Dpn1 (Fermentas) and transformed into *E. coli* DH5 α bacterial strain. Incorporation of desired mutation was confirmed by DNA sequencing. For protein expression and purification, vector construct was transformed into *E. coli* BL21 (DE3) cells (Novagen) and a single colony was inoculated in LB broth to obtained pre-inoculum culture. Protein was over-expressed in *E. coli* BL21 (DE3), and culture was grown at 37°C until O.D₆₀₀ reached between 0.6–0.8, followed by induction with 0.4 mM IPTG at 18°C overnight. Harvested bacterial pellet was re-suspended in 10 mM HEPES buffer containing 300 mM NaCl, 5 mM BME, 0.1 mM EDTA and 5% ethylene glycol at pH 7.5 (HNBEEG buffer). Cells were disrupted by sonication (Branson Sonifier) and supernatant was collected after centrifugation. Soluble protein was passed through the pre equilibrated glutathione resin and then washed with HNBEEG buffer to remove impurities. Bound fusion protein was cleaved with TEV protease to elute the protein of interest. Protein was further

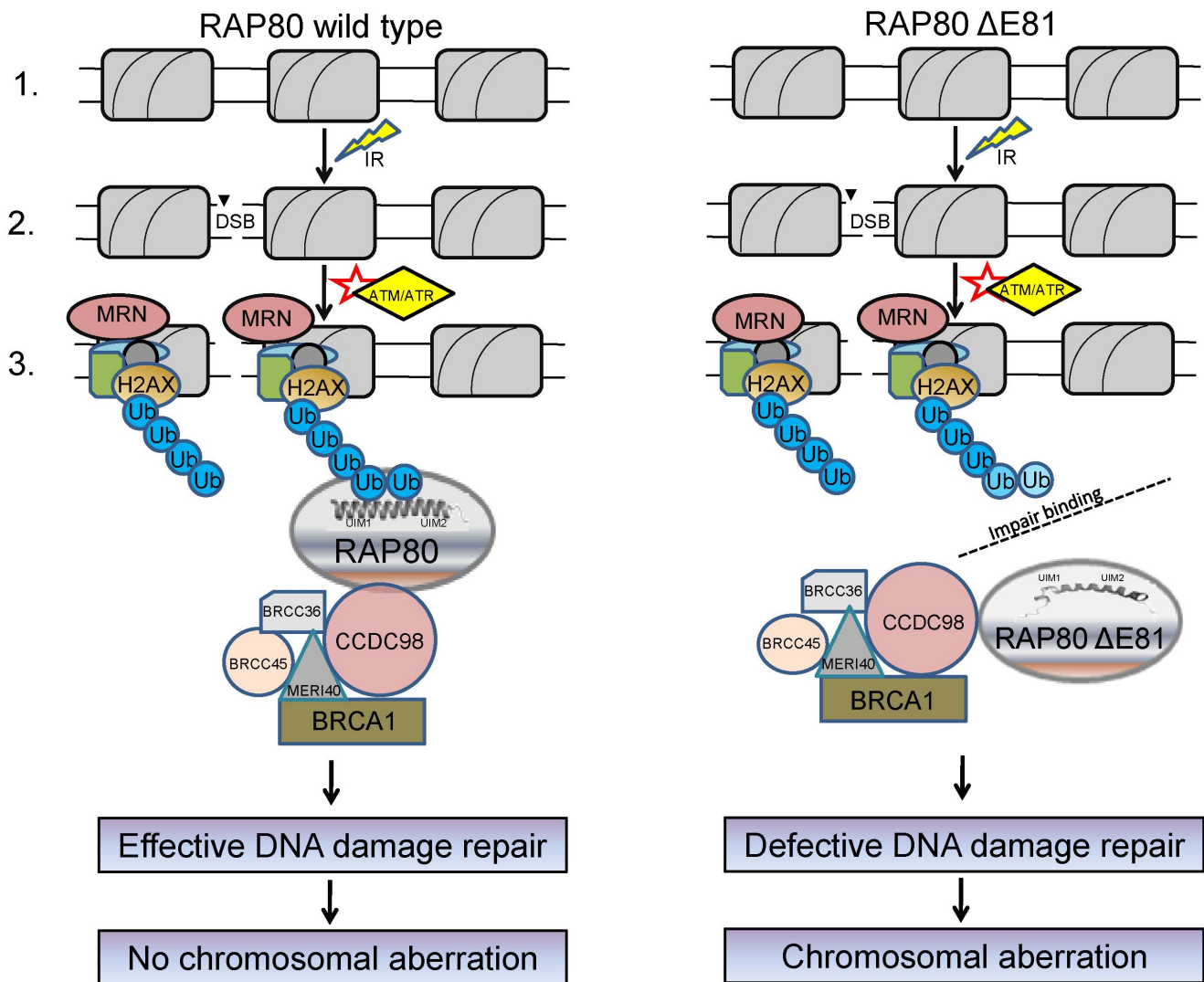


Figure 6. Anticipated mechanism of consequence due to RAP80 Δ E81. The model elucidate a possible mechanism of chromosomal aberration due to RAP80 Δ E81 mutation. In the wild-type RAP80: **Step1**, showed the intact nucleosome complex; **Step 2**, double strand break due to ionization radiation; **Step 3**, ATM/ATR kinase activation and assembly of various damage repair proteins at DNA double strand break (DSB) site followed by formation of polyubiquitin chain(s) on histone(s) (H2AX). The polyubiquitin chain(s) are recognized by RAP80 UIMs motif thereby recruiting the BRCA1 complex to the DNA damage site. However, in case of Δ E81 mutation, interaction between polyubiquitin chain and RAP80 UIM altered due to structural distortion in α -helix which further leads to defective recruitment of the BRCA1 complex. Error-prone DNA damage repair increases the chances of chromosomal aberration and hence the tumorigenesis.

doi:10.1371/journal.pone.0072707.g006

passed through a gel filtration column (Superdex 200, GE) to remove aggregates, etc. and was analyzed using SDS-PAGE for purity.

Protein Estimation

Quantification of RAP80 wild type and Δ E81 were performed with Bradford protein estimation protocol according to manufacturer's (expedon) instructions. Several dilutions of BSA were prepared as a standard reference. The absorbance was recorded in three sets at λ 595 nm using a spectrophotometer (Shimadzu). Average values were considered, and concentration of sample was determined by intra-plotation of BSA standard curve [43] [44].

Molecular Modeling and docking

Protein structures RAP80 (Δ E81UIMs, 79–124 amino acids) was modeled using homology modeling server considering NMR structure (PDB ID; 2RR9) as template. Good-quality models were selected based on overall stereo chemistry, and validated using Ramachandran plot and protein structure validation server “**SAVES**” (Metaserver for analyzing and validating protein structures, <http://nihserver.mbi.ucla.edu/SAVES/>). SAVES mainly comprises five programs, Procheck, What_check, Errat, Varify_3D and Prove. Modeled structure was simulated for 5 ns using Desmond software (Schrodinger) and superimposed on wild type complex. PDBsum was obtained to analyze the interactions.

Limited proteolysis

Equal concentration of RAP80 wild type and Δ E81 (0.2 mg/ml) was incubated with Trypsin and chymotrypsin separately so that final concentrations of proteases were 40 μ g/ μ l and 10 μ g/ μ l respectively. Reaction mixture was incubated for different time 0, 10, 30, 60, 180, 360 minutes at 37°C (trypsin) and 25°C (chymotrypsin), respectively. Reaction was terminated individually by adding 1 mM PMSF (sigma-Aldrich). Samples were heated by adding equal volume of laemmli buffer and analyzed by SDS-PAGE. This experiment was performed in three sets with control which was untreated with proteases [45] [46].

Surface Plasmon Resonance

Interaction studies between RAP80 wild type, Δ E81 and di-Ub (K63 linked) were performed using BIAcore 3000 (GE). A total of 5 μ g ligand (Di-Ub K-63 linked) was immobilized on CM5 sensor chip using amide coupling method. Different concentration (0, 100, 200, 400, 800, 1600 nM) of RAP80 wild type and Δ E81 (analytes) were passed on the chip at a flow rate of 20 μ l/min. Interaction was quantified in terms of Response unit (RU). Sensor chip was regenerated with 2 M glycine pH 2.0. Sensogram was obtained after blank correction. The experiment was repeated thrice.

GST pull down assay

Bacterial pellet of GST-RAP80 wild type and Δ E81 were resuspended in HNBEEG buffer and sonicated. Soluble fusion protein(s) bound on glutathione resin (0.5 mg/ml) was used to capture prey {Di-Ub (K-63 linked) 10 μ g, Boston Biochem}. Resin was pre-equilibrated with same buffer and loaded on SDS-PAGE. Complex was transferred to PVDF membrane (Millipore) and was probed with anti-ubiquitin antibody (Abcam). The experiment was repeated thrice by taking GST as control.

Circular Dichroism

Far-UV CD spectrum were recorded using a Circular Dichroism (CD) polarimeter (Jasco J-810, Japan). 10 μ M protein (in 2.5 mM HEPES pH 7.5, 50 mM NaCl) was scanned in a wavelength range of 200–240 nm at 10°C. Average blank corrected data of three independent scans were considered. Molar ellipticity was calculated, and data analysis was done using DichroWeb server (<http://dichroweb.cryst.bbk.ac.uk>) [47] [48] [49] [50] [51]. For thermal denaturation, wild type and Δ E81 protein (10 μ M) were unfolded in a temperature range of 10–60°C at 218 nm wavelength. Fraction unfolded was calculated at the different temperatures. The experiment was performed three times

independently, and an average data was considered. Data fitting was done according to two-state transition model, and thermodynamic parameters were calculated.

ANS Fluorescence spectroscopy

The ANS (1-Anilino-8-Naphthalene Sulfonate) fluorescence was monitored using a Fluorescence spectrophotometer (Horiba, USA) at an excitation wavelength of 360 nm. For thermal denaturation, 2 μ M protein (wild type and Δ E81) was incubated with 10 μ M ANS for 10 min and emission scans were recorded from wavelength 400–600 nm in a temperature range of 5–60°C. Thermodynamic parameters were obtained by curve fitting according to two-state transition models [52]. These experiments were performed three times independently, and average blank corrected data was considered for curve fitting in two-state transition models [53].

Differential Scanning Calorimetry

Thermal unfolding of wild type and Δ E81 was done using Differential Scanning Calorimetry (Setaram μ DSC3 evo, USA). Protein and buffer were filtered and degassed prior to the scan. A total of 2 mg protein (RAP80 wild type) and 0.2 mg (Δ E81) in solution form was allowed to unfold in 5–60°C temperature range with a temperature increment rate of 1°C/min. The experiment was repeated thrice independently. Data was fitted locally by “CALISTO” software according to two-state transition model. The thermodynamic reversibility of the protein unfolding was determined by heating the sample just above the transition maximum, cooling instantaneously, and then reheating. Thermal denaturation transitions were found irreversible due to absence of transition(s) in second run.

Acknowledgments

We thank DBT-BTIS facility at ACTREC for providing necessary software to this study. We are thankful to Smita Mahale and Jennifer-NIIRH for SPR facility, M.V Hosur and Lata-BARC for DSC experiment and data analysis.

Author Contributions

Conceived and designed the experiments: V AKV. Performed the experiments: V RK LRY PN AKV. Analyzed the data: V PG SKW AKV. Wrote the paper: V AKV.

References

- Shiotani B, Zou L (2009) ATR signaling at a glance. *J Cell Sci* 122: 301–304.
- Wang B, Hurov K, Hofmann K, Elledge SJ (2009) NBS1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control. *Genes Dev* 23: 729–739.
- Harper JW, Elledge SJ (2007) The DNA damage response: ten years after. *Mol Cell* 28: 739–745.
- Paul TT, Rogakou EP, Yamazaki V, Kirchgessner CU, Gellert M, et al. (2000) A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Curr Biol* 10: 886–895.
- Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* 273: 5858–5868.
- Petrini JH, Stracker TH (2003) The cellular response to DNA double-strand breaks: defining the sensors and mediators. *Trends Cell Biol* 13: 458–462.
- Huen MS, Grant R, Manke I, Minn K, Yu X, et al. (2007) RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly. *Cell* 131: 901–914.
- Kolas NK, Chapman JR, Nakada S, Ylanko J, Chahwan R, et al. (2007) Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science* 318: 1637–1640.
- Maidland N, Bekker-Jensen S, Fastrup H, Melander F, Bartek J, et al. (2007) RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell* 131: 887–900.
- Wang B, Elledge SJ (2007) Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brca1/Brc36 complex in response to DNA damage. *Proc Natl Acad Sci U S A* 104: 20759–20763.
- Lou Z, Minter-Dykhouse K, Franco S, Gostissa M, Rivera MA, et al. (2006) MDC1 maintains genomic stability by participating in the amplification of ATM-dependent DNA damage signals. *Mol Cell* 21: 187–200.
- Bassing CH, Alt FW (2004) H2AX may function as an anchor to hold broken chromosomal DNA ends in close proximity. *Cell Cycle* 3: 149–153.
- Bassing CH, Chua KF, Sekiguchi J, Suh H, Whitlow SR, et al. (2002) Increased ionizing radiation sensitivity and genomic instability in the absence of histone H2AX. *Proc Natl Acad Sci U S A* 99: 8173–8178.
- Stucki M, Jackson SP (2006) gammaH2AX and MDC1: anchoring the DNA-damage-response machinery to broken chromosomes. *DNA Repair (Amst)* 5: 534–543.
- Yan Z, Kim YS, Jetten AM (2002) RAP80, a novel nuclear protein that interacts with the retinoid-related testis-associated receptor. *J Biol Chem* 277: 32379–32388.

16. Kim H, Huang J, Chen J (2007) CCDC98 is a BRCA1-BRCT domain-binding protein involved in the DNA damage response. *Nat Struct Mol Biol* 14: 710–715.
17. Sobhian B, Shao G, Lilli DR, Culhane AC, Moreau LA, et al. (2007) RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* 316: 1198–1202.
18. Wang B, Matsuoka S, Ballif BA, Zhang D, Smogorzewska A, et al. (2007) Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science* 316: 1194–1198.
19. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, et al. (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 266: 66–71.
20. Monteiro AN, August A, Hanafusa H (1996) Evidence for a transcriptional activation function of BRCA1 C-terminal region. *Proc Natl Acad Sci U S A* 93: 13595–13599.
21. Venkitaraman AR (1999) Breast cancer genes and DNA repair. *Science* 286: 1100–1102.
22. Foulkes WD (2004) BRCA1 functions as a breast stem cell regulator. *J Med Genet* 41: 1–5.
23. Narod SA, Foulkes WD (2004) BRCA1 and BRCA2: 1994 and beyond. *Nat Rev Cancer* 4: 665–676.
24. Haile DT, Parvin JD (1999) Activation of transcription in vitro by the BRCA1 carboxyl-terminal domain. *J Biol Chem* 274: 2113–2117.
25. Xu X, Weaver Z, Linke SP, Li C, Gotay J, et al. (1999) Centrosome amplification and a defective G2-M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform-deficient cells. *Mol Cell* 3: 389–395.
26. Scully R, Ganesan S, Vlasakova K, Chen J, Socolovsky M, et al. (1999) Genetic analysis of BRCA1 function in a defined tumor cell line. *Mol Cell* 4: 1093–1099.
27. Kim H, Chen J, Yu X (2007) Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. *Science* 316: 1202–1205.
28. Yan J, Menendez D, Yang XP, Resnick MA, Jetten AM (2009) A regulatory loop composed of RAP80-HDM2-p53 provides RAP80-enhanced p53 degradation by HDM2 in response to DNA damage. *J Biol Chem* 284: 19280–19289.
29. Yin Z, Menendez D, Resnick MA, French JE, Janardhan KS, et al. (2012) RAP80 Is Critical in Maintaining Genomic Stability and Suppressing Tumor Development. *Cancer Res* 72: 5080–5090.
30. Nikkila J, Coleman KA, Morrissey D, Pylkas K, Erkkö H, et al. (2009) Familial breast cancer screening reveals an alteration in the RAP80 UIM domain that impairs DNA damage response function. *Oncogene* 28: 1843–1852.
31. Yan J, Kim YS, Yang XP, Albers M, Koegl M, et al. (2007) Ubiquitin-interaction motifs of RAP80 are critical in its regulation of estrogen receptor alpha. *Nucleic Acids Res* 35: 1673–1686.
32. Coleman KA, Greenberg RA (2011) The BRCA1-RAP80 complex regulates DNA repair mechanism utilization by restricting end resection. *J Biol Chem* 286: 13669–13680.
33. Bian C, Wu R, Cho K, Yu X (2012) Loss of BRCA1-A complex function in RAP80 null tumor cells. *PLoS One* 7: e40406.
34. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22: 195–201.
35. Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T (2009) The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* 37: D387–392.
36. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7: 95–99.
37. Ramachandran GN, Sasisekharan V (1968) Conformation of polypeptides and proteins. *Adv Protein Chem* 23: 283–438.
38. Markin CJ, Xiao W, Spyropoulos L (2010) Mechanism for recognition of polyubiquitin chains: balancing affinity through interplay between multivalent binding and dynamics. *J Am Chem Soc* 132: 11247–11258.
39. Sato Y, Yoshikawa A, Yamashita M, Yamagata A, Fukai S (2009) Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by NZF domains of TAB2 and TAB3. *EMBO J* 28: 3903–3909.
40. Patil R, Das S, Stanley A, Yadav L, Sudhakar A, et al. (2010) Optimized hydrophobic interactions and hydrogen bonding at the target-ligand interface leads the pathways of drug-designing. *PLoS One* 5: e12029.
41. Sekiyama N, Jee J, Isogai S, Akagi K, Huang TH, et al. (2012) NMR analysis of Lys63-linked polyubiquitin recognition by the tandem ubiquitin-interacting motifs of Rap80. *J Biomol NMR* 52: 339–350.
42. Hurley JH, Lee S, Prag G (2006) Ubiquitin-binding domains. *Biochem J* 399: 361–372.
43. Noble JE, Bailey MJ (2009) Quantitation of protein. *Methods Enzymol* 463: 73–95.
44. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248–254.
45. Havlis J, Thomas H, Seib M, Shevchenko A (2003) Fast-response proteomics by accelerated in-gel digestion of proteins. *Anal Chem* 75: 1300–1306.
46. Jimenez CR, Huang L, Qiu Y, Burlingame AL (2001) In-gel digestion of proteins for MALDI-MS fingerprint mapping. *Curr Protoc Protein Sci Chapter* 16: Unit 16 14.
47. Whitmore L, Wallace BA (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* 89: 392–400.
48. Lobley A, Whitmore L, Wallace BA (2002) DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* 18: 211–212.
49. Whitmore L, Wallace BA (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res* 32: W668–673.
50. Stephens PJ, McKenna CE, McKenna MC, Nguyen HT, Devlin F (1981) Circular dichroism and magnetic circular dichroism of reduced molybdenum-iron protein of *Azotobacter vinelandii* nitrogenase. *Biochemistry* 20: 2857–2864.
51. Sreerama N, Venyaminov SY, Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis. *Anal Biochem* 287: 243–251.
52. Pace CN (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* 131: 266–280.
53. Walters J, Milam SL, Clark AC (2009) Practical approaches to protein folding and assembly: spectroscopic strategies in thermodynamics and kinetics. *Methods Enzymol* 455: 1–39.

This article was downloaded by: [vikrant .]

On: 22 October 2013, At: 22:05

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Biomolecular Structure and Dynamics

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/tbsd20>

Structural and functional characterization of the MERIT40 to understand its role in DNA repair

Vikrant^a, Pallavi Nakhwa^a, Dilip C. Badgajar^a, Rajan Kumar^a, Khushboo K.S. Rathore^a & Ashok K. Varma^a

^a Tata Memorial Centre, Advanced Centre for Treatment, Research and Education in Cancer, Kharghar, Navi Mumbai, Maharashtra, 410 210, India.

Published online: 15 Oct 2013.

To cite this article: Vikrant, Pallavi Nakhwa, Dilip C. Badgajar, Rajan Kumar, Khushboo K.S. Rathore & Ashok K. Varma , Journal of Biomolecular Structure and Dynamics (2013): Structural and functional characterization of the MERIT40 to understand its role in DNA repair, Journal of Biomolecular Structure and Dynamics, DOI: 10.1080/07391102.2013.843473

To link to this article: <http://dx.doi.org/10.1080/07391102.2013.843473>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Structural and functional characterization of the MERIT40 to understand its role in DNA repair

Vikrant, Pallavi Nakhwa, Dilip C. Badgajar, Rajan Kumar, Khushboo K.S. Rathore and Ashok K. Varma*

Tata Memorial Centre, Advanced Centre for Treatment, Research and Education in Cancer, Kharghar, Navi Mumbai, Maharashtra 410 210, India

Communicated by Ramaswamy H. Sarma

(Received 8 July 2013; accepted 8 September 2013)

MERIT40 (MEdiator of RAP80 Interaction and Targeting 40) is a novel associate of the BRCA1-complex and plays an essential role in DNA damage repair. It is the least characterized protein of BRCA1-complex and mainly responsible for maintaining the complex integrity. However, its structural and functional aspects of regulating the complex stability still remain elusive. Here, we carried out a comprehensive examination of *MERIT40* biophysical properties and identified its novel interacting partner which would help to understand its role in BRCA1-complex. The recombinant protein was purified by affinity chromatography and unfolding pathway was determined using spectroscopic and calorimetric methods. Molecular model was generated using combinatorial approaches of modeling, and monomer–monomer docking was carried out to identify dimeric interface. Disordered region of *MERIT40* was hatchet using trypsin and chymotrypsin to illustrate the existence of stable domain whose function was speculated through DALI search. Our findings suggest that *MERIT40* forms a dimer in a concentration-independent manner. Its central region shows remarkable stability towards the protease digestion and has structural similarity with vWA-like region, a domain mainly present in complement activation factors. *MERIT40* undergoes a three-state unfolding transition pathway with a dimeric intermediate. It interacts with adaptor molecule of BRCA1-complex, called ABRAXAS, thus help in extending the bridging interaction among various members which further stabilizes the whole complex. The results presented in this paper provide first-hand information on structural and folding behavior of *MERIT40*. These findings will help in elucidating the role of protein–protein interactions in stabilization of BRCA1-complex.

Keywords: genomic stability; *MERIT40*; BRCA1; RAP80; DNA damage repair (DDR)

1. Introduction

Genomic integrity is an essential prerequisite for the survival of living organisms. Failure to do so causes various genetic disorders and cancers (Shiotani & Zou, 2009). The number of mutations per gene per cell division of human cells has been estimated at approximately 2.0×10^{-7} (DeMars & Held, 1972; Zhang & Simon, 2005). Despite this high rate of mutagenesis, the mammalian genome exhibits extreme fidelity, which can be attributed due to DNA damage response (DDR) (Martin, Laroche, Suka, Grunstein, & Gasser, 1999; Mills, Sinclair, & Guarente, 1999). DDR is a highly regulated network of multiple proteins that detects damage and respond by generating a cascade of events, including cell cycle checkpoint activation, DNA repair mechanism and cell death (Bartek & Lukas, 2007; Zhou & Elledge,

2000). It involves a hierarchy of recruitment and assembly of many mediators and effectors at the DNA damage site(s) (Paull et al., 2000). It depends on a number of post-translational modifications, including phosphorylation, ubiquitination and acetylation, many of which are triggered by activation of the ataxia telangiectasia-mutated kinase (Harper & Elledge, 2007; Huen & Chen, 2008; Yan & Jetten, 2008; Zhou & Elledge, 2000). It subsequently endorses the assembly of UBC13, RNF8, and MDC1 on the phosphorylated histone H2AX (Chen et al., 1998). The outcome of this signaling event is the formation of polyubiquitin chains on the H2AX by Ubc13 and RNF8 (Huen et al., 2007; Kolas et al., 2007; Mailand et al., 2007; Wang & Elledge, 2007). These polyubiquitin chains are recognized by receptor-associated protein 80 (RAP80), which then assembles the

*Corresponding author. Email: avarma@actrec.gov.in

BRCA1 complex at the DNA damage site (Harper & Elledge, 2007). Assembly of the BRCA1 complex at the damage site is a vital process for homologous recombination repair (Hu et al., 2011).

The BRCA1 complex consist of six proteins RAP80, ABRAXAS, BRCA1, MERIT40 (Mediator of RAP80 Interaction and Targeting 40), BRCC36, and BRCC45, all of which are either directly or indirectly associated with each other (Feng, Huang, & Chen, 2009; Shao et al., 2009; Wang, Hurov, Hofmann, & Elledge, 2009). RAP80 contains N-terminal ubiquitin-interacting motifs (UIMs), a central Abraxas-Interacting Region (AIR) and a C-terminal zinc finger motif. The translocation of BRCA1 complex involves recognition of K63-linked polyubiquitin chains by UIMs of RAP80 (Huen et al., 2007; Mailand et al., 2007; Wang & Elledge, 2007; Yan & Jetten, 2008). ABRAXAS acts as an adapter molecule and its N-terminal associates with the AIR of RAP80, and phosphorylated p(S)XXF motif at the C-terminus associates with the BRCT domain of BRCA1 (Kim, Huang, & Chen, 2007; Wang & Elledge, 2007). BRCA1 interacts with numerous molecules and consequently plays a role in multiple pathways such as the DNA damage repair, cell cycle regulation, and transcription regulation (Haile & Parvin, 1999; Miki et al., 1994; Monteiro, August, & Hanafusa, 1996; Venkitaraman, 1999; Xu et al., 1999). It acts as a tumor suppressor and has been found to be associated with hereditary breast and ovarian cancer (Foulkes, 2004; Venkitaraman, 2002). Furthermore, BRCA1 BRCT deletion mutant showed perturbed behavior towards sub-nuclear co-localization with H2AX, thus highlighting its importance in DNA damage repair (Paull et al., 2000; Rogakou, Boon, Redon, & Bonner, 1999; Scully et al., 1997, 1999).

MERIT40 is a key component of the BRCA1 complex and is essential to maintain its integrity. Knockdown of MERIT40 significantly reduces the RAP80 and ABRAXAS proteins level, consequently affecting the integrity of BRCA1 complex (Feng et al., 2009; Shao et al., 2009). It has also been shown to considerably lower the stability of BRCA1 complex compared to any other member of this complex in MERIT40 down-regulated cells (Shao et al., 2009). To elucidate its role in BRCA1 complex, structural and functional properties were analyzed using interdisciplinary approach. Recombinant MERIT40 showed concentration-independent formation of dimer and followed a three-state transition pathway with dimeric intermediate. It was found to interact with ABRAXAS which would be responsible for stabilization of BRCA1 complex. It is the first report of direct interaction of MERIT40 with ABRAXAS, which act as a bridging molecule in the complex. Our findings provide insight into the role of uncharacterized MERIT40 in the BRCA1 complex and also established

its biophysical characteristics. These findings will also shed light on the mechanisms that alter the stability of this protein complex and consequently its DNA damage repair function.

2. Results and discussion

MERIT40 is a key module in the BRCA1 complex and plays a decisive role in its stabilization. It facilitates recruitment of BRCA1 to the site of DNA damage and favors homologous recombination repair. With an aim to understand the role of MERIT40 in BRCA1 complex integrity and hence in DNA repair pathway, we underwent an in-depth structural and functional characterization. Protein structure and function are highly related and forecasting of protein function often requires knowledge of structure and physiochemical properties (Hegyi & Gerstein, 1999; Wrzeszczynski & Rost, 2009). Hitherto, MERIT40 is the least characterized protein of BRCA1 complex and no structural information exists in the literature. It showed least homology with existing proteins (data not shown) in the database and hence could not be modeled using homology modeling. MERIT40 structure was modeled using Robetta server which employs combinatorial approaches for model generation (Kim, Chivian, & Baker, 2004) (Figure 1). The overall quality and reliability of the model was assessed with respect to energy and stereo chemical geometry (PROVE_Plot Z-score RMS: 30.936, overall quality factor (ERRAT): 84.936, Verify_3D (of 87.88% residues) score > .2). The Ramachandran plot showed that phi-psi torsion angles of 99.7% of the residues were present in the allowed region (Morris, MacArthur, Hutchinson, & Thornton, 1992) and none of the residue was in disallowed region (Figure 1 supplementary) (Ramachandran, Ramakrishnan, & Sasisekharan, 1963; Ramachandran & Sasisekharan,

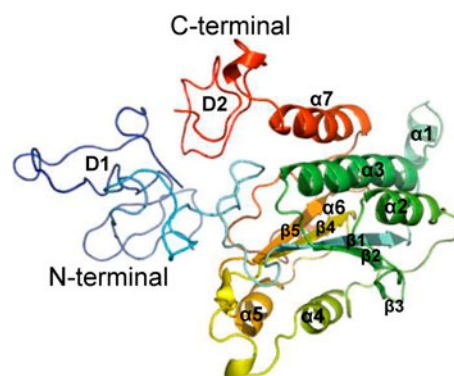


Figure 1. Modeled structure of MERIT40.

Note: Full length protein structure was modelled from Robetta server; Modeled structure showed 5- β sheets, 7- α helices, and disordered regions at N-terminal and C-terminal (α 1-7: alpha helices, β 1-5: beta sheets, D1-Diordered 1, D2-Diordered 2).

1968). The observed coordinates in Procheck analysis were found to be satisfactory (>80%) and Verify 3D profile revealed that the average score remained above .2 and never dropped below zero, which is significant. Furthermore, structural geometry was evaluated using Mol Probity web server (<http://molprobity.biochem.duke.edu/>) which showed all the parameters in required range (Table 1 supplementary). This indicates that the model quality was reliable and it could be used for further studies such as docking and domain organization.

2.1. Domain organization and plausible role of MERIT40 in biological processes

Intrinsic disorder is fairly a common phenomenon in many proteins, especially those related to cancer and testis (Rajagopalan, Mooney, Parekh, Getzenberg, & Kulkarni, 2011). Intrinsically disordered proteins lack rigid three-dimensional structures either along their entire length or in localized regions when set free of partners in solution (Uversky & Dunker, 2010). Due to the presence of many possible conformations and interactions, they play important biological roles in major cellular processes such as transcription, cell cycle regulation, signal transduction, and DNA repair (Uversky, Oldfield, & Dunker, 2008). Since MERIT40 is a member of cancer-related BRCA1 complex and one of member RAP80 has reported association with retinoid-related, testis-associated receptor, it would be more likely that MERIT40 might have disordered regions (Yan, Kim, & Jetten, 2002). To assess this possibility and mapping of disordered region, we analyzed the structure of MERIT40 using *in silico* and *in vitro* approaches. MERIT40 model suggested that N- and C-terminal region was having disorderedness while its middle domain exhibits a well-folded β/α conformation. Disordered regions of the N-terminal (~70 amino acids) and C-terminal (~30 amino acids) displayed a random structure pattern (Figure 1). The ordered region of the MERIT40 (~71–299 residues) encompasses four parallel, one anti-parallel β -sheet and seven α -helices. Our modeled structure provides the crucial information of presence of disordered region, hence it would be interesting to see whether the same scenario exists *in vitro* condition. Limited proteolysis is one of the widely used approach for identification and mapping of disordered regions of proteins in solution (Fontana, de Laureto, Spolaore, & Frare, 2012). In general, a compact globular domain of protein significantly resists the protease digestion, while disordered region undergoes rapid digestion due to more accessibility of protease sites (Fontana et al., 2012). To confirm the disorder of the N- and C-terminal regions and to identify the stable domain, we purified MERIT40 using affinity chromatography (Figure 2 supplementary) and subjected it to limited proteolysis assay. Trypsin and chymotrypsin were used as proteases to

locate the compact globular domain(s) in solution (Fontana et al., 1997). Both proteases rapidly digested MERIT40 (amino acids 1–329) to a proteolytically stable domain. MALDI-TOF/TOF analysis indicated that chymotrypsin digested the first 70 amino acids from the N-terminal region and a short stretch of 31 amino acid from the C-terminal (Figure 2(A)). However, trypsin digestion generated the same pattern at the N-terminal leaving the C-terminal intact (Figure 2(B)). Mass spectrometry identification of a stable fragment obtained after limited proteolysis suggested that the probable region was from 71–298 residues (Figure 3 supplementary). This stable fragment is likely defined in the range of MERIT40 vWFA-like domain (95–298), which showed homology with the von Willebrand factor (vWA) domain of the Rpn10 proteasome subunit (Hofmann & Bucher, 1998; Wang et al., 2009). Thus, the stable fragment of MERIT40 (71–298) may also possess proteasome-like activity, thereby raising the possibility of its involvement in the ubiquitination and deubiquitination processes along with RAP80 UIM.

Experimental determination of probable function of a novel protein is an extremely challenging task and most often the guidance is sought from *in silico* tools. Structure-based function prediction is one of the approaches for identifying the role of novel proteins in biological processes (Sael, Chitale, & Kihara, 2012). DALI server is an important tool that predicts the probable structural homologs of the novel protein (Holm & Rosenstrom, 2010). Structurally similar proteins most likely play similar role in a particular biological occurrence (Sael et al., 2012). MERIT40 is a recently annotated protein and hitherto its possible role in biological processes is not yet reported. DALI search (Holm & Rosenstrom, 2010) was carried out to find the structural homologs of MERIT40, and speculate its possible function. Three structural homologs Cobra Venom Factor (PDB ID: 3hrz), Complement Factor B (PDB ID: 2ok5) and Complement C3 Beta Chain (PDB ID: 2xwj) were observed (Figure 3). Cobra Venom Factor (CVF) is a complement-activating protein in Cobra venom which plays a crucial role in host defense (Vogel & Fritzinger, 2010). It resembles structurally and functionally with Complement C3 that plays an important role in complement activation and immunity (Table 1) (Vogel, Smith, & Muller-Eberhard, 1984). The β -chain of Complement C3 is homologous to the α -chain of CVF, and the C-terminal and N-terminal portions of Complement C3 α -chain are homologous to the α - and β -chains of CVF, respectively (Eggertsen, Lind, & Sjoquist, 1981; Fritzinger, Bredehorst, & Vogel, 1994). Complement fragment (C3b)/CVF interacts with Factor B (FB), a single-domain serine protease (SP) which circulates in plasma, and forms the proconvertase complex during complement activation. FB consists of a central vWA domain and a C-terminal

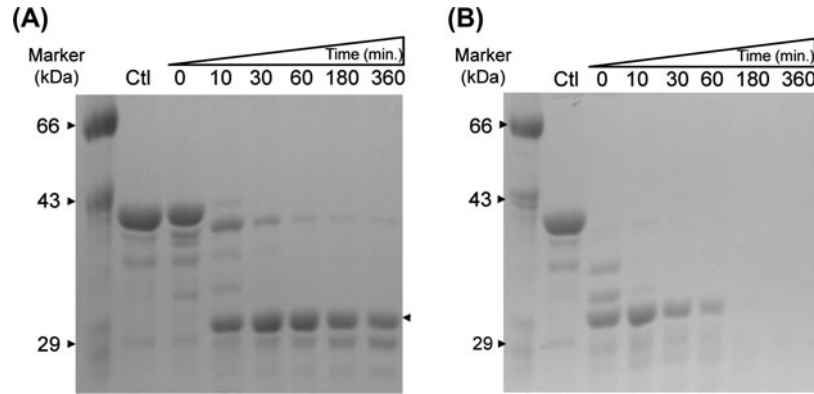


Figure 2. Limited proteolysis of MERIT40.

Note: Full length protein was subjected to digestion with Trypsin and Chymotrypsin, and digestion product was analyzed by 12% SDS-PAGE and MALDI-TOF (A) Time-dependent digestion of purified MERIT40 with chymotrypsin as a protease; (B) with trypsin (single arrow represents the main digestion product).

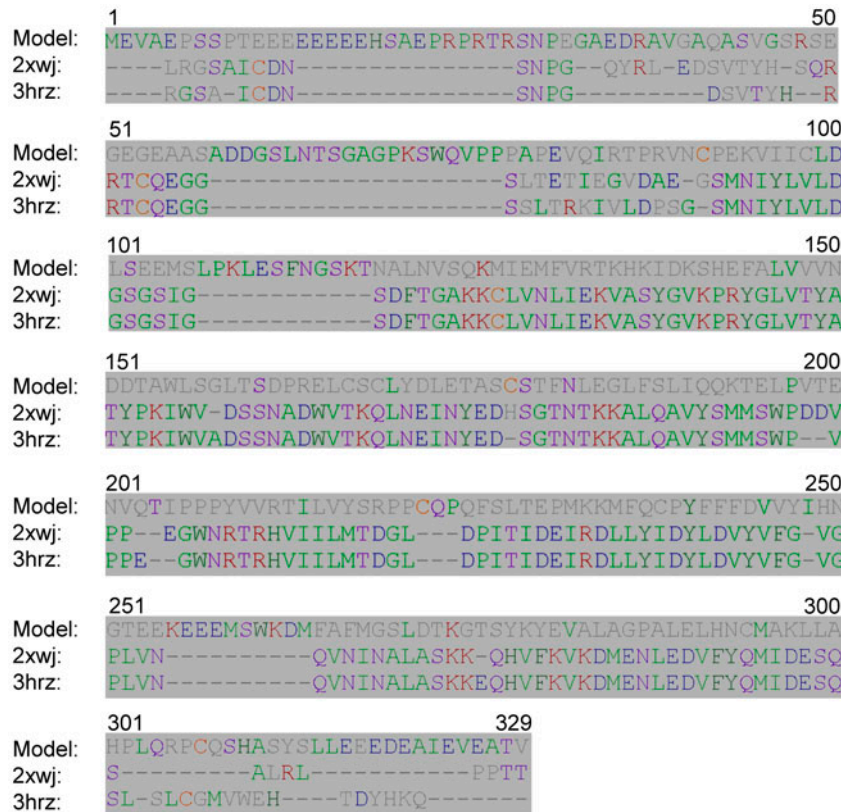


Figure 3. Multiple structural alignment of MERIT40 model with DALI searched templates. Protein model was submitted to the DALI server and best structurally aligned templates were selected on the basis of Z-score and % identity. The model showed highest structural alignment with Complement C3 beta (PDB: 2xwj) and Cvf (PDB: 3hrz).

trypsin-like SP domain that form the protease segment. Structural similarity of MERIT40 vWFA-like domain with a protease complex suggests its possible role in proteosome-like activity and involvement in ubiquitination processes.

Our molecular model properties were in line with the information obtained from limited proteolysis and DALI finding, thus strengthening the reliability for its use as a structure model in protein-protein or protein-ligand interaction studies. However, X-ray crystallographic analysis

Table 1. Structural comparison of MERIT40 with structurally similar protein using DALI.

Protein	PDB ID	Z-score	Sequence identity (%)	RMSD
Complement C3 beta chain	2XWJ	15.2	11	3.7
Cobra Venom Factor	3HRZ	14.9	10	4.1
Complement factor B	2OK5	13.7	10	4.2
Complement C2a fragment	2I6S	13.2	10	3.8
Complement C2	2ODP	13.0	10	3.9

is required to derive the exact structural coordinates of this protein.

2.2. Oligomeric behavior of MERIT40

Oligomerization is a common property found among disordered and exposed hydrophobic patches containing proteins (Zerovnik, 2011). Such protein may form several oligomeric species which may or may not have biological significance. To determine the magnitude of oligomerization and types of different oligomeric species, MERIT40 was characterized using size exclusion chromatography, native gel electrophoresis, glutaraldehyde crosslinking, mass spectrometry, dynamic light scattering (DLS), and *in silico* tools. Gel filtration profile showed the presence of prominent monomeric species (Figure 4 (A) supplementary), however, a second peak corresponding to dimer was also observed (Table 2). Furthermore, to confirm the presence of dimeric population and to delineate the multimeric behavior of MERIT40, native gel electrophoresis was performed. The presence of two distinct population on native gel corresponding to monomer and dimer provide evidence that MERIT40 may also exists as a dimer. In order to evaluate the proportion of dimeric population and to confirm whether their formation occur in a concentration-dependent manner, different concentrations of MERIT40 were analyzed on native gel. MERIT40 showed a concentration-independent formation of dimer and the dimeric proportion remained same irrespective of increasing concentration from 0.1 to 10 mg/ml (Figure 4(A), 4(B)). To support these finding and re-evaluate the dimer formation more precisely at different concentrations, increasing concentrations of protein (.1–10 mg/ml) were treated with glutaraldehyde and run on SDS-PAGE. It was observed that

the dimeric fractions remained same, irrespective of the changes in concentration, thus validating MERIT40 dimerization is concentration independent (Figure 4(C)). To further substantiate these findings, MALDI-TOF mass spectrometry was performed which revealed the presence of monomeric and dimeric population (Table 2, Table 2 supplementary, Figure 4(B), 4(C), 4(D) supplementary). It is often observed that molecules undergo dimerization due to intermolecular disulfide linkage (Wootton & Yoo, 2003). To investigate further whether dimer formation is occurring due to disulfide linkage, we treated one sample with β -ME and loaded on native page. There was no effect found on dimeric proportion after β -ME treatment, indicating dimer formation probably not occurred due to disulfide linkage (Figure 4(A)). Although, the exact magnitude of dimerization could not be determined, approximate values were determined using densitometric analysis of native gel. The monomeric and dimeric protein fractions were found to be approximately 60 and 40% of the total protein, respectively, across a range of concentrations (0.1–10 mg/ml) (Figure 4(B)). Similar proportion was also observed on gel filtration chromatography (Figure 4(A) supplementary). Thus it can be concluded, MERIT40 exists as a dimer but no relative concentration dependency exist for dimer formation over a range of concentration (0.1–10 mg/ml). In case of MERIT40, the plausible mechanism of concentration-independent dimer formation is allosteric regulation in which conformation of dimeric form is different from a simple association of monomeric–monomeric species (Corsepius & Lorimer, 2013). This does not allowed the dimeric formation beyond a fixed proportion which could be due to redirecting equilibrium towards monomer with increasing concentration.

Table 2. Molecular weight estimation of purified protein.

	Theoretical Mol. Wt. (kDa) ^a	V_e/V_o ^b	Experimental derived Mol. Wt. (kDa)	
			Gel filtration chromatography	Mass spectrometry (MALDI-TOF)
Monomer	36.60	1.48387	39.81	36.61
Dimer	73.12	1.59139	75.85	73.2

Note: V_e/V_o : Elution volume/Void volume ratio in gel filtration chromatography (superdex 200 16/60)

^aDetermined from ProtParam, Expasy.

^bDetermined from standard myoglobin, ovalbumin, albumin, IgG, Ferritin.

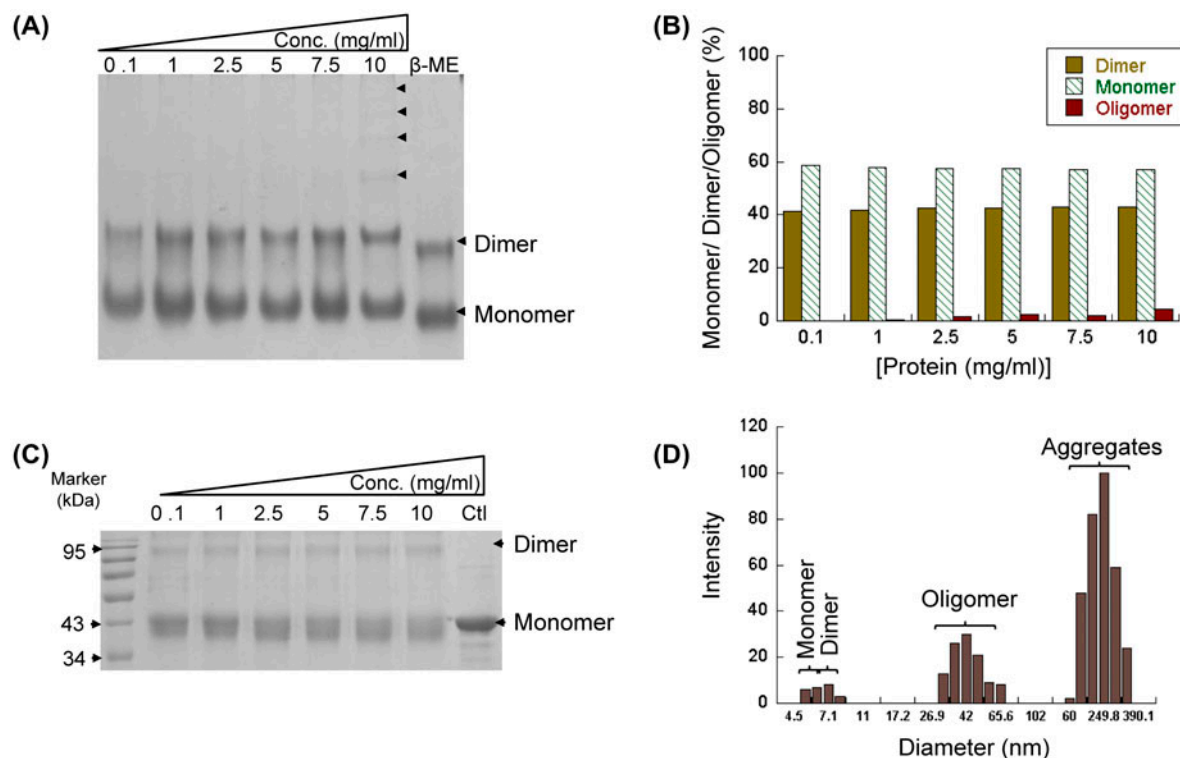


Figure 4. Oligomeric characterization of MERIT40 with native gel electrophoresis, and glutaraldehyde cross-linking followed by SDS-PAGE.

In each lane 20 μ g protein was loaded with different concentrations of protein. (A) Native gel (non denaturing) profile showing monomer, dimer and oligomer formation at different protein concentrations; (B) densitometry profile of the same showing percentage of different species. (C) Glutaraldehyde cross-linking reaction product on SDS-PAGE. (D) DLS of MERIT40 for molecular size measurement. Distribution pattern of different protein mer(s) of MERIT40 in DLS. About 1 mg/ml protein was scanned for molecular size measurement and identification of different mer(s) present in protein solution. Last lane in native gel is β -ME treated sample and second last lane showed the presence of nonspecific oligomers at high concentration, indicated by multiple arrowhead. Ctl: control as untreated with glutaraldehyde.

To determine the heterogeneity of purified MERIT40, we performed DLS, and by taking advantage of its sensitivity, we observed that the majority of samples were composed of three different clusters. The first cluster corresponds to the monomeric and dimeric populations, while the second and third clusters correspond to higher-order oligomers or aggregates (Figure 4(D)), (Figure 5 supplementary). The diameter of the monomers determined from modeled structure using *in silico* approach (Average 5–6 nm) and the effective hydrodynamic diameter determined by the DLS analysis (5–6 nm range) showed good concurrence with each other (Figures 1, 4 (D)). However, concentration-dependent DLS measurement could not be performed due to aggregation of the protein at higher concentrations.

Since we have established the existence of dimeric species along with the monomer in our *in vitro* study, we were keen to observe the same using *in silico* approach. In order to determine the *in silico* symmetrical dimer formation, we performed blind docking of mono-

mer molecules of MERIT40 after molecular dynamic simulation using Patchdock and Symmdock servers (Schneidman-Duhovny, Inbar, Nussinov, & Wolfson, 2005). To our surprise, only dimeric (diameter: 6–8 nm) form was observed in the docked structure (Figure 6 supplementary), further supporting our *in vitro* finding and reliability of our modeled structure. The dimeric structure was stabilized by the prime involvement of Gly115, Gln125, Lys126 and Arg133 residues at the interface. To validate these findings, we performed *in silico* substitution of these residues with bulky and charge repulsive (Gln115, Ile125, Glu126, and Glu133, respectively) and Ala residues, followed by docking of monomers. Substitution of either residue in monomers disfavored the *in silico* dimerization of MERIT40 (docking scores (Zhang, Vasmatzis, Cornette, & DeLisi, 1997) of various substitution: wild type 14,640, Gly115Gln 13,430, Gly115Ala 13,408, Gln125Ile 14,620, Gln125Ala 14,614, Lys126Glu 14,596, Lys126Ala 14,308, Arg133Glu 14,110, Arg133Ala

13,392) (Figure 6 supplementary). The energy of dimeric complex due to Gly115Ala (-19.56 kJ/mol), Gln125Ala (-6.78 kJ/mol), Lys126Ala ($.84$ kJ/mol), and Arg133Ala (-19.43 kJ/mol) was found higher as compared to wild type (-47.92 kJ/mol), indicating relative less stability of dimers due to substitution with alanine residue. Due to the limitations of *in silico* methods in the absence of X-ray structure, special modes of dimerization that result in formation of unexpectedly low- or high-diameter molecule(s) could not be explored in detail. Moreover, asymmetric dimerization model and their interacting residues could not be mapped for the same reason.

2.3. Structural characterization of MERIT40

To characterize the secondary structural constituents of MERIT40, purified MERIT40 was used to acquire far-UV circular dichroism (CD) spectrum (Figure 5(A)). The CD spectrum was analyzed using DichroWeb server, in which the percentage of α -helices and β -sheets was found to be 20 and 16.6%, respectively. Furthermore, secondary structural components of modeled structure were determined using the PROSS server (<http://roselab.jhu.edu/utis/runpross.html>), which showed the percentage of α -helices and β -sheets to be 23 and 15%, respectively. The results from the two independently performed methods were in good agreement. These observations strongly support the faithfulness of the secondary structure constituents of our modeled structure.

Tryptophan and tyrosine are the most frequently used intrinsic fluorophores to study the microenvironmental changes induced in the protein due to external agents like temperature and chaotrophs (Muniz et al., 2011). These aromatic residues are generally buried inside the protein and are extremely sensitive to modification

around their microenvironment. To study the three-dimensional structure of MERIT40, Trp and Tyr microenvironment were monitored using fluorescence spectroscopy. An emission maxima at λ 338 nm was obtained for the native protein, while a red shift at λ 346 nm was observed in case of unfolded protein (Figure 5(B)). This indicates that most of the tryptophan and tyrosine residues of the native protein are probably buried inside the hydrophobic core. To obtain the stability parameters and folding pattern, thermal denaturation was performed using CD and fluorescence spectroscopic probes. In Far-UV CD data, maximum change in ellipticity at different temperature was observed at 208 nm, which was therefore used to monitor the changes in protein secondary structure at varied temperatures (15 – 80 °C) (Figure 6 (A)). The T_m (melting temperature) of 48 ± 4.1 °C was obtained by data fitting, and thermodynamic parameters were calculated ($\Delta G_{H_2O}^\circ$ 4.8 ± 1.0 kcal/mol, ΔH 144 ± 2.5 kcal/mol) and listed in (Table 3) (Figure 7(A) supplementary). To further validate these finding, protein was unfolded in a temperature gradient of 15 – 80 °C in fluorescence spectrophotometer and Trp, Tyr emission maxima was monitored at λ 280 nm. The changes in emission maxima of protein at various temperature were used to determine its folding pathway and calculation of thermodynamic parameters. Fluorescence spectroscopy showed a T_m value of 52 ± 3.2 °C and energy parameters as listed in ($\Delta G_{H_2O}^\circ$ 4.6 ± 1.5 kcal/mol, ΔH 142 ± 2.3 kcal/mol) (Table 3) (Figure 6(A)), (Figure 7(B) supplementary). These parameters indicate that MERIT40 possessed moderate stability, and its enthalpy and free energy values were similar to globular protein (Pace, Shirley, McNutt, & Gajiwala, 1996). Both methods demonstrate that the protein most likely unfolds via the formation of

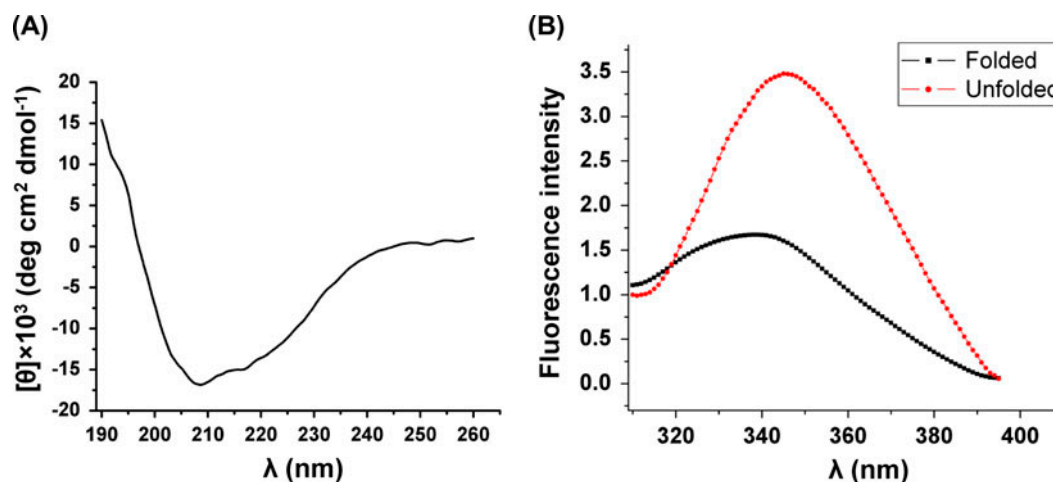


Figure 5. Far-UV CD spectra, and urea unfolding using fluorescence, of MERIT40.

(A) Molar ellipticity was calculated and plotted against the wavelength. (B) Fluorescence emission maxima of native and unfolded protein. Native protein showed emission maxima of 338 nm while unfolded protein undergoes a red shift up to 346 nm.

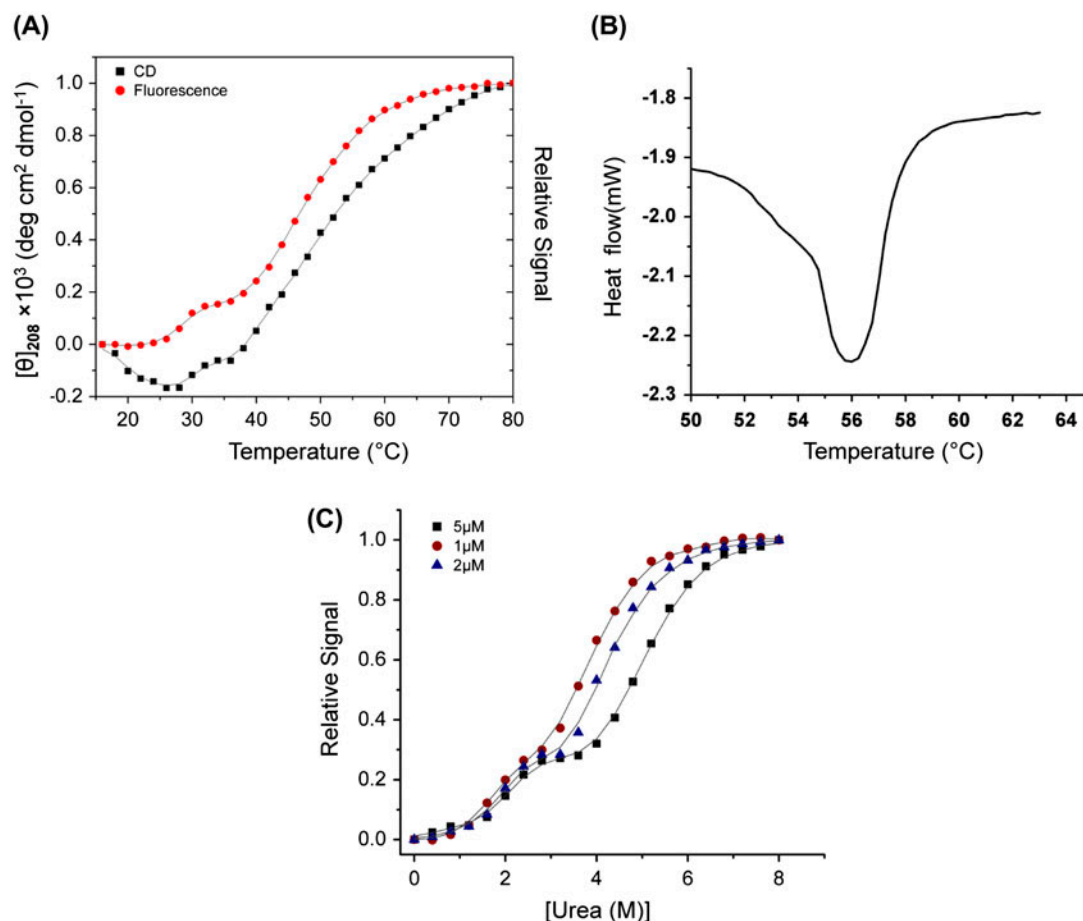


Figure 6. Denaturation profile of MERIT40 using CD, fluorescence spectroscopy and DSC.

(A) Overlay of protein fractions unfolded at different temperature in CD and fluorescence spectroscopy, Far-UV CD spectrum were recorded at 208 nm wavelength showed maximum change in ellipticity. Trp and Tyr fluorescence emission maxima were monitored at excitation wavelength of 280 nm at different temperature. Relative signal represents fraction unfolded for fluorescence. (B) DSC thermogram. A total of 1 mg protein in solution form was heated with 1 °C /min rate and heat change was recorded. (C) Chemical unfolding at different urea concentration. The unfolding curve represents two transitions. Concentration dependency of second transition indicates the formation of dimeric intermediate during unfolding process. Relative signal represent fraction unfolded.

Table 3. Thermal parameters of protein unfolding.

Methods	Unfolding process	T_m (°C)	C_m (M)	$\Delta G_{H_2O}^c$ (kcal/mol)	ΔH (kcal/mol)
CD	Thermal	48 ± 4.1	—	4.8 ± 1.0	144 ± 2.5
Fluorescence	Thermal	52 ± 3.2	—	4.6 ± 1.5	142 ± 2.3
Fluorescence	Chemical	—	$3.4 \pm .45$	3.8 ± 1.0	155 ± 1.9
DSC	Thermal	$55 \pm .9$	—	—	150 ± 2.0

an intermediate species as indicated by abrupt increase of signal in unfolding curve. In order to validate these findings, we assessed the unfolding profile of native protein using differential scanning calorimetry (DSC). 1 mg/ml protein was allowed to unfold in a temperature gradient of 15–80 °C in calorimetry and unfolding transition was obtained. Peak maxima during unfolding transition were considered as temperature of melting (T_m).

DSC thermogram suggested overlapping of two transitions, which could be corresponding to an intermediate species formed during the unfolding process (Figure 6 (B)). The melting temperature ($55 \pm .9$ °C) and enthalpy parameters (ΔH 150 ± 2.0 kcal/mol) were derived using Calisto software. The T_m value determined by DSC and fluorescence were similar and unfolding profile indicates that protein form an intermediate species (Table 3).

MERIT40 showed *in vitro* and *in silico* dimer formation, hence, it could be possible that it may form dimeric intermediate during unfolding process. To test this hypothesis, a concentration-dependent urea denaturation study was performed to find the type(s) of intermediate species formed during unfolding process. Protein was incubated in different urea concentration (0–8 M) till equilibrium was achieved and emission maximum using fluorescence were recorded. Change in average emission wavelength at different urea concentrations was plotted to calculate unfolded fraction. A biphasic unfolding curve with two transitions was observed, and protein concentration dependency in the second transition indicates that dissociation of MERIT40 subunits occurs after formation of a dimeric intermediate (Walters, Milam, & Clark, 2009) (Figure 6(C)). The chemical denaturation achieved using urea as the unfolding agent substantiates the thermal denaturation findings (Figure 6(A), 6(B), 6(C), (Figure 7(C) supplementary).

2.4. Pathogenicity of MERIT40 variants

Mutations or variants are known to exist almost in every member of BRCA1 complex but their pathogenic implications have not been characterized in details (Nikkila et al., 2009; Solyom, Patterson-Fortin, Pylkas, Greenberg, & Winqvist, 2010). One of the most familiar variant of BRCA1, M1775R has been reported to have role in cancer susceptibility (Williams & Glover, 2003). Disease implication caused by these variant arises either due to alteration of entire domain structure or localized alteration in weak intermolecular or intramolecular interactions (Williams & Glover, 2003). To characterize the pathogenicity of sequence variants identified in the patient cohort and to evaluate the importance of weak intramolecular interactions in disease susceptibility, we analyzed seven recently reported sequence variants of MERIT40 (c.342C>T, c.344+41A>T, c.393C>T, c.787 – 6C>T, c.821A>G, c.837G>A, c.87G>A). It

is reported that none of the above variant has any role in the predisposing individual to disease (Solyom et al., 2010). All the variants are silent in nature except of c.821A>G in exon 9, which lead to the substitution of Arg to Lys at 274 position. However, change in Lys to Arg does not bring any pathological implication as observed in cohort of patient in previous report (Solyom et al., 2010). We intended to determine the cause to neutrality of K274R variant. To evaluate this, we analyzed the molecular environment around the Lys274 residue and changes as a result of substitution with Arg. Structural analysis of our modeled structure unraveled that Lys274 is buried in the hydrophobic core of the protein which probably restrict its binding with other protein (Figure 7(A)). Lys274 forms hydrophobic interactions with Pro77, Gly75 and Thr273. We found no changes in the structural conformation and hydrophobic environment around Lys274 residue of MERIT40 as a result of substitution with Arg, and majority of interactions with adjacent residues remained unaltered (Figure 7(B)). Since, intramolecular interactions remained unaffected; it could be assumed that MERIT 40 K274R should be clinically insignificant. This assumption is in agreement with the findings of a previous report which state that mutations predisposing to breast cancer are either very rare or absent in the coding region of *MERIT40* (Solyom et al., 2010). This evidence further augments the validity of our modeled structure since it showed a good agreement with reported data.

2.5. MERIT40 stabilize BRCA1 complex through direct interaction with ABRAXAS

Protein–protein interaction in BRCA1 complex plays very significant role and alteration in interaction profile leads to destabilization of entire complex (Shao et al., 2009). MERIT40 is an essential component of BRCA1 complex; however, how MERIT40 stabilize this complex still elusive. Hitherto, direct binding partner of

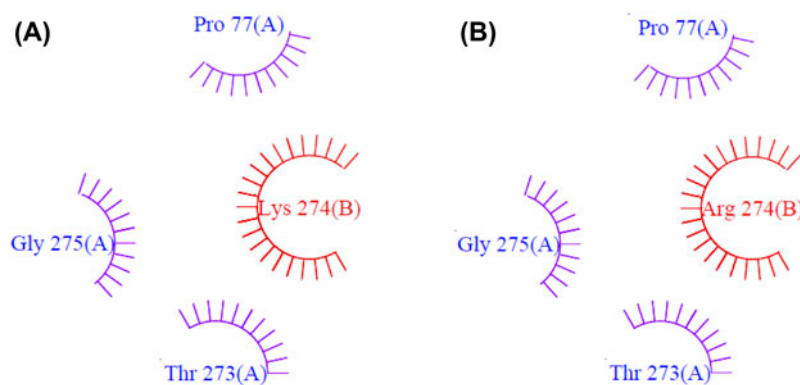


Figure 7. Mutational analysis of MERIT40. Note: Ligplot showed no change in hydrophobic environment in wild type lys274 and variant lys274arg. (A) ligplot of Lys274. (B) ligplot of Arg274. The spokes indicate residues involved in non-bonding interactions.

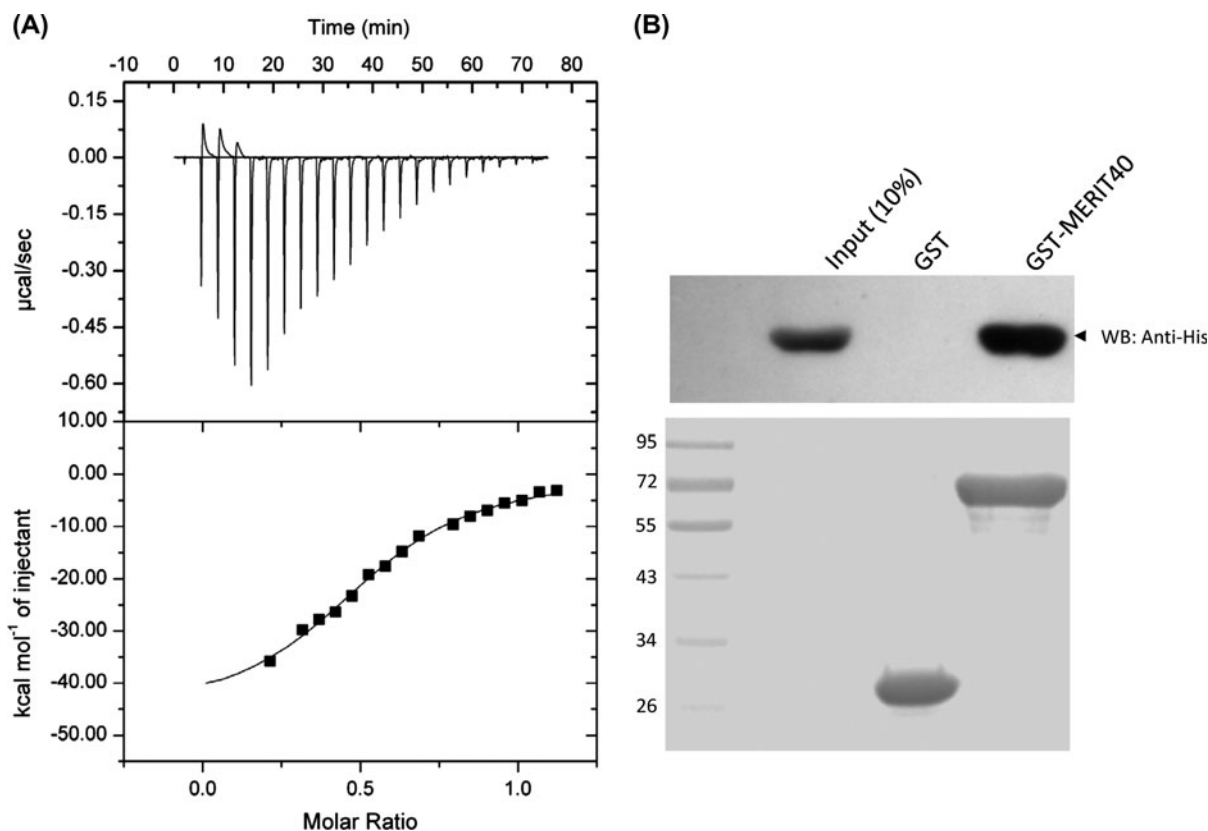


Figure 8. Identification of binding partner of MERIT40.

Note: Interaction analysis of MERIT40 with ABRAXAS using (A) Isothermal Titration Calorimetry and (B) GST pull down assay. The lower panel of figure B shown the GST and GST-MERIT40 loaded on SDS-PAGE followed by transferred to PVDF membrane and ponceau staining. Upper panel showed MERIT40 bound protein (His tag-ABRAXAS) detected through western blotting. GST was taken as control.

MERIT40 is not reported through which it could be stabilizing the complex. To evaluate this, we performed interaction analysis of MERIT40 with ABRAXAS (6-409) using Isothermal Titration Calorimetry. MERIT40 was titrated with ABRAXAS (10-fold molar ratio) at constant temperature and heat changes were determined. MERIT40 showed direct interaction with ABRAXAS as indicated with significant heat change during reaction. The binding affinity between MERIT40 and ABRAXAS was found ($K = 3.90 \pm .76 \text{ E}^5 \text{ M}^{-1}$) which suggests a modest interaction between them (Figure 8(A)). In order to validate these finding, we performed pull down assay, which confirm the interaction between MERIT40 and ABRAXAS (Figure 8(B)). It is reported that ABRAXAS form phosphodependent interaction with BRCA1 BRCT and phospho-independent interaction with RAP80, thus acting as a bridging molecule in BRCA1 complex (Wang et al., 2007). Binding interaction between MERIT40 and ABRAXAS probably helps in extension of bridging interaction among various members of BRCA1 complex and thereby maintaining its integrity.

To best of our knowledge, this is the first report of the characterization and identification of binding partner of this protein. MERIT40 can be classified as an intrinsically disordered protein due to the presence of N-and C-terminal disorder region. However, its middle region showed well-defined compact structure which indicates a good possibility of crystallization of middle region for X-ray diffraction study. Structural homologous of MERIT40 suggests its plausible role in complement activation pathway, nevertheless, definite function determination needs further experimental evidences. It exhibits concentration-independent dimerization and unfolds through a dimeric intermediate. It perhaps set up a interaction network in BRCA1 complex which is being utilized for stabilization of ABRAXAS since knockdown of MERIT40 significantly reduces the ABRAXAS and RAP80 levels (Feng et al., 2009; Shao et al., 2009). Stabilization of ABRAXAS probably further helps in maintaining the integrity of BRCA1-complex. This study will provide insights into the diverse interactions involved among various members which are essential for DNA repair function of BRCA1 complex. At last but not

the least, MERIT40 could be a multifunction molecule having role in DNA damage repair and complement activation.

3. Materials and methods

All the chemicals were purchased from Sigma-Aldrich, unless otherwise specified. Chemicals were analytical grade or molecular biology grade with more than 99.99% purity. Double distilled, .44 μ M filtered Milli-Q water (Millipore, USA) was used for making buffers. Protein and buffer solutions were either centrifuged or .44 μ M (Millipore, USA) filtered and degassed prior to use in any experiment depending upon the need.

3.1. Cloning, protein expression, and purification

MERIT 40 (C19orf62) gene (Junjie Chen, Yale University School of Medicine, New Haven) was PCR amplified (Thermocycler, Biorad) with primers having TEV protease site incorporated (E-N-L-Y-F-Q/S) in it along with BamH1 and EcoR1 restriction sites. PCR-amplified product and bacterial expression pGEX-kT vector (kind gift from Dr John Ladas laboratory, USA) were digested and ligated. Positive clones were subjected to restriction digestion and further confirmed by DNA sequencing. Similar cloning procedure was performed with abraxas (16-1227) (Junjie Chen, Yale University School of Medicine, New Haven) and the digested PCR product having Nco1 and Xho1 sites was ligated in vector pET28a(+) (Novagen). Plasmid DNA was transformed into bacterial expression strain *E.coli* Rosetta 2 (DE3) cells (Novagen) and single colony was inoculated in LB medium (Himedia, India) to obtain pre-innoculum culture. Bulk culture was grown at 37 °C till OD₆₀₀ reached between .6 and .8, and thereafter induced with .4 mM IPTG (Himedia, India). Cells were induced overnight at 24 °C and harvested by centrifugation. Pellet was resuspended in 10 mM HEPES buffer containing 300 mM NaCl, 10 mM β -ME, .1 mM EDTA, and 5% glycerol pH 7.5. Cells were disrupted by sonication (Branson sonifier) and supernatant was separated by centrifugation. Soluble protein was passed through pre-equilibrated glutathione resin/Ni-NTA (Novagen) with same buffer. To remove the tag, his tag TEV protease was added for 3hrs to perform on-beads cleavage. Protein was further purified by gel filtration chromatography (superdex 200, GE) and analyzed by SDS-PAGE (coomassie staining) for purity.

3.2. Protein characterization

Protein concentration was determined by NanoDrop 1000 (Thermo scientific, USA) and Bradford method. Absorbance of protein sample(s) was recorded after blank correction and concentration was determined by considering extinction coefficient (Grimsley & Pace,

2004, Chapter 3). For Bradford protein estimation, BSA was taken as standard and absorbance of different concentration of BSA solution was measured using spectrophotometer (Shimadzu, USA) according to manufacturer instruction (Sigma-Aldrich). The experiment was performed in triplicates and concentration of protein of interest was determined by interpolation of standard curve (Bradford, 1976; Noble & Bailey, 2009).

3.3. Molecular modeling, DALI search, and mutational analysis

Protein structure was modeled using Robetta server (<http://rosetta.bakerlab.org/>) (Kim et al., 2004). Selected models based on overall stereo chemical quality were validated using Ramachandran plot and protein structure validation server "SAVES" (Metaserver for analyzing and validating protein structures, <http://nihserver.mbi.ucla.edu/SAVES/>). SAVES is an online structure validation server that involves primarily five programs, Procheck, What_check, Errat, Verify_3D, and Prove. Protein was simulated for 5 ns using Desmond (Schrodinger). Monomer-Monomer geometrical docking was carried out using Patchdock & SymmDock server (<http://bio-info3d.cs.tau.ac.il/SymmDock/>) (Schneidman-Duhovny et al., 2005). Dimer interface residue(s) was identified with the help of PDBsum. Each residue was further mutated with bulky, charge repulsive, and alanine residue followed by MD simulation. Monomer-monomer docking was carried out and best docked structures were evaluated and selected on the basis of docking score and atomic contact energy. Monomer and dimer radius were measured from the docked structure using PyMol software. Structural homologs of MERIT40 were searched using DALI server (Holm & Rosenstrom, 2010).

For Mutational studies, mutation was incorporated in modeled structure corresponding to a particular variant and structural analysis was performed using Ligplot (<http://www.ebi.ac.uk/thornton-srv/software/LIGPLOT/>) to observe changes in molecular environment.

3.4. Limited proteolysis

Trypsin and chymotrypsin were mixed with purified protein (1 mg/ml) separately with their final concentration of 40 and 10 μ g/ μ l, respectively, and untreated protein was taken as control. Reaction mixtures were incubated at 37 °C (trypsin) and 25 °C (chymotrypsin) for different time period, 0, 10, 30, 60, 180, and 360 min. Reactions were terminated by adding 2 μ l of 200 mM PMSF (sigma-Aldrich). Samples were heated with aemmli buffer and analyzed on SDS-PAGE (coomassie staining) (Vikrant et al., 2013). Band corresponding to 29 kDa was considered as most stable fragment and was further subjected to trypsin digestion followed by mass spectrometry. Domain

of interest was identified by Mascot analysis with Biotoool software (Bruker Daltonics, USA). This experiment was repeated thrice to eliminate any error in domain prediction (Havlis, Thomas, Sebela, & Shevchenko, 2003; Jimenez, Huang, Qiu, & Burlingame, 2001, Chapter 16).

3.5. Glutaraldehyde cross linking and native gel electrophoresis

Purified protein was concentrated and various dilutions were prepared (.1, 1, 2.5, 5, 7.5, and 10 mg/ml). Reaction mixtures with 20 µg of protein in 10 mM HEPES buffer (pH 7.5) were treated with freshly prepared solution of glutaraldehyde (final concentration .1%) for 2 min at, 37 °C (Feeney, 1987). This reaction was terminated by addition of 5 µl of 2 M Tris-HCl, pH 8.0. Cross-linked product was mixed with equal amount of Laemmli buffer and analyzed on 12% SDS-PAGE.

For native gel same set of protein dilution (as above) without adding glutaraldehyde and tris was used for running on 7.5% basic-native gel (non-denaturing). Protein samples were mixed with laemmli buffer without SDS and β-ME and run at low voltage in cold room. In order to see disulfide bond-specific multimerization, one sample was incubated with β-ME before loading on gel (coomassie staining). These experiments were done thrice with different preparation.

3.6. Circular dichroism

Far-UV CD spectroscopy and thermal denaturation was performed using CD polarimeter (Jasco J-810, Japan). About 10 µM protein (in 2.5 mM Hepes pH 7.5, 50 mM NaCl) spectrum were taken in a scanning range of 190–260 nm at 20 °C. Blank correction was done to eliminate the effect of buffer components, average data of three independent scans were considered. Mean residual ellipticity was determined and data was analyzed using DichroWeb server (<http://dichroweb.cryst.bbk.ac.uk>) (Lobley, Whitmore, & Wallace, 2002; Sreerama, Venyaminov, & Woody, 2000; Stephens, McKenna, McKenna, Nguyen, & Devlin, 1981; Whitmore & Wallace, 2004, 2008). For thermal denaturation, purified protein (10 µM) was allowed to unfold in a temperature range of 15–80 °C in scanning wavelength of 208–240 nm. The experiment was repeated three times independently and average data were considered.

3.7. Fluorescence spectroscopy

The tryptophan and tyrosine positions were monitored using fluorescence spectrophotometer (Horiba, USA) at excitation wavelength of 280 nm and temperature 20 °C. Fluorescence emission scans were recorded from 310–400 nm wavelength. For thermal denaturation, 2 µM protein was unfolded in a temperature range of 15–80 °C and emission maximum were obtained. Averaged

normalized data of three independent experiments were considered after blank correction. Data were analyzed by fitting in three state transition model and thermal parameters were calculated.

For chemical denaturation, 10 M urea (Sigma-Aldrich) stock was prepared and protein dilutions were prepared as describe previously (Walters et al., 2009). Emission scans were collected (as described above) and was blank subtracted in ordered to increase signal-to-noise ratio. Thermodynamic parameters were obtained by curve fitting according in three state transition model.

3.8. Thermal and chemical denaturation data analysis and curve fitting

Data fitting (CD, Fluorescence) was done using Origin (OriginLab, Northampton, MA) for three state transition model and thermal parameters were calculated by considering native (*N*) to intermediate (*I*) and intermediate to unfolded (*U*) transition (Pace & Shaw, 2000)



Here, K_1 and K_2 represent the equilibrium constants for the two reactions.

In order to visualize different spectroscopic signals on a same scale, data were normalized using following equation.

$$f_F = (Y_X - Y_U) \div (Y_N - Y_U) \quad (2)$$

In this case, Y_x and Y_U are the signal being normalized and the signal of the unfolded protein respectively, and Y_N is the signal of the native protein. f_F represents the fraction folded at a particular temperature or denaturant concentration. A plot of the normalized signal (fraction folded, f_F) vs. the denaturant concentration or temperature, generate the unfolding curve for each of the spectroscopic probes.

The equilibrium constant (K) can be derived from f_F using following equations.

$$K = f_U / f_F = f_U / (1 - f_U) \quad (3)$$

where f_U represents the fraction unfolded at particular temperature of the denaturant concentration.

Following equations were employed for calculation of free energy (ΔG) and enthalpy (ΔH) of unfolding reaction from K .

$$\Delta G = -RT \ln K \quad (4)$$

$$d(\ln K)/d(1/T) = -\Delta H/R \quad (5)$$

where R is the gas constant ($1.985 \text{ calK}^{-1} \text{ mol}^{-1}$) and T is the temperature (Kelvin)

$$\Delta G = \Delta G^{\text{H}_2\text{O}} - m[\text{urea}] \quad (6)$$

where $\Delta G^{\text{H}_2\text{O}}$ is defined as an estimate of the conformational stability of a protein in the absence of denaturant (i.e. Intercept of linear extrapolation of ΔG value), m is the slope of the plot ΔG vs. urea and it measures the dependency of ΔG on urea concentration (Pace, 1986).

$$Y = \{(\gamma_F + m_F T) + (\gamma_U + m_U T) \exp(\Delta H/R(1/T_m - 1/T))\} / (1 + \exp(\Delta H/R(1/T_m - 1/T))) \quad (7)$$

$$Y = \{(\gamma_F + m_F T + c_F T^2) + (\gamma_U + m_U T + s_U T^2) \times \exp(\Delta H/R(1/T_m - 1/T))\} / (1 + \exp(\Delta H/R(1/T_m - 1/T))) \quad (8)$$

where Y represents the optical/fluorescence signal at particular temperature T and γ_F , γ_U , and c_F are the temperature independent constants. m_F and m_U are the slopes of pre and post-transition baselines. ΔH is the enthalpy of reaction and T_m is the temperature of melting. Equations (7) and (8) assumed a linear and parabolic dependence of pre- and post-transitions baselines on temperature, respectively. These were referred as linear and parabolic model of pre- and post-transitions baselines (Yadav & Ahmad, 2000).

3.9. Differential scanning calorimetry

Thermal unfolding of protein was done using DSC (Setaram μ DSC3 evo, USA) in the 15–80 °C temperature range. Protein and buffers were degassed prior undertaking the scan in order to avoid air bubble interference. A total of 1 mg protein in solution form was run with a temperature increment rate of 1 °C/min. The experiment was replicated thrice independently and blank correction was done. Observed values of furnace temperature vs. heat flow diagram were fitted locally by “CALISTO” software (AKTS) and enthalpy value was calculated. The thermodynamic reversibility of the transition was determined by heating the sample to a temperature that was little over the transition maximum, then cooling immediately and reheating. Thermal denaturation transitions were found irreversible as indicated by the lack of transition in the second run.

3.10. Dynamic light scattering

Molecular size measurement was done using Brookhaven 90plus particle size analyzer (Brookhaven, NY). Protein and buffer solution were filtered ($.44 \mu\text{m}$) and degassed

prior to measurement. About 1 mg/ml protein was scanned at an interval of 5 min for 15 min and effective diameter of each measurement was considered. The experiment was repeated in three independent sets.

3.11. Isothermal titration calorimetry

MERIT40 was taken as titrant (30 μM) and ABRAXAS (290 μM) as titre onto a Microcal ITC200. The reaction was allowed to proceed at 25 °C with a stirring speed 1000 rpm. A total of 23 injections were (each 2 μl) with a equilibration time of 4 s. Blank correction was done and affinity was determined using one site binding model in origin software. The experiment was repeated thrice. All the required reagents were degassed prior to setting the reaction.

3.12. GST pull down assay

Bacterial pellet of GST and GST-MERIT40 were resuspended in 10 mM HEPES buffer containing 300 mM NaCl, 10 mM β -ME, .1 mM EDTA and 5% glycerol pH 7.5, and sonicated. Soluble fusion protein(s) bound on glutathione resin (.5 mg/ml) was used to capture prey (His-tag-ABRAXAS)(50 μg). Resin was washed with same buffer and loaded on SDS-PAGE. Complex was transferred to PVDF membrane (Millipore, USA) and was probed with anti-His antibody (Abcam, USA). The experiment was repeated twice by taking GST as control.

Funding

Funding for this study was supported by DBT (No. BT/PR10765/BRB/664/2008 and BT/PR12565/BID/07/303/2009), Seed in Air grant from TMC. We thank DBT-BTIS facility at ACTREC for providing the necessary software for this study. Mr Vikrant thanks CSIR for fellowship (09/513 (0072)/2008-EMR-I). We thank Dr Hosur and Dr Lata (BARC) for providing DSC facility and data analysis.

Supplementary material

The supplementary material for this paper is available online at <http://dx.doi.org/10.1080/07391102.2013.843473>.

References

- Bartek, J., & Lukas, J. (2007). DNA damage checkpoints: From initiation to recovery or adaptation. *Current Opinion in Cell Biology*, 19, 238–245.
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72, 248–254.
- Chen, J., Silver, D. P., Walpita, D., Cantor, S. B., Gazdar, A. F., Tomlinson, G., ... Scully, R. (1998). Stable interaction between the products of the BRCA1 and BRCA2 tumor suppressor genes in mitotic and meiotic cells. *Molecular Cell*, 2, 317–328.

- Corsepius, N. C., & Lorimer, G. H. (2013). Measuring how much work the chaperone GroEL can do. *Proceedings of the National Academy of Sciences of the United States of America*, 110, E2451–2459.
- DeMars, R., & Held, K. R. (1972). The spontaneous azaguanine-resistant mutants of diploid human fibroblasts. *Human-genetik*, 16, 87–110.
- Eggertsen, G., Lind, P., & Sjoquist, J. (1981). Molecular characterization of the complement activating protein in the venom of the Indian cobra (*Naja N. siamensis*). *Molecular Immunology*, 18, 125–133.
- Feeney, R. E. (1987). Chemical modification of proteins: Comments and perspectives. *International Journal of Peptide and Protein Research*, 29, 145–161.
- Feng, L., Huang, J., & Chen, J. (2009). MERIT40 facilitates BRCA1 localization and DNA damage repair. *Genes & Development*, 23, 719–728.
- Fontana, A., de Laureto, P. P., Spolaore, B., & Frare, E. (2012). Identifying disordered regions in proteins by limited proteolysis. *Methods in Molecular Biology*, 896, 297–318.
- Fontana, A., Zamboni, M., Polverino de Laureto, P., De Filippis, V., Clementi, A., & Scaramella, E. (1997). Probing the conformational state of apomyoglobin by limited proteolysis. *Journal of Molecular Biology*, 266, 223–230.
- Foulkes, W. D. (2004). BRCA1 functions as a breast stem cell regulator. *Journal of Medical Genetics*, 41, 1–5.
- Fritzinger, D. C., Bredehorst, R., & Vogel, C. W. (1994). Molecular cloning and derived primary structure of cobra venom factor. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 12775–12779.
- Grimsley, G. R., & Pace, C. N. (2004). Spectrophotometric determination of protein concentration. *Current Protocols in Protein Science*, Unit 3. 1, 3.1.1–3.1.9.
- Haile, D. T., & Parvin, J. D. (1999). Activation of transcription *in vitro* by the BRCA1 carboxyl-terminal domain. *Journal of Biological Chemistry*, 274, 2113–2117.
- Harper, J. W., & Elledge, S. J. (2007). The DNA damage response: Ten years after. *Molecular Cell*, 28, 739–745.
- Havlis, J., Thomas, H., Sebela, M., & Shevchenko, A. (2003). Fast-response proteomics by accelerated in-gel digestion of proteins. *Analytical Chemistry*, 75, 1300–1306.
- Hegyi, H., & Gerstein, M. (1999). The relationship between protein structure and function: A comprehensive survey with application to the yeast genome. *Journal of Molecular Biology*, 288, 147–164.
- Hofmann, K., & Bucher, P. (1998). The PCI domain: A common theme in three multiprotein complexes. *Trends in Biochemical Sciences*, 23, 204–205.
- Holm, L., & Rosenstrom, P. (2010). Dali server: Conservation mapping in 3D. *Nucleic Acids Research*, 38, W545–549.
- Hu, Y., Scully, R., Sobhian, B., Xie, A., Shestakova, E., & Livingston, D. M. (2011). RAP80-directed tuning of BRCA1 homologous recombination function at ionizing radiation-induced nuclear foci. *Genes & Development*, 25, 685–700.
- Huen, M. S., & Chen, J. (2008). The DNA damage response pathways: At the crossroad of protein modifications. *Cell Research*, 18, 8–16.
- Huen, M. S., Grant, R., Manke, I., Minn, K., Yu, X., Yaffe, M. B., & Chen, J. (2007). RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly. *Cell*, 131, 901–914.
- Jimenez, C. R., Huang, L., Qiu, Y., & Burlingame, A. L. (2001). In-gel digestion of proteins for MALDI-MS fingerprint mapping. *Current protocols in protein science*, Unit 16. 14, 1604–1614.
- Kim, D. E., Chivian, D., & Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Research*, 32, W526–531.
- Kim, H., Huang, J., & Chen, J. (2007). CCDC98 is a BRCA1-BRCT domain-binding protein involved in the DNA damage response. *Nature Structural & Molecular Biology*, 14, 710–715.
- Kolas, N. K., Chapman, J. R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F. D., ... Durocher, D. (2007). Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science*, 318, 1637–1640.
- Lobley, A., Whitmore, L., & Wallace, B. A. (2002). DICHROWEB: An interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics*, 18, 211–212.
- Mailand, N., Bekker-Jensen, S., Fastrup, H., Melander, F., Bartek, J., Lukas, C., & Lukas, J. (2007). RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell*, 131, 887–900.
- Martin, S. G., Laroche, T., Suka, N., Grunstein, M., & Gasser, S. M. (1999). Relocalization of telomeric Ku and SIR proteins in response to DNA strand breaks in yeast. *Cell*, 97, 621–633.
- Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., Ding, W., ... Bogden, R. (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, 266, 66–71.
- Mills, K. D., Sinclair, D. A., & Guarente, L. (1999). MEC1-dependent redistribution of the Sir3 silencing protein from telomeres to DNA double-strand breaks. *Cell*, 97, 609–620.
- Monteiro, A. N., August, A., & Hanafusa, H. (1996). Evidence for a transcriptional activation function of BRCA1 C-terminal region. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 13595–13599.
- Morris, A. L., MacArthur, M. W., Hutchinson, E. G., & Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. *Proteins*, 12, 345–364.
- Muniz, V. A., Srinivasan, S., Boswell, S. A., Meinhold, D. W., Childs, T., Osuna, R., & Colon, W. (2011). The role of the local environment of engineered Tyr to Trp substitutions for probing the denaturation mechanism of FIS. *Protein Science*, 20, 302–312.
- Nikkila, J., Coleman, K. A., Morrissey, D., Pylkas, K., Erkkö, H., Messick, T. E., ... Greenberg, R. A. (2009). Familial breast cancer screening reveals an alteration in the RAP80 UIM domain that impairs DNA damage response function. *Oncogene*, 28, 1843–1852.
- Noble, J. E., & Bailey, M. J. (2009). Quantitation of protein. *Methods in Enzymology*, 463, 73–95.
- Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in Enzymology*, 131, 266–280.
- Pace, C. N., & Shaw, K. L. (2000). Linear extrapolation method of analyzing solvent denaturation curves. *Proteins Supplements*, 4, 1–7.
- Pace, C. N., Shirley, B. A., McNutt, M., & Gajiwala, K. (1996). Forces contributing to the conformational stability of proteins. *FASEB Journal*, 10, 75–83.

- Paull, T. T., Rogakou, E. P., Yamazaki, V., Kirchgessner, C. U., Gellert, M., & Bonner, W. M. (2000). A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Current Biology*, 10, 886–895.
- Rajagopalan, K., Mooney, S. M., Parekh, N., Getzenberg, R. H., & Kulkarni, P. (2011). A majority of the cancer/testis antigens are intrinsically disordered proteins. *Journal of Cellular Biochemistry*, 112, 3256–3267.
- Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*, 7, 95–99.
- Ramachandran, G. N., & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Advances in Protein Chemistry*, 23, 283–438.
- Rogakou, E. P., Boon, C., Redon, C., & Bonner, W. M. (1999). Megabase chromatin domains involved in DNA double-strand breaks *in vivo*. *Journal of Cell Biology*, 146, 905–916.
- Sael, L., Chitale, M., & Kihara, D. (2012). Structure- and sequence-based function prediction for non-homologous proteins. *Journal of Structural and Functional Genomics*, 13, 111–123.
- Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., & Wolfson, H. J. (2005). PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Research*, 33, W363–367.
- Scully, R., Chen, J., Ochs, R. L., Keegan, K., Hoekstra, M., Feunteun, J., & Livingston, D. M. (1997). Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell*, 90, 425–435.
- Scully, R., Ganesan, S., Vlasakova, K., Chen, J., Socolovsky, M., & Livingston, D. M. (1999). Genetic analysis of BRCA1 function in a defined tumor cell line. *Molecular Cell*, 4, 1093–1099.
- Shao, G., Patterson-Fortin, J., Messick, T. E., Feng, D., Shanbhag, N., Wang, Y., & Greenberg, R. A. (2009). MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks. *Genes & Development*, 23, 740–754.
- Shiotani, B., & Zou, L. (2009). ATR signaling at a glance. *Journal of Cell Science*, 122, 301–304.
- Solyom, S., Patterson-Fortin, J., Pylkas, K., Greenberg, R. A., & Winqvist, R. (2010). Mutation screening of the MERIT40 gene encoding a novel BRCA1 and RAP80 interacting protein in breast cancer families. *Breast Cancer Research and Treatment*, 120, 165–168.
- Sreerama, N., Venyaminov, S. Y., & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: Inclusion of denatured proteins with native proteins in the analysis. *Analytical Biochemistry*, 287, 243–251.
- Stephens, P. J., McKenna, C. E., McKenna, M. C., Nguyen, H. T., & Devlin, F. (1981). Circular dichroism and magnetic circular dichroism of reduced molybdenum-iron protein of *Azotobacter vinelandii* nitrogenase. *Biochemistry*, 20, 2857–2864.
- Uversky, V. N., & Dunker, A. K. (2010). Understanding protein non-folding. *Biochimica et Biophysica Acta*, 1804, 1231–1264.
- Uversky, V. N., Oldfield, C. J., & Dunker, A. K. (2008). Intrinsically disordered proteins in human diseases: Introducing the D2 concept. *Annual Review of Biophysics*, 37, 215–246.
- Venkitaraman, A. R. (1999). Breast cancer genes and DNA repair. *Science*, 286, 1100–1102.
- Venkitaraman, A. R. (2002). Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell*, 108, 171–182.
- Vikrant, Kumar, R., Yadav, L. R., Nakhwa, P., Waghmare, S. K., Goyal, P., & Varma, A. K. (2013). Structural and functional implication of RAP80 DeltaGlu81 mutation. *PLoS One*, 8, e72707.
- Vogel, C. W., & Fritzinger, D. C. (2010). Cobra venom factor: Structure, function, and humanization for therapeutic complement depletion. *Toxicon*, 56, 1198–1222.
- Vogel, C. W., Smith, C. A., & Muller-Eberhard, H. J. (1984). Cobra venom factor: Structural homology with the third component of human complement. *The Journal of Immunology*, 133, 3235–3241.
- Walters, J., Milam, S. L., & Clark, A. C. (2009). Practical approaches to protein folding and assembly: Spectroscopic strategies in thermodynamics and kinetics. *Methods in Enzymology*, 455, 1–39.
- Wang, B., & Elledge, S. J. (2007). Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brcal/Brcc36 complex in response to DNA damage. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 20759–20763.
- Wang, B., Hurov, K., Hofmann, K., & Elledge, S. J. (2009). NBA1, a new player in the Brcal A complex, is required for DNA damage resistance and checkpoint control. *Genes & Development*, 23, 729–739.
- Wang, B., Matsuoka, S., Ballif, B. A., Zhang, D., Smogorzewska, A., Gygi, S. P., & Elledge, S. J. (2007). Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science*, 316, 1194–1198.
- Whitmore, L., & Wallace, B. A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research*, 32, W668–673.
- Whitmore, L., & Wallace, B. A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers*, 89, 392–400.
- Williams, R. S., & Glover, J. N. (2003). Structural consequences of a cancer-causing BRCA1-BRCT missense mutation. *Journal of Biological Chemistry*, 278, 2630–2635.
- Wootton, S. K., & Yoo, D. (2003). Homo-oligomerization of the porcine reproductive and respiratory syndrome virus nucleocapsid protein and the role of disulfide linkages. *Journal of Virology*, 77, 4546–4557.
- Wrzeszczynski, K. O., & Rost, B. (2009). Cell cycle kinases predicted from conserved biophysical properties. *Proteins*, 74, 655–668.
- Xu, X., Weaver, Z., Linke, S. P., Li, C., Gotay, J., Wang, X. W., ... Deng, C. X. (1999). Centrosome amplification and a defective G2-M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform-deficient cells. *Molecular Cell*, 3, 389–395.
- Yadav, S., & Ahmad, F. (2000). A new method for the determination of stability parameters of proteins from their heat-induced denaturation curves. *Analytical Biochemistry*, 283, 207–213.
- Yan, J., & Jetten, A. M. (2008). RAP80 and RNF8, key players in the recruitment of repair proteins to DNA damage sites. *Cancer Letters*, 271, 179–190.
- Yan, Z., Kim, Y. S., & Jetten, A. M. (2002). RAP80, a novel nuclear protein that interacts with the retinoid-related testis-associated receptor. *Journal of Biological Chemistry*, 277, 32379–32388.

- Zerovnik, E. (2011). Oligomerization preceding amyloid fibril formation: A process in common to intrinsically disordered and globular proteins. *Network*, 22, 154–161.
- Zhang, C., Vasmatzis, G., Cornette, J. L., & DeLisi, C. (1997). Determination of atomic desolvation energies from the structures of crystallized proteins. *Journal of Molecular Biology*, 267, 707–726.
- Zhang, X., & Simon, R. (2005). Estimating the number of rate limiting genomic changes for human breast cancer. *Breast Cancer Research and Treatment*, 91, 121–124.
- Zhou, B. B., & Elledge, S. J. (2000). The DNA damage response: Putting checkpoints in perspective. *Nature*, 408, 433–439.



Role of MERIT40 in stabilization of BRCA1 complex: A protein–protein interaction study



Vikrant, Ulka U. Sawant, Ashok K. Varma*

Advanced Centre for Treatment, Research and Education in Cancer, Kharghar, Navi Mumbai, Maharashtra 410210, India

ARTICLE INFO

Article history:

Received 13 March 2014

Available online 22 March 2014

Keywords:

BRCA1-complex stability

MERIT40

RAP80

DNA damage response (DDR)

ABSTRACT

MERIT40 is a novel associate of the BRCA1-complex, thus play an essential role in DNA damage repair mechanism. It is the least implicit protein and its structural and functional aspects of regulating the stability of BRCA1–MERIT40 complex remain equivocal. Analysis of protein–protein interactions between BRCA1 and its cellular binding partners like ABRAXAS, RAP80 and MERIT40 would help to understand the role of protein complex integrity in DNA repair mechanism. The recombinant proteins were purified and their structural aspects were elucidated by spectroscopic methods. Interaction analysis was carried out to determine binding partners of MERIT40. MERIT40 showed interaction with bridging molecule, called ABRAXAS, thus generate a scaffold among various members which further stabilizes the entire complex. It acts as an adapter molecule by interacting with BRCA1–BRCT in non-phosphorylation dependent manner. The feature enlighten on structural and interaction profile of BRCA1-complex member to elucidate their role in complex stability and DNA repair process.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Sensing and responding to DNA damage is the prerequisite for genomic integrity of living organism and their survival [1,2]. ATM (ataxia telangiectasia mutated) and ATR (ATM and Rad3-related) are the pair of related kinases which are activated in the response to DNA damage and acts as sensor molecules in DNA repair cascade [3–5]. Phosphorylation of the initial target H2AX (a histone variant) and MDC1 (Mediator of DNA damage complex 1) by ATM and ATR accumulate γ -H2AX at the damage site [6–8]. γ -H2AX acts as docking site for MDC1 [8–11], a protein having the phospho binding module, which subsequently endorses the congregation of RNF8 and UBC13 [6,12–14]. Furthermore, aftereffect of this signaling event is the formation of polyubiquitin chains on the H2AX by UBC13 and RNF8 [15–18]. K-63 linked polyubiquitin chain on γ -H2AX is exclusively recognized by ubiquitin interacting motif (UIM) of RAP80 (receptor-associated protein 80) [19,20]. RAP80 is a member of BRCA1 complex which encompasses ABRAXAS, MERIT40, and BRCA1. RAP80 has two functional UIMs motifs at N-terminal and a central AIR domain (Abraxas-Interacting Region) [20,21]. ABRAXAS comprises sequence consensus of pS-X-X-F motif at the C-terminal through which it mediates the

interaction with BRCA1 [22]. Similar consensus sequence is present in the BACH1, CtIP which is required for interaction with BRCA1–BRCT domains in phosphorylated dependent manner [23–25]. BRCT domains are the evolutionary conserved domain, identified in several DNA damage repair proteins and mediate protein–protein interactions by binding to phosphoproteins having pSer/(Thr)-X-X-Phe motifs [26,27]. BRCT repeats are indispensable for the tumor repressor functions of BRCA1 and obligatory for its translocation to IRIF (Ionizing Radiation Induced Foci) [28].

Genetic variants discovered in two major breast cancer susceptibility genes BRCA1 and BRCA2 are accounted approximately 5–10% of breast cancer cases [29]. Inherited mutations in the BRCA1 gene alone accounts for more than 50% familial breast cancer and 20–30% inherited ovarian cancers. However, for familial breast cancer, not more than 20% of the cases are reported for germ-line mutations in BRCA1 [30] which concentrates thoughtfulness towards the role of BRCA1 interacting protein in DNA damage repair and cancer. BRCT mutations responsible for breast and ovarian cancer found to have disrupted association between BRCA1 and RAP80 [20,31,32]. Deletion mutation at pSer/(Thr)-X-X-Phe motif in ABRAXAS prevents association of RAP80 and BRCA1, suggesting BRCA1–RAP80 complex is ABRAXAS dependent [22]. Knockdown studies with RAP80, ABRAXAS and BRCA1 have demonstrated that RAP80 is necessary for foci formation of ABRAXAS and BRCA1. As a result of UIM deletion, RAP80 mutant failed to recruit BRCA1 complex at DNA damage site. Cells knockdown for BRCA1 neither shown reduced level of RAP80 foci nor ABRAXAS

* Corresponding author. Address: Tata Memorial Centre, Advanced Centre for Treatment, Research and Education in Cancer, Navi Mumbai, Maharashtra, India. Fax: +91 22 2740 5085.

E-mail address: avarma@actrec.gov.in (A.K. Varma).

foci formation after the DNA damage [28]. It suggests that ABRAXAS act downstream of RAP80 while upstream of BRCA1. Co-immunoprecipitation studies supported the fact that ABRAXAS acts as a linker between BRCA1 and RAP80, suggesting their co-existence in a complex [28]. ABRAXAS is also involved in BRCA1 dependent G2/M checkpoint, suggesting ABRAXAS function in a collaborative manner with BRCA1 [28]. It is currently indistinct how the RAP80 AIR facilitates DSB localization of ABRAXAS, and the entire repertoire of RAP80 AIR-associated proteins control BRCA1-complex DDR functions.

MERIT40 is a mediator required for stabilization of BRCA1-complex [33,34]. SiRNA mediated knockdown of MERIT40 reduces ABRAXAS level drastically at the RAP80 foci, suggesting aborted binding of ABRAXAS to RAP80. It highlights the importance of MERIT40 for RAP80 mediated BRCA1 recruitment at DSB [34]. The indispensable role of MERIT40 in BRCA1 complex integrity was further established when cells knockdown for MERIT40 has shown reduced ABRAXAS–BRCC36 interaction. MERIT40 is also essential for localization of BRCA1-complex at DNA damage site [34]. MERIT40 knockdown unravel the negative effect on the RAP80, BRCC36 and ABRAXAS complex irrespective of BRCA1. Conversely, RAP80 or BRCC36 depletion did not appear to have any influence on the endogenous MERIT40 level, showing MERIT40 has a greater prominence in BRCA1-complex stability than others [34]. Altogether, these assumptions leads the fact which support the vulnerability of BRCA1 associated protein involvement in DNA damage repair. BRCA1-complex plays foremost role in DNA repair, and effective progression of this process sequentially depends on every member.

In the present study, an attempt has been made to understand the functional aspects of the MERIT40 in BRCA1 complex stabilization. We have purified recombinant MERIT40, BRCA1-BRCT, RAP80 and ABRAXAS to establish their role in BRCA1 complex formation. To explore the possibility of crosstalk, a qualitative and quantitative interaction relationship was explored in detail among various members. MERIT40 was found to interact with ABRAXAS and BRCA1-BRCT while no bonding with RAP80 was observed. Our findings provide insights into the stability of MERIT40 and BRCA1 complex as well as the various interactions that maintain the complex integrity. These findings will be helpful in understanding the functional interplay of DNA repair proteins. It will further elucidate mechanisms that abrogate the synergy of the BRCA1 complex members, consequently its repair function.

2. Materials and methods

All the chemicals were analytical or molecular biology grade and purchased from Sigma–Aldrich, unless otherwise specified. For buffer preparation, double distilled, 0.44 μ M filtered Milli-Q water (Millipore, USA) was used. Protein and buffer solutions were filtered and degassed before proceeding to experiment.

2.1. Gene cloning, protein expression and purification

RAP80 ranging from 1 to 405 amino acids, henceforth will referred as RAP80 (kind gift from J. Chen), was PCR amplified (Thermocycler, Bio-Rad) and sub-cloned into pGEX-kT (Kind gift from John A.A. Ladias) expression vector after double restriction digestion with BamH1 and EcoR1. For PCR amplification, primers (Sigma–Aldrich) having TEV protease site (E-N-L-Y-F-Q/S) were used, and ligation was done with T4 DNA ligase. Positive clones were selected on the basis of restriction digestion and DNA sequencing. Similar cloning procedure was followed for MERIT40 and BRCA1 BRCT. Different regions of ABRAXAS ranging from 6 to 268, 6 to 301 and 6 to 373 amino acids, henceforth will referred as ABRAXAS, was sub-cloned in pET28(a+)vector. BRCA1-BRCT

(1653–1853) and MERIT40 was sub-cloned into pGEX-kT. For protein expression and purification, vector construct was transformed into *Escherichia coli* Rosetta (2DE3) cells (Novagen). Protein expression was carried out by inducing bulk culture at *O.D*₆₀₀ between 0.6 and 0.8 with 0.4 mM IPTG (Himedia, India) for 16 h at 24 °C. Culture was harvested and re-suspended in 10 mM HEPES buffer containing 300 mM NaCl, 5 mM BME, 0.1 mM EDTA pH 7.5 (HNBE buffer). Cells suspension was sonicated (Branson Sonifier) thrice with 1 min sonication and 1 min rest in ice at duty cycle of 50 and power 60 unit. Supernatant was collected after centrifugation of culture at 18,000 rpm for 45 min. Soluble protein was passed through the pre-equilibrated glutathione resin/Ni-NTA (Novagen) with HNBE buffer. Fusion protein was allowed to bind with the affinity column and thereafter washed with 10 column volumes of HNBE buffer to remove non-specifically bound proteins. Bound fusion protein was cleaved with TEV protease to elute native protein. Elution fractions were further concentrated and passed on FPLC (AKTA, GE) through the size exclusion chromatography (Superdex 200, GE) to acquire homogenous and pure protein. Purified protein was analyzed on SDS–PAGE.

2.2. Protein estimation

Purified protein was quantified with Bradford protein estimation protocol (Expedon) and Nanodrop (Thermo scientific) as per manufacturer instructions. For Bradford method, different dilutions of BSA were prepared and absorbance was recorded at 595 nm. Concentration of protein of interest was determined by interpolation of standard curve. The experiment was performed in three sets [35], [36,37] and [38,39].

2.3. Fluorescence spectroscopy

Chemical denaturation of protein of interest (2 μ M) was carried out with urea as unfolding agent. Emission maxima of native and unfolded protein (with 8 M urea) was recorded using fluorescence spectrophotometer (Horiba, USA) [40]. Protein was excited at 280 nm and fluorescence emission spectra were collected from 310 to 400 nm range. The experiment was repeated three times.

2.4. Circular Dichroism spectroscopy

Secondary structure characteristic of protein was analyzed using Circular Dichroism (CD) polarimeter (Jasco J-810, Japan). Far-UV CD scan of protein (10 μ M) was collected in the 200–240 nm wavelength range. 10 μ M individual protein in 2.5 mM HEPES pH 7.5, 50 mM NaCl buffer at 10 °C was scanned for secondary structure characterization. Average blank corrected data of three independent scans was submitted to DichroWeb server to predict the secondary structure composition [41–45].

2.5. GST pull down assay

GST (control), GST-BRCT and GST-MERIT40 were expressed in bacterial system (as describe above) and pellets were resuspended in buffer (10 mM HEPES, 250 mM NaCl, 10 mM β -ME, 0.1 mM EDTA and 5% glycerol) having pH 7.5 followed by sonication. Soluble fraction was separated using centrifugation and incubated with glutathione resin. Bound protein(s) (0.1 mg/ml) was used to capture prey (his-tag-ABRAXAS for GST-MERIT40, his-tag-MERIT40 for GST-BRCT) (10 μ g) and RAP80 (10 μ g, for GST-MERIT40), and thereafter resin was washed with buffer and loaded on SDS–PAGE. Interacting partner was probed with anti-His antibody (for ABRAXAS and MERIT40) and anti-RAP80 antibody (Abcam, USA). The experiment was repeated three times.

2.6. Isothermal titration calorimeter

Interaction studies between MERIT40 and ABRAXAS, BRCT and RAP80 was performed using Isothermal Titration Calorimetry (MicroCal200, GE). All the samples were degassed thoroughly before setting up titration reaction. 25 μ M BRCT was kept in cell and allowed to titrate with MERIT40 (250 μ M) through repeated injection (2 μ l each) at 25 $^{\circ}$ C with spacing of 2 s. Stirrer speed was constant at 1000 rpm and equilibration time was 4 s in between successive injections. 25 μ M of RAP80 was kept in cell and allowed to interact with 250 μ M MERIT40. Similarly, MERIT40 was taken as titrant (30 μ M) and ABRAXAS (290 μ M) as titer with the above listed parameters. Data fitting was done using in-built origin 5.0 software and binding affinity was determined. The experiment was repeated twice independently.

3. Results

MERIT40 is an indispensable partner of the BRCA1 and helps in complex stability. It perpetuates the level of RAP80 and ABRAXAS and facilitates recruitment of BRCA1 at the site of DNA damage, hence favors homologous recombination repair. With an objective to understand the role of MERIT40, we have studied structural and functional characterization of BRCA1 complex members. MERIT40, BRCA1-BRCT, RAP80 and ABRAXAS native proteins were purified and analyzed on SDS-PAGE (Fig. 1A). Purified protein were gel extracted and subjected to MALDI-TOF mass spectrometry for identification [46,47]. MERIT40, BRCA1, RAP80 and ABRAXAS showed their correct identity in ExPASy server (<http://www.expasy.org/>) through mascot analysis (Table 1). It illustrates the faithfulness of purified protein for further use in structural and functional analysis.

3.1. Structural characterization

Secondary and tertiary structural characteristics are the important benchmark for defining folded-ness and functional upshot of a protein(s). A well folded protein showed a typical secondary

structure signature as well as localized behavior of aromatic and hydrophobic residues. Tyrosine and tryptophan are the most extensively used intrinsic fluorophores to study the micro-environmental changes induced in the protein due to external stimulus such as temperature and chaotrophs [48]. In a well folded conformation of protein, these residues are generally buried inside the hydrophobic core and are extremely sensitive to modification around their micro-environment.

In order to evaluate the secondary structural characteristics of MERIT40, RAP80, ABRAXAS, and BRCA1, far-UV CD spectrum were recorded. Secondary structural component of purified protein was analyzed using DichroWeb server (<http://dichroweb.cryst.bbk.ac.uk>) [41–45]. In MERIT40, percentage of α -helices and β -sheets were found to be 20% and 16.6% while BRCA1 BRCT showed it 22% and 27.1%, respectively (Fig. 1B). RAP80 and ABRAXAS showed the presence of fewer α -helices (21.1% and 15%), and displayed the prevalence of β -sheets (30.7% and 24%) (Fig. 1B). Tertiary structure of purified proteins were evaluated with respect to the position of Tyr and Trp residues. Emission spectrum of native and unfolded protein was monitored. Native MERIT40, RAP80, ABRAXAS and BRCA1 showed their emission maxima at 338, 340, 332 and 339 nm respectively, while undergoes a substantial red shift when unfolded (Fig. 1C). It indicates Trp and Tyr residue are buried in three dimensional hydrophobic environments [48]. Altogether, these results also deciphered a well folded state of MERIT40, RAP80, ABRAXAS and BRCA1.

3.2. Interaction analysis

Protein–protein interaction is immensely important to unravel the stability of protein complex at molecular level. BRCA1/BARD1 co-interdependency is an excellent notable example where BRCA1 cannot perform its regular functions in the absence of BARD1 [49]. ABRAXAS also interacts with BRCA1 BRCT through its sequences containing pS-X-X-F motif, and abrogation of this interaction cause severe DNA repair defect [22]. Interaction network in BRCA1-complex is highly crucial and alteration in even a single

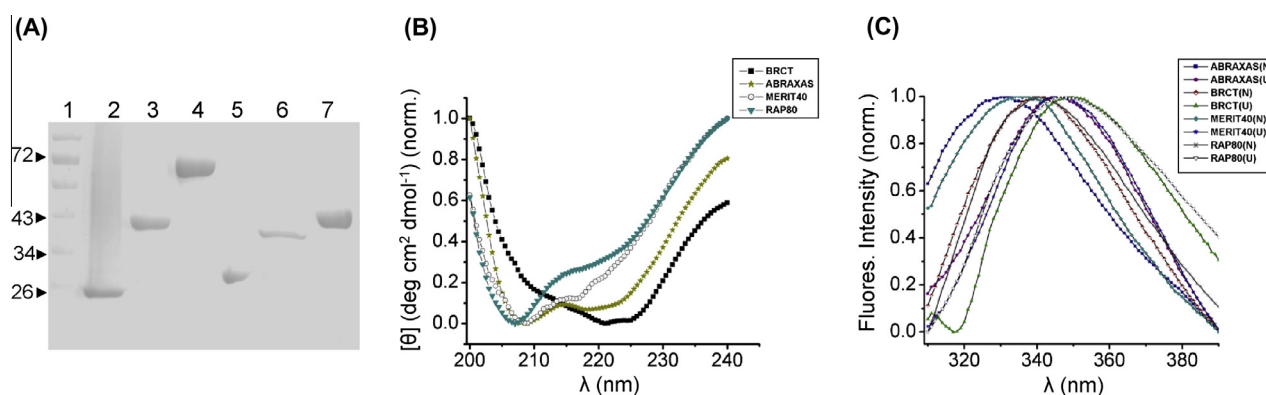


Fig. 1. Purification and structural characterization of ABRAXAS, BRCA1-BRCT, MERIT40 and RAP80. (A) Proteins were purified by affinity and gel filtration chromatography, and loaded on SDS-PAGE. Lane 1-marker, 2-BRCA1-BRCT, 3-MERIT40, 4-RAP80, 5-ABRAXAS (6–268), 6-ABRAXAS (6–301) and 7-ABRAXAS (6–373). (B) Overlay of far-UV CD spectrum and (C) emission maximum of BRCA1-complex proteins. N-native, U-unfolded with urea.

Table 1

Mass spectrometry profiles of MERIT40, BRCA1, ABRAXAS and RAP80 using Mascot analysis.

Protein	Match	Score	Nominal mass (M_r , Dalton)	Calculated pI	Sequence Coverage (%)
MERIT40	h-MERIT40	119	37,050	4.6	35.0
ABRAXAS	h-ABRAXAS	164	47,033	6.6	31.5
RAP80	h-RAP80	85	80,875	5.3	20.1
BRCA1	h-BRCA1	63	208,000	5.0	24.0

h: Human.

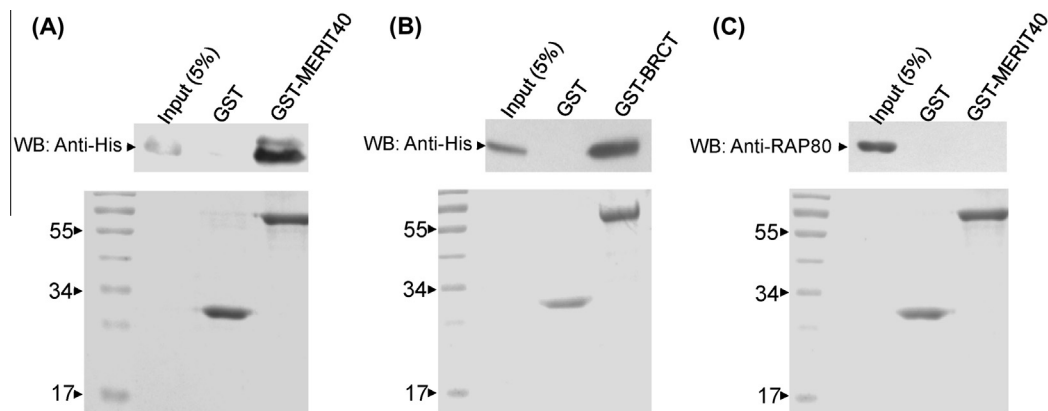


Fig. 2. Pull down assay to determined direct binding partner(s) of MERIT40. Protein–protein interaction analysis of MERIT40 with (A) ABRAXAS, (B) BRCA1-BRCT and (C) RAP80. Lower panel showing the loaded bait protein on SDS–PAGE stained with ponceau. In panel (A)–(C); GST-MERIT40, GST-BRCT and GST-MERIT40 were taken as bait and western blotting was performed against the his-ABRAXAS (prey), his-MERIT40 (prey) and RAP80 (prey), respectively.

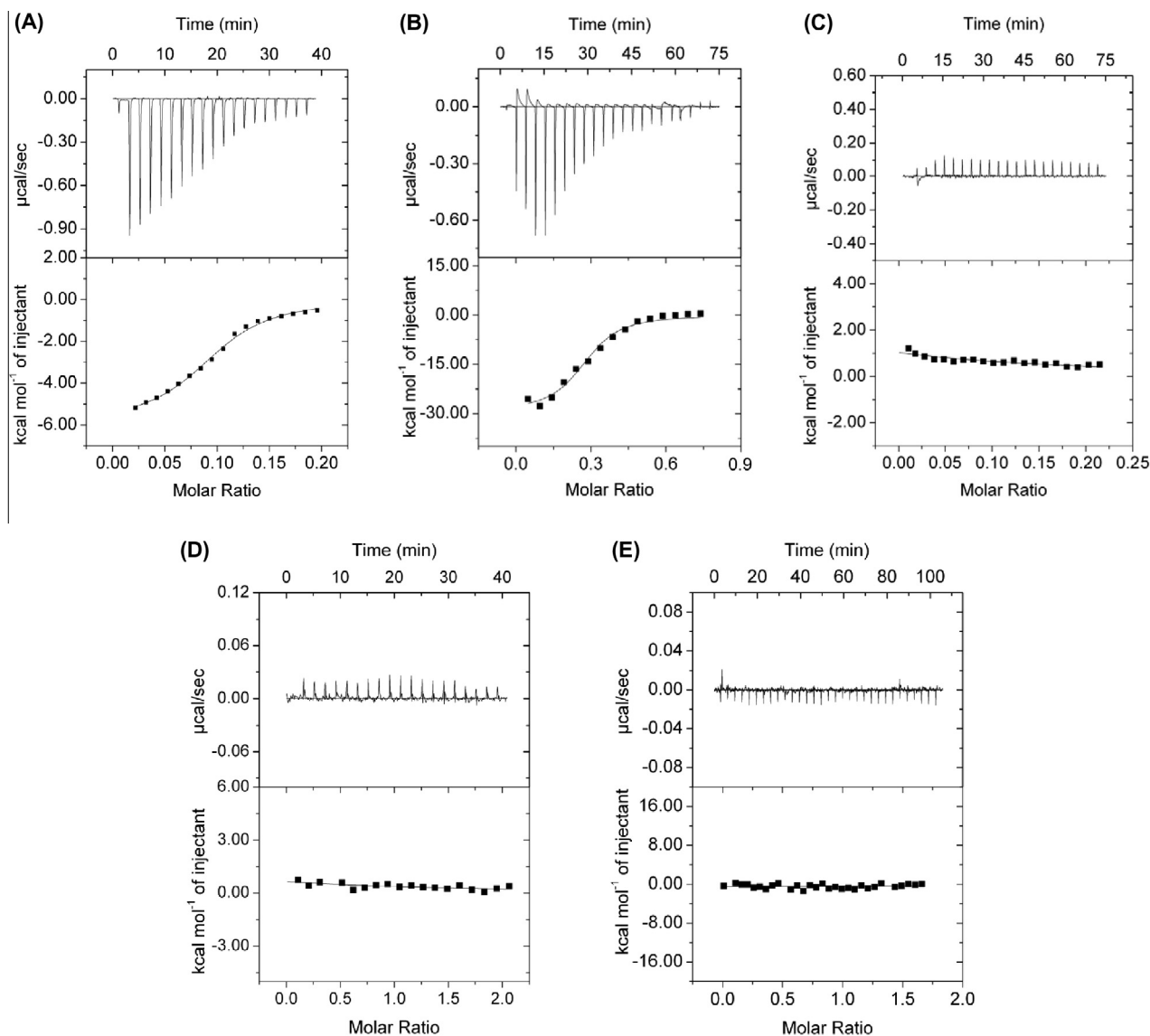


Fig. 3. Isothermal titration calorimetry of MERIT40 with BRCA1-complex members to determined binding interactions and affinity constants. Protein–protein interaction analysis of MERIT40 with (A) BRCA1-BRCT, (B) ABRAXAS (6–373), (C) ABRAXAS (6–301), (D) ABRAXAS (6–268) and (E) RAP80. The binding isotherm showed direct interaction of MERIT40 with ABRAXAS and BRCA1-BRCT.

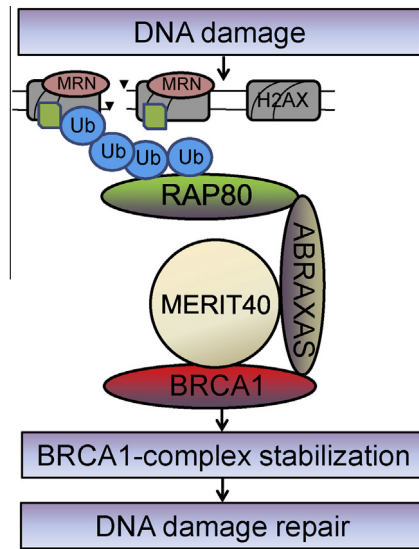


Fig. 4. Anticipated model of MERIT40 mediated BRCA1-complex stabilization. The representative illustrate a possible mechanism of complex stabilization through direct interaction of MERIT40 with BRCA1 and ABRAXAS. Ionization radiation induced DNA double strand break (pointed arrowhead) causes ATM/ATR kinase activation followed by assembly of various damage repair proteins at nucleosome complex. In the sustenance of MERIT40, the polyubiquitin chain(s) formed on H2AX are recognized by RAP80 thereby recruiting the intact BRCA1-complex to the DNA damage site.

amino acids of any of the member in the complex leads to destabilization of whole complex [34]. MERIT40 is a novel associate of this complex, however it remain elusive how MERIT40 establish a network among various member which maintain the BRCA1-complex integrity. To appraise this, we have carried out the interaction analysis of MERIT40 with RAP80, ABRAXAS and BRCA1 BRCT using pull down assay and Isothermal Titration Calorimetry.

A quantitative pull down interaction analysis has established the direct interaction between MERIT40 and BRCA1-BRCT, ABRAXAS (Fig. 2A, B). However, we have not observed significant binding between MERIT40 and RAP80, indicating their interaction is either absent or highly transient (Fig. 2C). To validate these findings, we have perform isothermal titration calorimetry among members of BRCA1 complex. The observed ITC isotherm of MERIT40 and BRCA1-BRCT interaction depict a low binding affinity ($K_d = 9.8 \pm 0.66 \mu\text{M}$) but found somewhat cooperative (Fig. 3A). This further creates a possibility of MERIT40 existence in neighborhood of direct BRCT binding proteins, such as ABRAXAS. Interaction profile of MERIT40 and ABRAXAS (6–373) suggest that ABRAXAS C-terminal region involve in interaction with MERIT40 with a moderate binding affinity ($K_d = 1.7 \pm 0.28 \mu\text{M}$) (Fig. 3B). To further narrow down the minimal binding domain of ABRAXAS with MERIT40, we have performed interaction analysis with shorter region (6–301, 6–268). Absence of interaction between MERIT40, and ABRAXAS (6–301) and (6–268) amino acids showed that probable binding domain of ABRAXAS resides are at the C-terminal regions only while N-terminal of ABRAXAS mainly interacts with RAP80 (Fig. 3C, D) [22]. Since RAP80 is an upstream player in BRCA1-complex, it would be interesting to observe the binding profile of RAP80 and MERIT40. Surprisingly, MERIT40 does not showed interaction with RAP80 (Fig. 3E). Collectively, these binding profiles suggest that MERIT40 showed direct interaction with BRCA1-BRCT as well as ABRAXAS.

4. Discussion

Interactive relationship of MERIT40 with ABRAXAS as well as BRCA1-BRCT mediates the formation of stable BRCA1 complex.

However, MERIT40 does not exhibit consensus sequence pS-X-X-F motif as ABRAXAS, therefore, it might be interacting with BRCA1-BRCT through some other motif. Direct interaction of MERIT40 with ABRAXAS and BRCA1-BRCT categories it as a mediator molecule required for the stabilization of the BRCA1-complex. Knockdown of MERIT40 leads to reduced ABRAXAS foci as well as abrogating ABRAXAS–RAP80 interaction. However, BRCA1 level remain unaffected implying MERIT40 is required for integrity and stability of the BRCA1-complex at DNA damage site. Hence, it can be concluded that MERIT40 is the core mediator molecule and obligatory for organization of stable BRCA1-complex through binding with ABRAXAS and BRCA1-BRCT. Crosstalk between MERIT40 and ABRAXAS probably help in bridging the interactions among BRCA1-complex (Fig. 4). Our study will be helpful in better understanding the accomplishment of MERIT40 in BRCA1-complex and hence DNA damage repair. It will further divulge the molecular organization of this complex and novel associate require maintaining its integrity and stability.

Acknowledgments

Funding for this study was supported by DBT-India (Nos. BT/PR10765/BRB/664/2008 and BT/PR12565/BID/07/303/2009), Seed in Air grant from TMC. We thank DBT-BTIS facility at ACTREC for providing necessary software to this study. Mr. Vikrant thanks CSIR for fellowship (09/513 (0072)/2008-EMR-1). We thanks Dr. Hosur and Dr. Lata (BARC) for data analysis.

References

- [1] B.B. Zhou, S.J. Elledge, The DNA damage response: putting checkpoints in perspective, *Nature* 408 (2000) 433–439.
- [2] J. Rouse, S.P. Jackson, Interfaces between the detection, signaling, and repair of DNA damage, *Science* 297 (2002) 547–551.
- [3] S. Matsuo, B.A. Ballif, A. Smogorzewska, E.R. McDonald 3rd, K.E. Hurov, J. Luo, C.E. Bakalarski, Z. Zhao, N. Solimini, Y. Lerenthal, Y. Shiloh, S.P. Gygi, S.J. Elledge, ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage, *Science* 316 (2007) 1160–1166.
- [4] B. Shiotani, L. Zou, ATR signaling at a glance, *J. Cell Sci.* 122 (2009) 301–304.
- [5] K.K. Khanna, M.F. Lavin, S.P. Jackson, T.D. Mulhern, ATM, a central controller of cellular responses to DNA damage, *Cell Death Differ.* 8 (2001) 1052–1065.
- [6] S. Burma, B.P. Chen, M. Murphy, A. Kurimasa, D.J. Chen, ATM phosphorylates histone H2AX in response to DNA double-strand breaks, *J. Biol. Chem.* 276 (2001) 42462–42467.
- [7] M. Stucki, S.P. Jackson, GammaH2AX and MDC1: anchoring the DNA-damage-response machinery to broken chromosomes, *DNA Repair (Amsterdam)* 5 (2006) 534–543.
- [8] Z. Lou, K. Minter-Dykhouse, S. Franco, M. Gostissa, M.A. Rivera, A. Celeste, J.P. Manis, J. van Deursen, A. Nussenzweig, T.T. Paull, F.W. Alt, J. Chen, MDC1 maintains genomic stability by participating in the amplification of ATM-dependent DNA damage signals, *Mol. Cell* 21 (2006) 187–200.
- [9] C.H. Bassing, F.W. Alt, H2AX may function as an anchor to hold broken chromosomal DNA ends in close proximity, *Cell Cycle* 3 (2004) 149–153.
- [10] G.S. Stewart, B. Wang, C.R. Bignell, A.M. Taylor, S.J. Elledge, MDC1 is a mediator of the mammalian DNA damage checkpoint, *Nature* 421 (2003) 961–966.
- [11] M. Stucki, J.A. Clapperton, D. Mohammad, M.B. Yaffe, S.J. Smerdon, S.P. Jackson, MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks, *Cell* 123 (2005) 1213–1226.
- [12] E.P. Rogakou, D.R. Pilch, A.H. Orr, V.S. Ivanova, W.M. Bonner, DNA double-strand breaks induce histone H2AX phosphorylation on serine 139, *J. Biol. Chem.* 273 (1998) 5858–5868.
- [13] A. Celeste, O. Fernandez-Capetillo, M.J. Kruhlak, D.R. Pilch, D.W. Staudt, A. Lee, R.F. Bonner, W.M. Bonner, A. Nussenzweig, Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks, *Nat. Cell Biol.* 5 (2003) 675–679.
- [14] J. Chen, D.P. Silver, D. Walpita, S.B. Cantor, A.F. Gazdar, G. Tomlinson, F.J. Couch, B.L. Weber, T. Ashley, D.M. Livingston, R. Scully, Stable interaction between the products of the BRCA1 and BRCA2 tumor suppressor genes in mitotic and meiotic cells, *Mol. Cell* 2 (1998) 317–328.
- [15] M.S. Huen, R. Grant, I. Manke, K. Minn, X. Yu, M.B. Yaffe, J. Chen, RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly, *Cell* 131 (2007) 901–914.
- [16] N.K. Kolas, J.R. Chapman, S. Nakada, J. Ylanko, R. Chahwan, F.D. Sweeney, S. Panier, M. Mendez, J. Wildenhain, T.M. Thomson, L. Pelletier, S.P. Jackson, D. Durocher, Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase, *Science* 318 (2007) 1637–1640.

- [17] N. Mailand, S. Bekker-Jensen, H. Fastrup, F. Melander, J. Bartek, C. Lukas, J. Lukas, RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins, *Cell* 131 (2007) 887–900.
- [18] B. Wang, S.J. Elledge, Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brc1/Brc36 complex in response to DNA damage, *Proc. Natl. Acad. Sci. U.S.A* 104 (2007) 20759–20763.
- [19] H. Kim, J. Chen, X. Yu, Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response, *Science* 316 (2007) 1202–1205.
- [20] J. Yan, A.M. Jetten, RAP80 and RNF8, key players in the recruitment of repair proteins to DNA damage sites, *Cancer Lett.* 271 (2008) 179–190.
- [21] Z. Yan, Y.S. Kim, A.M. Jetten, RAP80, a novel nuclear protein that interacts with the retinoid-related testis-associated receptor, *J. Biol. Chem.* 277 (2002) 32379–32388.
- [22] H. Kim, J. Huang, J. Chen, CCDC98 is a BRCA1-BRCT domain-binding protein involved in the DNA damage response, *Nat. Struct. Mol. Biol.* 14 (2007) 710–715.
- [23] K.L. Gorrington, D.Y. Choong, G.J. Lindeman, J.E. Visvader, I.G. Campbell, Breast cancer risk and the BRCA1 interacting protein CtIP, *Breast Cancer Res. Treat.* 112 (2008) 351–352.
- [24] X. Yu, L.C. Wu, A.M. Bowcock, A. Aronheim, R. Baer, The C-terminal (BRCT) domains of BRCA1 interact in vivo with CtIP, a protein implicated in the CtBP pathway of transcriptional repression, *J. Biol. Chem.* 273 (1998) 25388–25392.
- [25] X. Yu, J. Chen, DNA damage-induced cell cycle checkpoint control requires CtIP, a phosphorylation-dependent binding partner of BRCA1 C-terminal domains, *Mol. Cell. Biol.* 24 (2004) 9478–9486.
- [26] X. Yu, C.C. Chini, M. He, G. Mer, J. Chen, The BRCT domain is a phospho-protein binding domain, *Science* 302 (2003) 639–642.
- [27] I.A. Manke, D.M. Lowery, A. Nguyen, M.B. Yaffe, BRCT repeats as phosphopeptide-binding modules involved in protein targeting, *Science* 302 (2003) 636–639.
- [28] Z. Liu, J. Wu, X. Yu, CCDC98 targets BRCA1 to DNA damage sites, *Nat. Struct. Mol. Biol.* 14 (2007) 716–720.
- [29] B. Newman, M.A. Austin, M. Lee, M.C. King, Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families, *Proc. Natl. Acad. Sci. U.S.A* 85 (1988) 3044–3048.
- [30] Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Anglian Breast Cancer Study Group, *Br. J. Cancer* 83 (2000) 1301–1308.
- [31] B. Sobhian, G. Shao, D.R. Lilli, A.C. Culhane, L.A. Moreau, B. Xia, D.M. Livingston, R.A. Greenberg, RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites, *Science* 316 (2007) 1198–1202.
- [32] J. Yan, X.P. Yang, Y.S. Kim, A.M. Jetten, RAP80 responds to DNA damage induced by both ionizing radiation and UV irradiation and is phosphorylated at Ser 205, *Cancer Res.* 68 (2008) 4269–4276.
- [33] L. Feng, J. Huang, J. Chen, MERIT40 facilitates BRCA1 localization and DNA damage repair, *Genes Dev.* 23 (2009) 719–728.
- [34] G. Shao, J. Patterson-Fortin, T.E. Messick, D. Feng, N. Shanbhag, Y. Wang, R.A. Greenberg, MERIT40 controls BRCA1–Rap80 complex integrity and recruitment to DNA double-strand breaks, *Genes Dev.* 23 (2009) 740–754.
- [35] J.E. Noble, M.J. Bailey, Quantitation of protein, *Methods Enzymol.* 463 (2009) 73–95.
- [36] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding, *Anal. Biochem.* 72 (1976) 248–254.
- [37] Vikrant, R. Kumar, L.R. Yadav, P. Nakhwa, S.K. Waghmare, P. Goyal, A.K. Varma, structural and functional implication of RAP80 DeltaGlu81 mutation, *PLoS One* 8 (2013) e72707.
- [38] G.R. Grimsley, C.N. Pace, Spectrophotometric determination of protein concentration, *Curr. Protoc. Protein Sci.* (2004) (Chapter 3, Unit 3.1).
- [39] Vikrant, P. Nakhwa, D.C. Badgular, R. Kumar, K.K. Rathore, A.K. Varma, Structural and functional characterization of the MERIT40 to understand its role in DNA repair, *J. Biomol. Struct. Dyn.* 3 (2013) 3–8.
- [40] J. Walters, S.L. Milam, A.C. Clark, Practical approaches to protein folding and assembly: spectroscopic strategies in thermodynamics and kinetics, *Methods Enzymol.* 455 (2009) 1–39.
- [41] L. Whitmore, B.A. Wallace, Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases, *Biopolymers* 89 (2008) 392–400.
- [42] A. Lobley, L. Whitmore, B.A. Wallace, DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra, *Bioinformatics* 18 (2002) 211–212.
- [43] L. Whitmore, B.A. Wallace, DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, *Nucleic Acids Res.* 32 (2004) W668–W673.
- [44] P.J. Stephens, C.E. McKenna, M.C. McKenna, H.T. Nguyen, F. Devlin, Circular dichroism and magnetic circular dichroism of reduced molybdenum-iron protein of *Azotobacter vinelandii* nitrogenase, *Biochemistry* 20 (1981) 2857–2864.
- [45] N. Sreerama, S.Y. Venyaminov, R.W. Woody, Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis, *Anal. Biochem.* 287 (2000) 243–251.
- [46] J. Havlis, H. Thomas, M. Sebela, A. Shevchenko, Fast-response proteomics by accelerated in-gel digestion of proteins, *Anal. Chem.* 75 (2003) 1300–1306.
- [47] C.R. Jimenez, L. Huang, Y. Qiu, A.L. Burlingame, In-gel digestion of proteins for MALDI-MS fingerprint mapping, *Curr. Protoc. Protein Sci.* (2001) (Chapter 16, Unit 16 14.).
- [48] V.A. Muniz, S. Srinivasan, S.A. Boswell, D.W. Meinhold, T. Childs, R. Osuna, W. Colon, The role of the local environment of engineered Tyr to Trp substitutions for probing the denaturation mechanism of FIS, *Protein Sci.* 20 (2011) 302–312.
- [49] A. Atipairin, B. Canyuk, A. Ratanaphan, The RING heterodimer BRCA1–BARD1 is a ubiquitin ligase inactivated by the platinum-based anticancer drugs, *Breast Cancer Res. Treat.* 126 (2011) 203–209.

This article was downloaded by: [vikrant .]

On: 21 August 2014, At: 12:52

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Biomolecular Structure and Dynamics

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/tbsd20>

Mislocalization of BRCA1-complex due to ABRAXAS Arg361Gln mutation

Vikrant^a, Rajan Kumar^a, Quadir Siddiqui^a, Nalini Singh^a, Sanjeev K. Waghmare^a & Ashok K. Varma^a

^a Advanced Centre for Treatment, Research and Education in Cancer, Tata Memorial Centre, Kharghar, Navi Mumbai, Maharashtra 410 210, India

Published online: 08 Aug 2014.

To cite this article: Vikrant, Rajan Kumar, Quadir Siddiqui, Nalini Singh, Sanjeev K. Waghmare & Ashok K. Varma (2014): Mislocalization of BRCA1-complex due to ABRAXAS Arg361Gln mutation, Journal of Biomolecular Structure and Dynamics, DOI: [10.1080/07391102.2014.945484](https://doi.org/10.1080/07391102.2014.945484)

To link to this article: <http://dx.doi.org/10.1080/07391102.2014.945484>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Mislocalization of BRCA1-complex due to ABRAXAS Arg361Gln mutation

Vikrant, Rajan Kumar, Quadir Siddiqui, Nalini Singh, Sanjeev K. Waghmare and Ashok K. Varma*

Advanced Centre for Treatment, Research and Education in Cancer, Tata Memorial Centre, Kharghar, Navi Mumbai, Maharashtra 410 210, India

Communicated by Ramaswamy H. Sarma

(Received 10 March 2014; accepted 14 July 2014)

ABRAXAS is an integral member of BRCA1-complex, which helps in its recruitment to the DNA damage site. It interacts with BRCA1 via its C-terminal phospho-peptide binding motif while the N-terminal associates with RAP80, and thereby recruits the BRCA1-complex at the site of DNA damage. Nonetheless, how ABRAXAS helps in the structural integrity of BRCA1-complex, and its DNA repair mechanism remains elusive. To elucidate the role of ABRAXAS in the DNA repair process, we characterized the ABRAXAS wild type and Arg361Gln mutant using *in silico* and *in vitro* approach. It has been observed that ABRAXAS Arg361Gln mutant is responsible for defective nuclear localization of BRCA1-complex, and hence important for DNA repair function. We found conformational changes in ABRAXAS mutant, which impaired binding to RAP80 and further disturb BRCA1-complex localization. The results presented in this paper will help to understand the cause of BRCA1 mislocalization, and various DNA repair defects that occur due to substitution. Comparative study of ABRAXAS wild type and mutant will provide helpful perspective for inhibitor designing that can potentially recompense the deleterious effect(s) of Arg361Gln mutation, and have therapeutic application.

Keywords: structural stability; protein–protein interactions; ABRAXAS; BRCA1; DNA repair

1. Introduction

DNA damage repair is an indispensable phenomenon requisite for maintaining genomic integrity and survival of living organisms in the biological system (Rouse & Jackson, 2002; Zhou & Elledge, 2000). Cells repair its damaged DNA by integrating several processes, including DNA replication, gene regulation, cell cycle, and apoptosis in a complex network of intercommunicating molecules, called DNA damage response (DDR) (Bartek, Bartkova, & Lukas, 2007; Bartek & Lukas, 2007). DDR primarily involves three groups of molecules such as sensor, mediator, and effector (Bartkova et al., 2007; Celeste et al., 2003; Sartori et al., 2007; Shiotani & Zou, 2009a, 2009b). Genotoxic stress causes the double-strand break (DSB) which leads to activation of ATM/ATR kinases. This event executes phosphorylation of histone variant such as H2AX, and MDC1, and promotes their phosphorylation-dependent interaction (Burma, Chen, Murphy, Kurimasa & Chen, 2001; Lou, Minter-Dykhouse, Wu, & Chen 2003; Paull et al., 2000; Rogakou, Pilch, Orr, Ivanova, & Bonner, 1998). γ -H2AX and MDC1 assembly creates a docking site for E3 ubiquitin ligase complex UBC13/RNF8, which further poly-ubiquitinates γ -H2AX (Huen et al., 2007; Kolas et al., 2007; Mailand et al., 2007). The K-63-linked polyubiquitin chain formed on

γ -H2AX is recognized by two tandem Ubiquitin Interacting Motif (UIMs) of adaptor molecule, RAP80 (Huen et al., 2007; Kim, Chen, & Yu, 2007; Kolas et al., 2007; Sato et al., 2009; Vikrant, Kumar et al., 2013). RAP80 has a middle ABRAXAS Interacting Region (AIR) domain through which it interacts with ABRAXAS. ABRAXAS acts as a bridging molecule in the BRCA1-complex which comprises RAP80, ABRAXAS, BRCA1, MERIT40, BRCC36 and BRCC45 (Badgujar, Sawant, Yadav, Hosur, & Varma, 2013; Celeste et al., 2003; Mailand et al., 2007; Sobhian et al., 2007; Vikrant, Sawant, & Varma, 2014; Wang et al., 2000, 2007).

BRCA1-complex foci formation is essential for execution of homologous recombination repair process after DNA damage. siRNA-mediated knockdown of ABRAXAS reduces the BRCA1-complex foci formation after IR-induced DNA damage (Wang & Elledge, 2007). RAP80 double mutant lacking UIM and AIR completely abolishes foci formation, inferring a potential role of ABRAXAS in BRCA1-complex formation (Wang & Elledge, 2007). ABRAXAS and RAP80 knockdown cells showed defective homologous recombination repair and become hypersensitive to IR and UV (Kim, Huang, & Chen, 2007; Wang et al., 2007). Role of ABRAXAS is also suspected in G2/M check point activation since

*Corresponding author. Email: avarma@actrec.gov.in

depletion of RAP80 and ABRAXAS displayed defective cell-cycle checkpoints (Kim, Chen et al., 2007; Wang et al., 2007). ABRAXAS harbors consensus sequence of pS-X-X-F (pS- phosphoserine, X-any amino acid) binding motif at its C-terminal, which interact with BRCA1-BRCT and bring about foci formation (Badgujar et al., 2013; Vikrant, Kumar et al., 2013). ABRAXAS acts upstream of BRCA1 and its knockdown significantly reduce accumulation of BRCA1-complex at the DNA damage site. There are reports that loss of ABRAXAS or BRCA1 could lead to similar phenotype, thereby projecting ABRAXAS as an excellent candidate gene for familial breast cancer study (Liu, Wu & Yu, 2007; Osorio et al., 2009; Sobhian et al., 2007; Wang et al., 2007).

Different mutations that disrupt interactions with phospho-binding partners of BRCA1-BRCT have been reported in the BIC (Breast Cancer Information Core) database (Coquelle, Green, & Glover, 2011; Nikkilä et al., 2009; Pylkas, Erkko, Nikkila, Solyom, & Winqvist, 2008; Williams & Glover, 2003). Familial mutation at the C-terminus of ABRAXAS in Finnish population (Arg361Gln) is suspected to disassemble the BRCA1-complex (Solyom et al., 2012). c.1082G>A alteration was observed among 3 of 125 (2.4%) breast cancer families; however, it was absent in 868 healthy female volunteers (Solyom et al., 2012). This variant impairs nuclear localization of BRCA1-complex and its DNA damage repair function, thereby can predispose an individual to repair defects (Solyom et al., 2012). Immunoprecipitation of epitope tagged ABRAXAS and its mutant Arg361Gln with BRCA1 and other core complex in the cytoplasm displays its failure to form foci in the nucleus after DNA damage, while nuclear localization of wild type is retained (Solyom et al., 2012). Failure to achieve nuclear retention adversely affects the G2/M checkpoint and homology-directed DNA repair, which reduces nuclear retention of ABRAXAS-interacting partners to DSB site (Solyom et al., 2012). Moreover, expression of the Arg361Gln variant causes hypersensitivity to ionizing radiation and reduces BRCA1 localization at the sites of DNA damage (Solyom et al., 2012).

ABRAXAS, as a newly emerging susceptibility gene to cancer predisposition, opens the vast perspective of studying the role and impact of genetic alteration in disease progression. In order to understand the deleterious effects of mutation and possible consequences, we have performed bio-physicochemical characterization and interaction analysis of wild type and mutant. To our knowledge, it is the first multi-model approach which elucidates the structural and functional implications of ABRAXAS and its mutant Arg361Gln. ABRAXAS mutant exhibited relatively less structure distortion and stability; however, sufficient to impair its binding with RAP80. The results presented here will help to

understand the role of Arg361Gln variant of ABRAXAS in BRCA1-complex recruitment and DNA damage repair function. It might further assist in classifying the severity of ABRAXAS Arg361Gln variant in breast cancer development.

2. Results and discussion

2.1. In silico analysis

ABRAXAS is the key member of BRCA1-complex, and share least structural homology with reported structures in the database, which limits full-length *in silico* homology modeling. Consequently, ABRAXAS (351-370) peptide was modeled using the structure of a short region of Amidohydrolase-2 as a template (PDB: 3S4T) which showed highest similarity with ABRAXAS mutation containing region (Kim, Chivian & Baker, 2004; Leaver-Fay et al., 2011). Modeled structure was validated through “SAVES” server, which showed acceptable Ramachandran plot and structural parameters (Ramachandran, Ramakrishnan, & Sasisekharan, 1963; Ramachandran & Sasisekharan, 1968) (Figure 1(A), Figure 1 supplementary). ABRAXAS wild type peptide has shown more of α -helices while the three-dimensional structure (tertiary structure) of mutant peptide appeared different with reduced helicity (Figure 1(A)). Since we have observed localized changes in the structural profile of mutant relative to wild type, it was intriguing to look into such amendments using *in silico* approach.

2.2. Unaltered oligomeric behavior of ABRAXAS mutant

Oligomerization is a common phenomenon observed among disordered and exposed hydrophobic patches containing proteins (Zerovnik, 2011). Substitution mutation generally causes distortion in secondary and tertiary structures, which may lead to the disorderness (Kolchanov, Soloviov, & Zharkikh, 1983; Zhukov, Jaroszewski, &

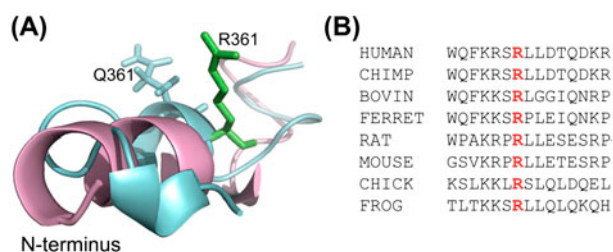


Figure 1. *In silico* structural analysis of ABRAXAS wild type and mutant.

Notes: (A) Superimposed structures of ABRAXAS wild type (pink) and mutant (cyan) peptides. (B) Multiple sequence alignment of ABRAXAS showed highly conserved nature of Arg361 (highlighted in red color) residue among different species.

Bierzynski, 2000). The exposed hydrophobic patches of protein are highly prone to form nonspecific interaction with each other and may result high-order oligomer. ABRAXAS Arg361Gln mutation results in substitution of a basic amino acid residue with nonpolar charged residue that might affect the monomeric nature of protein. However, results obtained from *in silico* study illustrate localized structural changes in mutant, which may affect the protein folding and oligomerization. To explore this prospect, we purified the ABRAXAS wild type and mutant using two-step purification involving affinity and gel filtration chromatography (Figure 2(A) and (B)). Most often, a mutant protein tends to show different solubility, binding, and elution profiles during purification than the wild type provided that the structural conformation of the protein is drastically altered. Interestingly, we did not observe any significant differences in solubility, binding, and elution profiles between wild type and mutant. In gel

filtration chromatography, both the proteins eluted as a monomer at the same elution volume indicating their similar molecular weight (Figure 2(B)) (Table 1). The experimentally determined molecular weight using gel filtration chromatography (wild type: 40KDa, Arg361Gln, 40 KDa) and theoretically derived molecular weight (wild type; 42.9827KDa, Arg361Gln; 42.9547) (Table 1) showed good concurrence, suggesting monomeric nature of both wild type and mutant. Coomassie stained SDS-PAGE of ABRAXAS wild type and mutant showed a single band corresponding to 43 KDa (Figure 2(A)). Therefore, it can be concluded that the Arg361Gln mutation does not affect the monomeric nature of protein.

To determine the oligomeric heterogeneity between wild type and mutant, and the effect of mutation on molecular dimensions, dynamic light scattering (DLS) was used, which provides precise measurement of the molecular dimension of biomolecules in the solution. We observed that the mutant and wild-type protein samples were composed of two different clusters, the first corresponded to the monomeric populations, whereas second for high-order oligomer or aggregates (Figure 2(C)). The observed hydrodynamic diameter of wild type was found to be $9.0 \pm .3$ nm, while the mutant showed increment up to $10.0 \pm .5$ nm. The relative increase in effective diameter of mutant compared to wild type indicates molecular expansion due to substitution mutation. However, the increment in hydrodynamic diameter did not affect the monomeric nature of mutant, as confirmed through size exclusion chromatography (Figure 2(B)). These findings support that Arg361Gln mutation probably incorporates a short- or long-range structural alteration in protein without changing its oligomeric characteristics. The alterations could be either at the secondary or tertiary structural level, and might have a local or global effect.

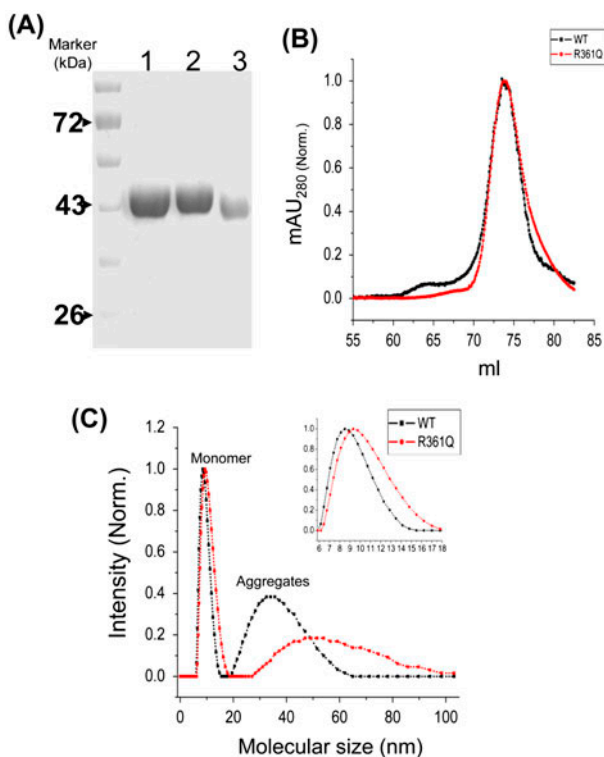


Figure 2. Purification and hydrodynamic diameter measurement of ABRAXAS wild type and mutant.

Notes: (A) SDS-PAGE showing purified protein of ABRAXAS (6-373) wild-type (lane 1), mutant (lane 2), and RAP80 (lane 3). (B) Overlay of gel filtration spectra of wild type and mutant (Superdex 200). Elution profile of both the protein was similar and showed the presence of single peak corresponding to monomer. (C) Overlay of DLS profile of wild type and mutant displayed different hydrodynamic radii, inset is showing the magnified region of the peak corresponding to monomeric population.

2.3. Structural organization

Limited-proteolysis based approach was used to determine the structural alterations in ABRAXAS wild type and mutant (Fontana, de Laureto, Spolaore, & Frare, 2012). In general, a compact globular domain of protein considerably resists the protease digestion, while a disordered or unstructured region, which might have arisen due to mutation, undergoes rapid digestion. If a mutation causes severe damage in the α -helices or β -sheets, it might incorporate disorderedness during protein folding. These unstructured regions undergo rapid digestion due to more accessibility of protease sites (Fontana et al., 2012). To understand domain integrity and determine the stability of ABRAXAS wild type and mutant against the protease digestion, limited digestion with trypsin and chymotrypsin proteases was performed. Wild-type and mutant proteins were treated with equal concentration of proteases for limited time, and the results were analyzed

Table 1. Molecular weight estimation of ABRAXAS wild type and mutant.

	Theoretical Mol. Wt. (kDa) ^a	Ve/Vo ^b	Experimental derived Mol. Wt. (kDa) Gel filtration chromatography
Wild type	42.98	1.6344	40 ± 3.2
Mutant	42.95	1.6344	40 ± 3.2

Note: Ve/Vo: Elution volume/Void volume ratio in gel filtration chromatography (superdex 200 16/60).

^aDetermined from Protparam, Expasy.

^bCompared with standard myoglobin, ovalbumin, albumin, IgG, Ferritin.

on SDS-PAGE (Figure 3(A–D)). ABRAXAS wild-type resistance toward protease digestion was analogous to mutant, which indicates the existence of similar structural domain(s). Furthermore, equivalent susceptibility of mutant protein toward protease digestion suggests that substitution of Arg361Gln does not destabilize the domain integrity of ABRAXAS. This certainly indicates that mutant and wild type are having similar structural pattern with equivalent number of α -helices and β -sheets. It also negates the possibility of major structural changes in the mutant protein, which otherwise could have exposed more protease sites. However, it does not eliminate the possibility of a long- or short-range structural alteration in mutant. Since the observed structural domain (s) organization in limited proteolysis was similar, it appears that wild type and mutant might be having equivalent or modestly changed secondary and tertiary

structures. To explore this likelihood, we compared the secondary structure of ABRAXAS wild type and mutant using far-UV Circular Dichroism (CD) (Figure 4(A)). It was observed that ABRAXAS wild type and mutant have well-defined α/β structure, with β -sheets characteristic being more prominent. Data analysis using Dichroweb server showed that wild type and mutant are having α -helices (wild type 15%, mutant 15.5%) and β -sheets (wild type 24%, mutant 23%) (Lobley, Whitmore & Wallace, 2002). Secondary structure comparison of wild type and mutant deciphered similar α -helical and β -sheets characteristics illustrating minor effect of ABRAXAS Arg361Gln mutation on overall protein secondary structure. These observations suggest that Arg361Gln mutation might be inducing short-range secondary structural changes in mutant, which might in turn collectively responsible for tertiary structure modification.

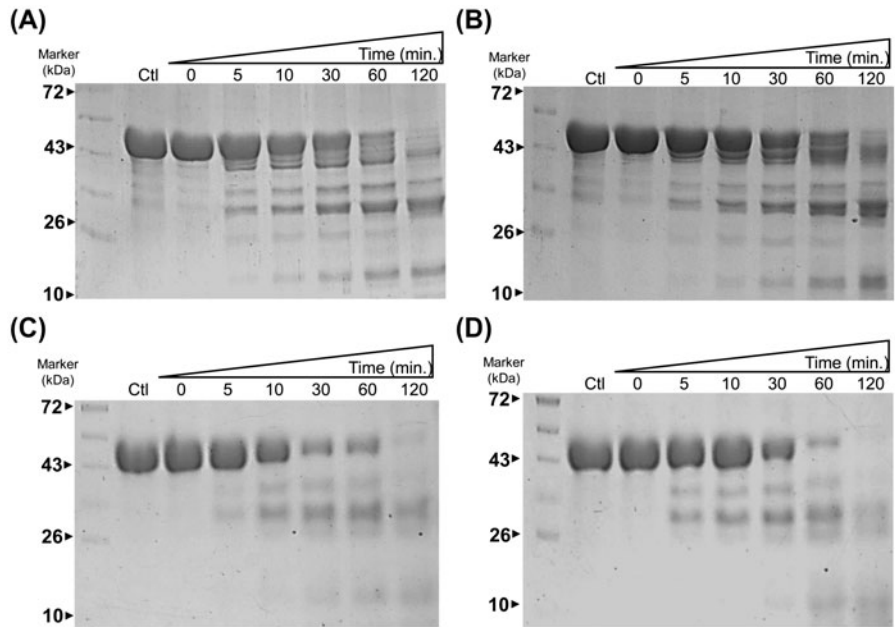


Figure 3. Resistance profile of ABRAXAS wild type and mutant towards protease digestion. Notes: Limited proteolysis of ABRAXAS (6-373) wild type (A, C) and mutant (B, D) using trypsin (A, B) and Chymotrypsin (C, D) as proteases. Wild type and mutant showed relatively equal resistance toward proteolysis as indicated by similar rate of disappearance in band intensity w.r.t time. Ctl- control as untreated with trypsin and chymotrypsin, respectively.

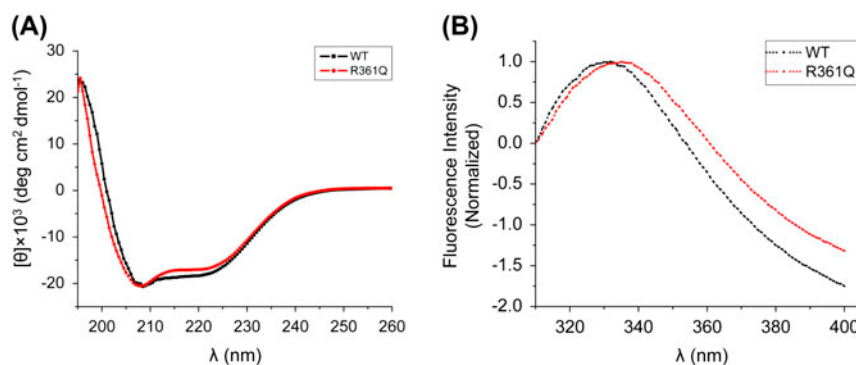


Figure 4. Secondary and tertiary structure evaluation of ABRAXAS wild type and mutant.

Notes: Comparison of secondary structural components of ABRAXAS (6-373) wild type and mutant. (A) Overlay of Far-UV CD spectrum of wild type and mutant. Wild type as well as mutant showed a well defined α/β characteristics and similar structural components. (B) Overlay of fluorescence emission spectrum of wild type and mutant. Wild type showed an emission maxima of 332 nm while mutant undergoes a red shift up to 334 nm.

Intrinsic fluorophores such as tryptophan and tyrosine are most frequently used to monitor the microenvironmental changes induced in the protein tertiary structure (Muniz et al., 2011). These aromatic residues are usually buried inside the protein hydrophobic core and are extremely sensitive to modification in their microenvironment. Since we observed similar secondary structural components in wild type and mutant, it would be interesting to look into the tertiary structure, which might have altered due to mutation. To study the three-dimensional structure of ABRAXAS wild-type and mutant, Trp and Tyr microenvironments were monitored using fluorescence spectroscopy. Protein's emission spectra were recorded at 280 nm to observe the cumulative effect of Trp and Tyr residues' position on the tertiary structure. An emission maxima at λ 332 nm for wild type and 334 nm for Arg361Gln suggested slight changes in Trp and Tyr microenvironments (Figure 4(B)). A red shift in emission maxima of mutant indicates relatively more exposure of Trp and Tyr toward protein hydrophilic surface (i.e. toward protein exterior). However, the shift was very small and might not be sufficient for large structural amendment. Thus, it can be speculated that tertiary structural components become moderately altered without causing any drastic conformational changes.

2.4. Thermal and chemical stability

ABRAXAS Arg361Gln mutation probably does not amend the secondary structural features of protein drastically, as revealed through CD. A mild alteration in tertiary structure was observed, which may or may not affect the stability profile of protein. However, topology of the native state can be the most imperative factor determining the folding profile, and similar structure proteins may unfold through a different pathway

(Zarrine-Afsar, Larson & Davidson, 2005). To evaluate this, thermal stability of ABRAXAS wild type and mutant was compared at secondary (CD) and tertiary structure (fluorescence) level. The spectrum obtained from CD corresponding to λ_{222} nm showed the maximum change in ellipticity and high signal-to-noise ratio. Change in molar ellipticity value was plotted against different temperatures (Figure 5(A)). Melting temperature (T_m), a temperature at which half of the protein assumed to be unfolded, and free energy ($\Delta G^\circ_{H_2O}$) of unfolding process were calculated using linear extrapolation method of curve fitting at temperature 273 K (Pace, 1986; Pace & Shaw, 2000). A higher T_m value usually indicates a more stable protein, and the path of unfolding curve describes about the folding pathway. Thermal stability of ABRAXAS wild-type (T_m 35.0 °C, $\Delta G^\circ_{H_2O}$ 2.1 \pm .50 kcal/mol) was found appreciably equivalent to mutant (T_m 34.8 °C, $\Delta G^\circ_{H_2O}$ 1.53 \pm .60 kcal/mol), and both endured similar folding pathway (Figure 2A, supplementary). For stability assessment of tertiary structure, emission maximum of wild type and mutant was monitored at different temperatures corresponding to λ_{280} and fraction unfolded was calculated. A plot of fraction unfolded against the temperature or unfolding agent concentration gives the information about folding pathway and stability. Fluorescence spectroscopy revealed comparable unfolding pattern but change in stability. The derived T_m value at 298 K temperature was 23.0 \pm 3.1 °C for mutant ($\Delta G^\circ_{H_2O}$ 1.25 \pm .50 kcal/mol) and 27.0 \pm 2.50 °C for wild-type ($\Delta G^\circ_{H_2O}$ 1.08 \pm .38 kcal/mol) (Figure 5(A); Figure 2A, supplementary). As reported earlier, the presence of a biphasic unfolding curve with two transitions indicates the formation of an intermediate species during unfolding process (Walters, Milam, & Clark, 2009). The resultant unfolding pattern can be considered as three-state transition (Walters et al., 2009).

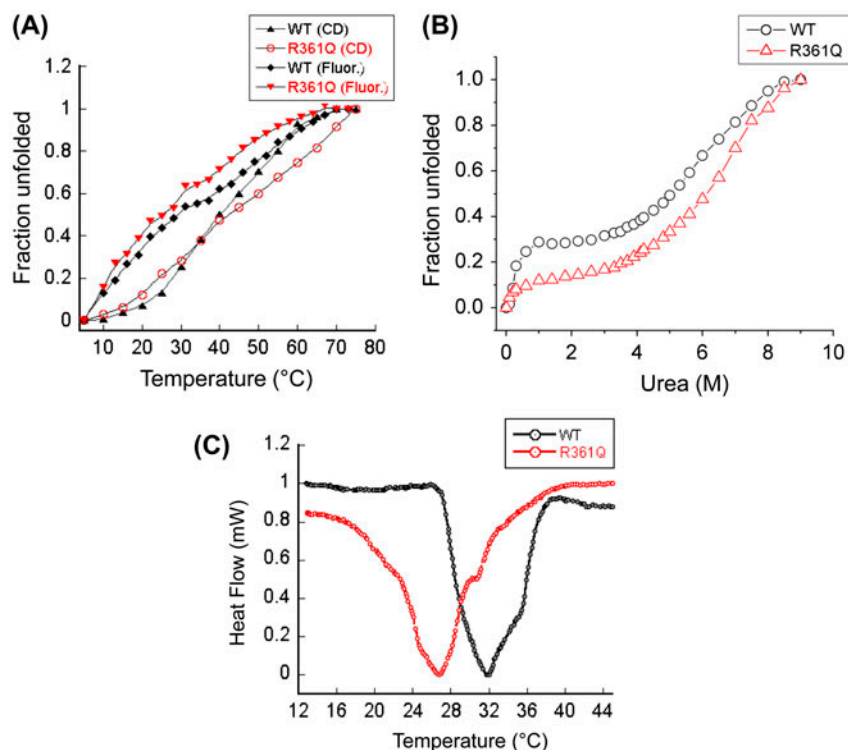


Figure 5. Thermal and chemical stability of ABRAXAS wild type and mutant.

Notes: (A) Thermal stability profile of wild type and mutant showing overlay of protein fraction unfolded at different temperatures. Wild type showed relatively higher thermal stability. (B) Chemical stability profile of wild type and mutant showing overlay of protein fraction unfolded at different urea concentrations. Wild type showed relatively high propensity of intermediate species existence as compared to mutant indicating different unfolding pathways. (C) DSC of wild type and mutant proteins showing a well defined transition around 32 and 27 °C, respectively.

Fluorescence-based thermal denaturation revealed that both wild type and mutant most likely unfold through an intermediate species, and undergo three-state transitions (Figure 5(A)). To determine the chemical stability of wild type and mutant, protein samples (2 μ M) were incubated at different concentrations of urea (0–8 M) until equilibrium achieved, and emission maxima were recorded at a excitation wavelength of λ_{280} . Thermodynamic parameters for ABRAXAS wild type $\Delta G^{\circ}_{H_2O}$ $2.54 \pm .36$ kcal/mol and mutant $\Delta G^{\circ}_{H_2O}$ $2.01 \pm .41$ kcal/mol were calculated by plotting fraction unfolded against urea concentration at temperature 298 K (Figure 2B, supplementary). Chemical stability of mutant was found different, nevertheless, both unfold through acquiring an intermediate species which was more predominate in case of wild-type (Figure 5(B)). Fluorescence-based unfolding study using thermal and chemical denaturation methods suggests a higher stability of wild type as compared to mutant. To substantiate these findings and evaluate the reversibility of unfolding of wild type and mutant, Differential Scanning Calorimetry (DSC) was performed. A peak maximum was considered as temperature of melting and area under the curve was used to determine enthalpy. Folding pathway was traced

according to the pattern of curve progression. DSC data showed overlapping of two transitions, which could be corresponding to an intermediate formed during unfolding process, and displayed the three-state reversible folding pathways of wild type and mutant. The observed T_m value for wild type was significantly higher than the mutant (wild type T_m 32.0 ± 2.5 °C, ΔH 440 ± 12 Kcal/mol; mutant T_m 27.0 ± 3.0 °C, ΔH 660 ± 15 kcal/mol) (Figure 5(C)). Furthermore, the observed intermediate species was reluctant to exist in case of mutant while predominantly shown by wild type. These results suggest that overall stability of ABRAXAS was different due to existence of Arg361Gln mutation, and undergo dissimilar folding pathway. Altogether, these findings suggest the relative existence of wild type in a more stable form. However, the type of intermediate species formed during unfolding transition, and its characterization is beyond the scope of study.

2.5. Binding interaction

ABRAXAS is a member of BRCA1-complex and involves in direct interaction with RAP80 and BRCA1 (Kim, Huang et al., 2007). ABRAXAS binds directly to RAP80

and BRCA1 BRCT, thus facilitate the recruitment of BRCA1-complex at the DNA damage site. In our study, we found that the secondary structural components of wild type and mutant remained same irrespective of the mutation. However, the three-dimensional folding pattern and unfolding pathway indicated less stability for the mutant protein. Since structural components and folding pathway monitored at the global level showed differences, there is a possibility that the binding interactions between ABRAXAS wild type or mutant with RAP80 would likely be affected. To test this assumption, His pull-down assay of native RAP80 (1-405) with His-tagged ABRAXAS wild type and mutant was performed. Interestingly, the binding of ABRAXAS wild type and RAP80 was found significantly higher compared to the mutant (Figure 6). This difference in binding profile suggested that structural alteration in mutant hinders the association between RAP80 and ABRAXAS.

ABRAXAS is the key member of BRCA1-complex and acts as a bridging molecule among various members. ABRAXAS expression was significantly correlated with lower chance of tumorigenesis in patients with advanced no small-cell lung cancer receiving firstline platinum-gemcitabine chemotherapy (Joerger et al., 2011). Knock-out studies of ABRAXAS result in defective recruitment of BRCA1-complex, and hence the DNA repair defect (Kim, Huang et al., 2007; Wang et al., 2007). Thus, it is a multifaceted molecule that plays an important role in BRCA1-mediated homologous recombination repair and cancer progression.

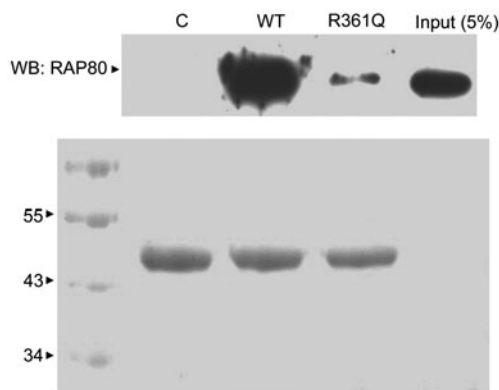


Figure 6. Binding analysis of ABRAXAS wild type and mutant with RAP80.

Notes: Histidine pull down assay followed by western blotting (Aygun, Svejstrup & Liu, 2008; Wang & Kirschner, 2013). Upper panel-ABRAXAS wild type or mutant was used as a bait and RAP80 as prey. RAP80 was probed with anti-RAP80 antibody. Heat denatured RAP80 was taken as control. Lower panel-Ponceau stained PVDF membrane showing the ABRAXAS His-tagged fusion protein as bait(s). Loading sequence is the same as in case of the upper panel. Wild type showed higher binding proportion compared to mutant.

In the present study, we have carried out a comprehensive examination of structural and functional properties of ABRAXAS, and their alterations due to mutation. The modeled structure of ABRAXAS wild type and mutant showed modest structural alteration. However, the observed structural changes in mutant do not alter its monomeric nature. The mutation containing region is a part of coiled-coil domain, which mainly involves in interaction with other BRCA1-complex members. Multiple sequence alignment of ABRAXAS considering various species in phylogenetic order showed highly conserved nature of Arg361 residue (Figure 1(B)) which signifies its importance in disease susceptibility. Wild type and mutant showed similar secondary structural composition, whereas the relative orientation of Trp and Tyr was slightly disturbed. This indicates that Arg361Gln mutation brings several localized changes in structural pattern of ABRAXAS, which altogether furnish a different conformational stability in a cumulative manner. Albeit, these conformation changes are very minor and could not be detected at the secondary structure level, but their relative positions were traced during global unfolding. The relative redundancy of intermediate species in case of wild type suggests the existence of different unfolding pathways which are partially followed by mutant as well as unfolding of wild type that was found to be more cooperative. Altogether, the modest localized changes in the mutant structure bring down its thermal and chemical stability which further perturbs the interaction with RAP80. The cumulative global changes in mutant structure were sufficient to disturb critical interaction necessary for BRCA1-complex integrity and localization. Therefore, in the presence of Arg361Gln mutation, ABRAXAS could not extend its bridging interaction(s) through RAP80, which perhaps prevent the recruitment of BRCA1-complex to the DNA damage site (Kim, Chen et al., 2007; Sobhian et al., 2007; Wang et al., 2007). Consequently, the nuclear retention of BRCA1 is adversely affected which further agitates G2/M checkpoint and homology-directed DNA repair (Figure 7(A) and (B)) (Solyom et al., 2012). To best of our knowledge, it is the first comprehensive report elucidating the structural and functional mechanism of mutational consequences and disease susceptibility. These findings would substantially list ABRAXAS as a new susceptibility gene to cancer predisposition. It also opens the vast perspective of considering the Arg361Gln mutation role in disease progression, such as cancer. It will further explore the opportunity of inhibitor design for therapeutic application that can recompense the effect of such deleterious mutation.

3. Materials and methods

High-quality molecular biology or analytical grade chemicals (>99.9%) were purchased from Sigma Aldrich

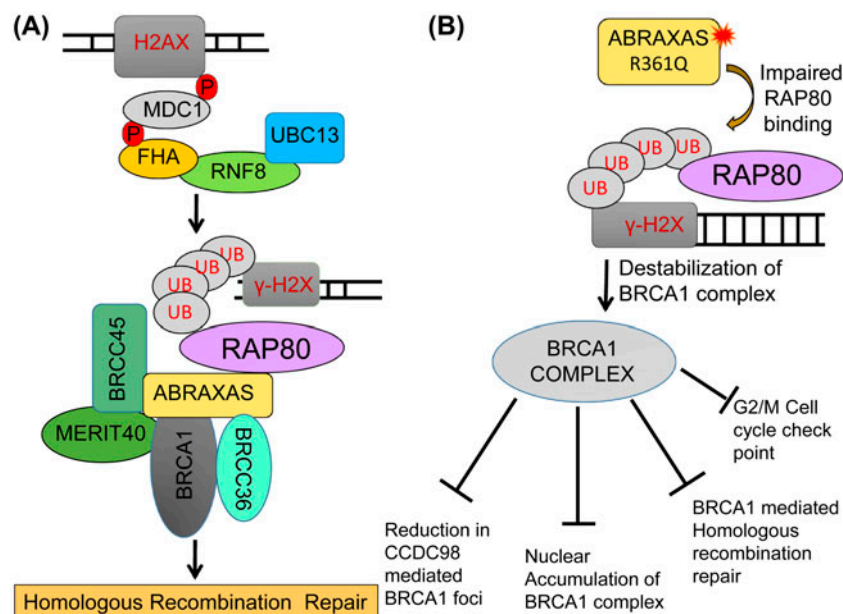


Figure 7. Proposed mechanism of consequence due to ABRAXAS Arg361Gln mutation.

Notes: (A) The figure illustrates the plausible mechanism of defective homologous recombination repair and other consequences due to Arg361Gln mutation. In case of wild-type, after the DNA damage, H2AX becomes phosphorylated by ATM/ATR kinase at nucleosome site which further recruits MDC1, RNF8, and UBC13. This assembly followed by formation of polyubiquitin chain(s) on histone(s) (H2AX) which is further recognized by tandem UIM motifs of RAP80. Since RAP80 is directly associated with ABRAXAS which in turn interacts with other BRCA1-complex members, thereby recruits the entire complex to the DNA damage site. (B) However, in case of Arg361Gln mutation, interaction between ABRAXAS and RAP80 is altered, which leads to defective recruitment of BRCA1-complex. Consequently, nuclear accumulation of BRCA1-complex significantly reduces which further leads to defective G2/M cell-cycle checkpoint, and an error prone homologous recombination repair.

(USA), unless specified. Restriction enzymes were purchased from Fermentas (USA).

3.1. Gene cloning, protein expression, and purification

PCR-amplified (Thermo cycler, Bio-Rad) gene product of ABRAXAS (6-373) (kind gift from J. Chen) was purified and digested with NcoI/XhoI. NcoI/XhoI digested vector pET-28a (Novagen) was ligated with PCR product at 1:3 molecular ratio and transformed into *E. coli* DH5 α . Positive clones were screened for insert release followed by confirmation with DNA sequencing. Site-directed mutagenesis was performed on wild-type template using mutagenic primers (Integrated DNA Technology) having c.1082G>A substitution. Amplicons were digested with DpnI (Fermentas, USA) and transformed into *E. coli* DH5 α bacterial strain. c.1082G>A substitution was confirmed using DNA sequencing. For recombinant protein expression, vector construct was transformed into *E. coli* Rosetta 2 (DE3) bacterial strain (Novagen). Single colony was inoculated and culture was grown till O.D₆₀₀ reached .6–.8, thereafter it was induced with IPTG and grown overnight at 25°C. Culture was harvested and pellet was re-suspended in

10 mM HEPES buffer containing 300 mM NaCl, 10 mM BME, .1 mM EDTA, pH 7.5 (HNBE buffer). Cells were disrupted by sonication (Branson Sonifier) and soluble fraction was incubated on pre-equilibrated resin (Ni-NTA in case of ABRAXAS, Novagen). Resin was washed thrice with HNBE buffer to remove non-specifically bound impurities. On bead cleavage was performed using TEV protease to remove affinity tag from native protein. Protein of interest was eluted gradientally with 100, 250, and 500 mM of imidazole in HNBE buffer pH 7.5. Protein was further purified on FPLC (AKTA) using size exclusion chromatography (Superdex 200, GE) to remove aggregates, etc. and analyzed on SDS-PAGE for purity.

3.2. Molecular modeling and docking

ABRAXAS (351-370) peptide was modeled using homology modeling by considering the template structure from the protein data bank (PDB ID: 3S4T). Authentic models were selected based on overall geometry and stereo-chemistry. The model was validated through protein structure validation server “SAVES” (Metaserver for analyzing and validating protein structures, <http://nihserv.mbi.ucla.edu/SAVES/>) (Ramachandran et al., 1963;

Ramachandran & Sasisekharan, 1968). Modeled structure was simulated using Schrodinger Desmond software. Wild type model was overlaid with the mutant to make a qualitative assumption about the structural changes.

3.3. Protein estimation

Purified protein was quantified with Bradford's protein estimation protocol as per manufacturer instructions (Expedon). Different dilutions of BSA were prepared as a standard reference, and absorbance was recorded using a spectrophotometer (Shimadzu) at O.D₅₉₅. Concentration of protein was determined by interpolation of a standard curve. Experiment was performed in triplicates (Bradford, 1976; Noble & Bailey, 2009).

3.4. Limited proteolysis

Wild type (1 mg/ml) and mutant (1 mg/ml) proteins were digested with trypsin (37 °C) and chymotrypsin (25 °C), such that the final concentration of proteases was 20 and 10 µg/µl in solution, respectively. Reaction was terminated at the different time point (0, 5, 10, 30, 60, 180, 240 min) by addition of 1 mM PMSF. Reaction product was mixed with equal volume of Laemmli buffer and heated prior to loading on SDS-PAGE. The experiment was repeated thrice by taking untreated wild type and mutant as controls (Havliš, Thomas, Šebela, & Shevchenko, 2003; Jimenez, Huang, Qiu, & Burlingame, 1998, chapter 16).

3.5. Fluorescence spectroscopy

Trp and Tyr residues' position in wild type and mutant was monitored using a Fluorescence spectrophotometer (Horiba, USA). Fluorescence emission scans were recorded from 310 to 400 nm wavelength range at a excitation wavelength of 280 nm, and the temperature 15 °C. For thermal denaturation, 2 µM protein was unfolded in a temperature range of 10–75 °C and emission spectra were acquired. Averaged blank corrected data were considered, and were analyzed by fitting in a three-state transition model to calculate thermal parameters (Pace, 1986).

For chemical denaturation, 10 M urea (Sigma–Aldrich) stock was prepared in 2.5 mM HEPES buffer containing 50 mM NaCl pH 7.5, and protein dilutions were prepared as described previously (Walters et al., 2009). Emission scans were collected and were further blank corrected. Thermodynamic parameters were obtained by curve fitting as per three-state transition model. The experiment was repeated three times, independently.

3.6. CD spectroscopy

CD polarimeter (Jasco J-810, Japan) was used to acquired Far-UV CD spectrum of purified proteins (10 µM) in a wavelength range 200–240 nm at 10 °C (buffer 2.5 mM HEPES pH 7.5, 50 mM NaCl). The average blank corrected data of three independent scans were considered, and Dichroweb server (<http://dichroweb.cryst.bbk.ac.uk>) (Lobley et al., 2002; Sreerama, Venyaminov, & Woody, 2000; Stephens, McKenna, McKenna, Nguyen & Devlin, 1981; Whitmore & Wallace, 2004, 2008) was used for estimation of secondary structure composition. For thermal denaturation, wild-type and mutant (10 µM) proteins were unfolded in a temperature range from 10 to 75 °C at 222 nm wavelength. Unfolding profile was obtained and curve fitting was performed. The experiment was replicated thrice.

3.7. Thermal and chemical denaturation data analysis and curve fitting

Thermal parameters were calculated by considering native (N) to intermediate (I) and intermediate to unfolded (U) transitions during unfolding process. Data analysis, curve fitting, and thermodynamic parameters determination were performed as described earlier (Pace & Shaw, 2000; Vikrant, Nakhwa et al., 2013).

3.8. Differential scanning calorimetry

Unfolding transition and enthalpy of reaction were determined using DSC (Setaram µDSC3 evo, USA). Samples were filtered and degassed prior to scanning. A total of 2 mg protein, wild type and mutant, was allowed to unfold from 10 to 75 °C temperature range with a temperature increment rate of 1 °C/minute. Data fitting was done using "CALISTO" software and enthalpy value was obtained. To determine the thermodynamic reversibility, samples were heated just above the transition maximum, cooled instantaneously, and reheated. The experiment was repeated thrice, independently.

3.9. Pull down assay

ABRAXAS (6-373) wild type and mutant were re-suspended in HNBEEG buffer (HNBE buffer containing 5% ethylene glycol and 5% glycerol) and lysed by sonication as mentioned above in section 1. Soluble His-tagged fusion protein ABRAXAS (6-373) wild type and mutant were allowed to bind on Ni-NTA resin (Novagen) and thereafter washed to remove impurities. Equal concentration of bound fusion protein (5 mg/ml) was used as bait and incubated with purified RAP80 (1-405) (prey protein, 50 µM concentration) for 5 h at 4 °C. Resin was washed with same buffer and loaded on SDS-PAGE. Protein complex was transferred over PVDF membrane

(Millipore) and probed with anti-RAP80 antibody (Abcam). Experiment was repeated thrice by taking heat denatured RAP80 (1-405) as control.

3.10. Dynamic light scattering

Molecular size measurement was done using a Malvern particle size analyzer (Zetasizer μ V). Samples were filtered (44 μ m) and degassed prior to scanning. Two milligram per milliliter of protein was scanned at a 5-min interval for 15 min, and effective diameter of peak of interest was considered. The experiment was repeated in three independent sets.

Supplementary material

The supplementary material for this paper is available online at <http://dx.doi.org/10.1080/07391102.2014.945484>.

Acknowledgements

We thank DBT-BTIS facility at ACTREC for providing necessary software to this study. Mr Vikrant thanks CSIR for fellowship (09/513 (0072)/2008-EMR-I).

Funding

Funding for this study was supported by DBT [grant number BT/PR10765/BRB/664/2008], [grant number BT/PR12565/BID/07/303/2009]; Seed in Air grant from TMC.

References

- Aygun, O., Svejstrup, J., & Liu, Y. (2008). A RECQ5-RNA polymerase II association identified by targeted proteomic analysis of human chromatin. *Proceedings of the National Academy of Sciences*, 105, 8580–8584.
- Badgujar, D. C., Sawant, U., Yadav, L., Hosur, M. V., & Varma, A. K. (2013). Preliminary crystallographic studies of BRCA1 BRCT-ABRAXAS complex. *Acta Crystallographica Section F Structural Biology and Crystallization Communications*, 69, 1401–1404.
- Bartek, J., & Lukas, J. (2007). DNA damage checkpoints: From initiation to recovery or adaptation. *Current Opinion in Cell Biology*, 19, 238–245.
- Bartek, J., Bartkova, J., & Lukas, J. (2007). DNA damage signalling guards against activated oncogenes and tumour progression. *Oncogene*, 26, 7773–7779.
- Bartkova, J., Horejsi, Z., Sehested, M., Nesland, J. M., Rajpert-De Meyts, E., Skakkebaek, N. E., ... Bartek, J. (2007). DNA damage response mediators MDC1 and 53BP1: Constitutive activation and aberrant loss in breast and lung cancer, but not in testicular germ cell tumours. *Oncogene*, 26, 7414–7422.
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72, 248–254.
- Burma, S., Chen, B. P., Murphy, M., Kurimasa, A., & Chen, D. J. (2001). ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *Journal of Biological Chemistry*, 276, 42462–42467.
- Celeste, A., Fernandez-Capetillo, O., Kruhlak, M. J., Pilch, D. R., Staudt, D. W., Lee, A., ... Nussenzweig, A. (2003). Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks. *Nature Cell Biology*, 5, 675–679.
- Coquelle, N., Green, R., & Glover, J. N. (2011). Impact of BRCA1 BRCT domain missense substitutions on phosphopeptide recognition. *Biochemistry*, 50, 4579–4589.
- Fontana, A., de Laureto, P. P., Spolaore, B., & Frare, E. (2012). Identifying disordered regions in proteins by limited proteolysis. *Methods in Molecular Biology*, 896, 297–318.
- Havliš, J., Thomas, H., Šebela, M., & Shevchenko, A. (2003). Fast-response proteomics by accelerated in-gel digestion of proteins. *Analytical Chemistry*, 75, 1300–1306.
- Huen, M. S., Grant, R., Manke, I., Minn, K., Yu, X., Yaffe, M. B., & Chen, J. (2007). RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly. *Cell*, 131, 901–914.
- Jimenez, C. R., Huang, L., Qiu, Y., & Burlingame, A. L. (1998). In-gel digestion of proteins for MALDI-MS fingerprint mapping. *Current Protocols in Protein Science*, Chapter 16, Unit 16.4, 16.4.1–16.4.5.
- Joerger, M., deJong, D., Burylo, A., Burgers, J. A., Baas, P., Huitema, A. D., ... Schellens, J. H. (2011). Tubulin, BRCA1, ERCC1, ABRAXAS, RAP80 mRNA expression, p53/p21 immunohistochemistry and clinical outcome in patients with advanced non small-cell lung cancer receiving first-line platinum-gemcitabine chemotherapy. *Lung Cancer*, 74, 310–317.
- Kim, D. E., Chivian, D., & Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Research*, 32, W526–531.
- Kim, H., Chen, J., & Yu, X. (2007). Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. *Science*, 316, 1202–1205.
- Kim, H., Huang, J., & Chen, J. (2007). CCDC98 is a BRCA1-BRCT domain-binding protein involved in the DNA damage response. *Nature Structural and Molecular Biology*, 14, 710–715.
- Kolas, N. K., Chapman, J. R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F. D., ... Durocher, D. (2007). Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science*, 318, 1637–1640.
- Kolchanov, N. A., Soloviov, V. V., & Zharkikh, A. A. (1983). The effects of mutations, deletions and insertions of single amino acids on the three-dimensional structure of globins. *FEBS Letters*, 161, 65–70.
- Leaver-Fay, A., Tyka, M., Lewis, S. M., Lange, O. F., Thompson, J., Jacak, R., ... Bradley, P. (2011). ROSETTA3: An object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymology*, 487, 545–574.
- Liu, Z., Wu, J., & Yu, X. (2007). CCDC98 targets BRCA1 to DNA damage sites. *Nature Structural and Molecular Biology*, 14, 716–720.
- Lobley, A., Whitmore, L., & Wallace, B. A. (2002). DICHRO-WEB: An interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics*, 18, 211–212.
- Lou, Z., Minter-Dykhouse, K., Wu, X., & Chen, J. (2003). MDC1 is coupled to activated CHK2 in mammalian DNA damage response pathways. *Nature*, 421, 957–961.
- Mailand, N., Bekker-Jensen, S., Fastrup, H., Melander, F., Bartek, J., Lukas, C., & Lukas, J. (2007). RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell*, 131, 887–900.
- Muñiz, V. A., Srinivasan, S., Boswell, S. A., Meinhold, D. W., Childs, T., Osuna, R., & Colón, W. (2011). The role of the

- local environment of engineered Tyr to Trp substitutions for probing the denaturation mechanism of FIS. *Protein Science*, 20, 302–312.
- Nikkilä, J., Coleman, K. A., Morrissey, D., Pylkäs, K., Erkkö, H., Messick, T. E., ... Greenberg, R. A. (2009). Familial breast cancer screening reveals an alteration in the RAP80 UIM domain that impairs DNA damage response function. *Oncogene*, 28, 1843–1852.
- Noble, J. E., & Bailey, M. J. (2009). Quantitation of protein. *Methods Enzymology*, 463, 73–95.
- Osorio, A., Barroso, A., García, M. J., Martínez-Delgado, B., Urioste, M., & Benítez, J. (2009). Evaluation of the BRCA1 interacting genes RAP80 and CCDC98 in familial breast cancer susceptibility. *Breast Cancer Research and Treatment*, 113, 371–376.
- Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymology*, 131, 266–280.
- Pace, C. N., & Shaw, K. L. (2000). Linear extrapolation method of analyzing solvent denaturation curves. *Proteins: Structure, Function, and Genetics*, 41, 1–7.
- Paull, T. T., Rogakou, E. P., Yamazaki, V., Kirchgessner, C. U., Gellert, M., & Bonner, W. M. (2000). A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Current Biology*, 10, 886–895.
- Pylkäs, K., Erkkö, H., Nikkilä, J., Sölyom, S., & Winqvist, R. (2008). Analysis of large deletions in BRCA1, BRCA2 and PALB2 genes in Finnish breast and ovarian cancer families. *BMC Cancer*, 8, 146.
- Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*, 7, 95–99.
- Ramachandran, G. N., & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Advances in Protein Chemistry*, 23, 283–437.
- Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S., & Bonner, W. M. (1998). DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *Journal of Biological Chemistry*, 273, 5858–5868.
- Rouse, J., & Jackson, S. P. (2002). Interfaces between the detection, signaling, and repair of DNA damage. *Science*, 297, 547–551.
- Sartori, A. A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., ... Jackson, S. P. (2007). Human CtIP promotes DNA end resection. *Nature*, 450, 509–514.
- Sato, Y., Yoshikawa, A., Mimura, H., Yamashita, M., Yamagata, A., & Fukai, S. (2009). Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by tandem UIMs of RAP80. *The EMBO Journal*, 28, 2461–2468.
- Shiotani, B., & Zou, L. (2009a). ATR signaling at a glance. *Journal of Cell Science*, 122, 301–304.
- Shiotani, B., & Zou, L. (2009b). Single-stranded DNA orchestrates an ATM-to-ATR switch at DNA breaks. *Molecular Cell*, 33, 547–558.
- Sobhanian, B., Shao, G., Lilli, D. R., Culhane, A. C., Moreau, L. A., Xia, B., ... Greenberg, R. A. (2007). RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science*, 316, 1198–1202.
- Sölyom, S., Aressy, B., Pylkäs, K., Patterson-Fortin, J., Hartikainen, J. M., Kallioniemi, A., ... Winqvist, R. (2012). Breast cancer-associated Abraxas mutation disrupts nuclear localization and DNA damage response functions. *Science Translational Medicine*, 4, 122ra123.
- Sreerama, N., Venyaminov, S. Y., & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: Inclusion of denatured proteins with native proteins in the analysis. *Analytical Biochemistry*, 287, 243–251.
- Stephens, P. J., McKenna, C. E., McKenna, M. C., Nguyen, H. T., & Devlin, F. (1981). Circular dichroism and magnetic circular dichroism of reduced molybdenum-iron protein of *Azotobacter vinelandii* nitrogenase. *Biochemistry*, 20, 2857–2864.
- Vikrant, Kumar, R., Yadav, L. R., Nakhwa, P., Waghmare, S. K., Goyal, P., & Varma, A. K. (2013). Structural and functional implication of RAP80 ΔGlu81 mutation. *PLoS One*, 8, e72707.
- Vikrant, Nakhwa, P., Badgujar, D. C., Kumar, R., Rathore, K. K., & Varma, A. K. (2013). Structural and functional characterization of the MERIT40 to understand its role in DNA repair. *Journal of Biomolecular Structure and Dynamics*, 1–16. doi:10.1080/07391102.2013.843473
- Vikrant, Sawant, U. U., & Varma, A. K. (2014). Role of MERIT40 in stabilization of BRCA1 complex: A protein–protein interaction study. *Biochemical and Biophysical Research Communications*, 446, 1139–1144.
- Walters, J., Milam, S. L., & Clark, A. C. (2009). Practical approaches to protein folding and assembly: Spectroscopic strategies in thermodynamics and kinetics. *Methods in Enzymology*, 455, 1–39.
- Wang, B., & Elledge, S. J. (2007). Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brcal/Brc36 complex in response to DNA damage. *Proceedings of the National Academy of Sciences*, 104, 20759–20763.
- Wang, B., Matsuoka, S., Ballif, B. A., Zhang, D., Smogorzewska, A., Gygi, S. P., & Elledge, S. J. (2007). Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science*, 316, 1194–1198.
- Wang, W., & Kirschner, M. W. (2013). Emi1 preferentially inhibits ubiquitin chain elongation by the anaphase-promoting complex. *Nature Cell Biology*, 15, 797–806.
- Wang, Y., Cortez, D., Yazdi, P., Neff, N., Elledge, S. J., & Qin, J. (2000). BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes and Development*, 14, 927–939.
- Whitmore, L., & Wallace, B. A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research*, 32, W668–W673.
- Whitmore, L., & Wallace, B. A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers*, 89, 392–400.
- Williams, R. S., & Glover, J. N. (2003). Structural consequences of a cancer-causing BRCA1-BRCT missense mutation. *Journal of Biological Chemistry*, 278, 2630–2635.
- Zarrine-Afsar, A., Larson, S. M., & Davidson, A. R. (2005). The family feud: Do proteins with similar structures fold via the same pathway? *Current Opinion in Structural Biology*, 15, 42–49.
- Zerovnik, E. (2011). Oligomerization preceding amyloid fibril formation: A process in common to intrinsically disordered and globular proteins. *Network*, 22, 154–161.
- Zhou, B. B., & Elledge, S. J. (2000). The DNA damage response: Putting checkpoints in perspective. *Nature*, 408, 433–439.
- Zhukov, I., Jaroszewski, L., & Bierzynski, A. (2000). Conservative mutation Met8 → Leu affects the folding process and structural stability of squash trypsin inhibitor CMTI-I. *Protein Science*, 9, 273–279.

Preliminary crystallographic studies of BRCA1 BRCT–ABRAXAS complex

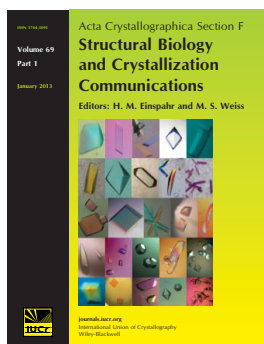
Dilip C. Badgujar, Ulka Sawant, Vikrant, Lumbini Yadav, M. V. Hosur and
Ashok K. Varma

Acta Cryst. (2013). **F69**, 1401–1404

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see <http://journals.iucr.org/services/authorrights.html>



Acta Crystallographica Section F: Structural Biology and Crystallization Communications is a rapid all-electronic journal, which provides a home for short communications on the crystallization and structure of biological macromolecules. Structures determined through structural genomics initiatives or from iterative studies such as those used in the pharmaceutical industry are particularly welcomed. Articles are available online when ready, making publication as fast as possible, and include unlimited free colour illustrations, movies and other enhancements. The editorial process is completely electronic with respect to deposition, submission, refereeing and publication.

Crystallography Journals **Online** is available from journals.iucr.org

**Dilip C. Badgajar, Ulka Sawant,
Vikrant, Lumbini Yadav, M. V.
Hosur and Ashok K. Varma***Advanced Centre for Treatment, Research and
Education in Cancer, Kharghar, Navi Mumbai
410 210, India

Correspondence e-mail: avarma@actrec.gov.in

Received 19 September 2013

Accepted 8 November 2013

Preliminary crystallographic studies of BRCA1 BRCT–ABRAXAS complex

The BRCA1 holoenzyme complex plays an important role in DNA damage repair. ABRAXAS is a newly discovered component of this complex and its C-terminal region directly binds to the BRCA1 BRCT domain. Single crystals of the BRCA1 BRCT–ABRAXAS complex grown by co-crystallization belonged to space group $P4_12_12$, with unit-cell parameters $a = b = 187.18$, $c = 85.31$ Å. Diffraction data were collected on the BM-14 beamline at the ESRF. Molecular-replacement calculations using *Phaser* led to three molecules in the asymmetric unit and a high solvent content of 76%.

1. Introduction

BRCA1 and BRCA2 are multifunctional proteins that are crucial for ensuring genome stability of human cells. Different domains of BRCA1 bind different interacting partners to perform a variety of functions including double-stranded DNA damage repair. Mutations in the BRCA1 and BRCA2 genes account for only 20% of familial breast and ovarian cancers, suggesting that the remainder arise for reasons that have still to be identified. Recent genetic analysis of cancer-patient families has identified *abraxas* as a new breast-cancer susceptibility gene and also as a novel target for therapeutic intervention (Solyom *et al.*, 2012). The protein ABRAXAS, recently discovered to be part of the BRCA1 DNA damage-repair complex, is one of the key adaptor proteins, which not only directly binds BRCA1 but also bridges interactions of BRCA1 with other interacting partners such as RAP80 and MERIT40. Through its interactions with RAP80, ABRAXAS is responsible for recruiting the BRCA1 complex to the DNA damage site (Shao *et al.*, 2009).

Cell lines containing the R361Q mutant of ABRAXAS were found to be hypersensitive to ionizing radiation. ABRAXAS is thus an essential factor in DNA damage resistance, repair and cell-cycle checkpoints (Eroles, 2013). ABRAXAS binds directly to the BRCA1 BRCT domain, which has been established to be a phosphopeptide-recognition module (Yu *et al.*, 2003; Wang *et al.*, 2007). The phosphopeptide-recognition sequence is pS- X_1 - X_2 -F (where pS is phosphorylated serine, X_1 is an aromatic amino acid, X_2 is a hydrophobic amino acid and F is phenylalanine; Yu *et al.*, 2003; Rodriguez *et al.*, 2003; Manke *et al.*, 2003). Two other proteins involved in DNA damage response and containing this signature motif for binding to BRCA1 BRCT are BACH1 and CtIP. However, ABRAXAS is unlike BACH1 and CtIP both in the location of the signature sequence and in the mechanism of localization at the DNA damage site. The signature sequence is located in the internal part of the sequences in BACH1 and CtIP, whereas it is at the carboxyl-terminus of ABRAXAS. Campbell and co-workers have studied the binding to BRCA1 and MDC1 of model tetrapeptides with and without free carboxyl ends. They observed that BRCA1 and MDC1 prefer oligopeptides with free carboxyl ends, with the preference being much greater in the case of MDC1. Through crystal structures of complexes, these authors have identified the environment around the conserved Arg1699 of BRCA1 as the factor determining the degree of specificity (Campbell *et al.*, 2010). In this context, the present study involving a longer oligopeptide with a free carboxy-terminus will provide information about the influence of additional amino acids on the interaction specificity.

© 2013 International Union of Crystallography
All rights reserved

Structural integrity of the BRCA1 BRCT domain is essential in performing the tumour suppressor function. The majority of breast and ovarian cancer-causing germline mutations in BRCA1 are located in the BRCA1 BRCT domain (Williams *et al.*, 2003). These mutations invariably disrupt phosphorylation-dependent interactions with cellular binding partners such as ABRAXAS, BACH1 and CtIP (Yu *et al.*, 2003; Manke *et al.*, 2003). ABRAXAS actually contains two potential ATM-phosphorylation serines, Ser404 and Ser406, at the C-terminus (Wang *et al.*, 2007). Through site-directed mutagenesis, phosphorylation at Ser406 alone was shown to be essential for binding to BRCA1 BRCT. Phosphorylation of Ser406 appears to be physiologically relevant, as the presence of phosphorylated Ser406 in cells exposed to ionizing radiation was confirmed using a phospho-specific antibody (Wang *et al.*, 2007). However, ABRAXAS phosphorylated at both the Ser404 and Ser406 residues was found to be selectively enriched on the incidence of ionizing radiation. To gain

insight into BRCT–phosphopeptide binding, we have undertaken biophysical and crystallographic studies of the interaction between the BRCA1 BRCT domain and differently (mono- and di-) phosphorylated C-terminal oligopeptides from ABRAXAS. Here, we report the crystallization of BRCA1 BRCT complexed with the ABRAXAS peptide phosphorylated at Ser406.

2. Materials and methods

2.1. Gene cloning of the BRCA1 BRCT domain

The BRCT domain (1646–1859) of BRCA1 was PCR amplified using the forward primer 5'-GTCGGATCCGAGAACCTGTACTTTCAGGGTGTGAACAAACGTATGTCCATG-3' and the reverse primer 5'-GTCGAATCCCTATTAGGGGATCTGGGGTATCATGTATGG-3'. The primers are designed in such a way as to accommodate a *Bam*HI restriction site and a TEV protease site (ENLYFQG) in the forward primer and an *Eco*RI site in the reverse primer. The PCR product of 687 base pairs and the empty vector pGEX-KT (a kind gift from Dr John Ladias) were digested by *Bam*HI and *Eco*RI and then ligated using the Quick Ligation Kit (New England Biolabs, USA). The potential clones were selected after screening for the presence of the insert and were finally confirmed by DNA sequencing.

2.2. Expression and purification of the BRCA1 BRCT domain protein

The BRCA1 BRCT protein was expressed in *Escherichia coli* bacterial strain BL21(DE3) using standard protocols. The BRCT domain fusion protein was purified from soluble cell-free lysate by affinity chromatography using glutathione Sepharose 4B resin (GE Healthcare). All purification steps were carried out at room temperature (~295 K). The GST fusion tag was removed by TEV protease and the reaction mixture was passed through a gel-filtration column (HiLoad Superdex 200) to obtain a monomeric population of BRCT. The TEV protease with a 6×His tag was removed by passing through a Ni²⁺–nitrotri-acetic acid (Ni–NTA) column. The purity of the BRCT domain was checked by resolving on 12% SDS–PAGE and the identity of the protein was confirmed by MALDI–TOF/TOF. The BRCT protein was concentrated to 25 mg ml^{−1} using a Centricon-10 (a centrifugal concentrator with 10 kDa molecular-mass cutoff) prior to crystallization.

2.3. Interaction analysis of BRCA1 BRCT with ABRAXAS

The oligopeptide of sequence (NH₂-G₃₉₉-F-G-E-Y-S-R-pS-P-T-F₄₀₉-COOH) from the C-terminus of ABRAXAS was obtained commercially (USV Biotech). The interaction between BRCA1 BRCT and ABRAXAS at 298 K was studied using an isothermal titration calorimeter (ITC-200; GE Healthcare). The protein and the peptide concentrations used were 0.020 and 0.200 mM, respectively. The ITC experiment was designed as a series of 18 injections, each of volume 2 µl, at time intervals of 210 s under constant stirring at 1000 rev min^{−1}. The heat of dilution was measured for subtraction during analysis of the exothermic data from the sample. The analysis software (*Origin 7.0*) provides values of the stoichiometry *N*, the binding constant *K_a* and the enthalpy change ΔH . All ITC experiments were repeated thrice and average values are reported.

2.4. Crystallization of the BRCA1 BRCT–ABRAXAS peptide complex

The ÄKTA FPLC-purified BRCA1 BRCT domain protein (25 mg ml^{−1}) was mixed with the ABRAXAS peptide (10 mg ml^{−1}) in a 1:1.5 molar ratio and incubated at 277 K overnight. Screening for

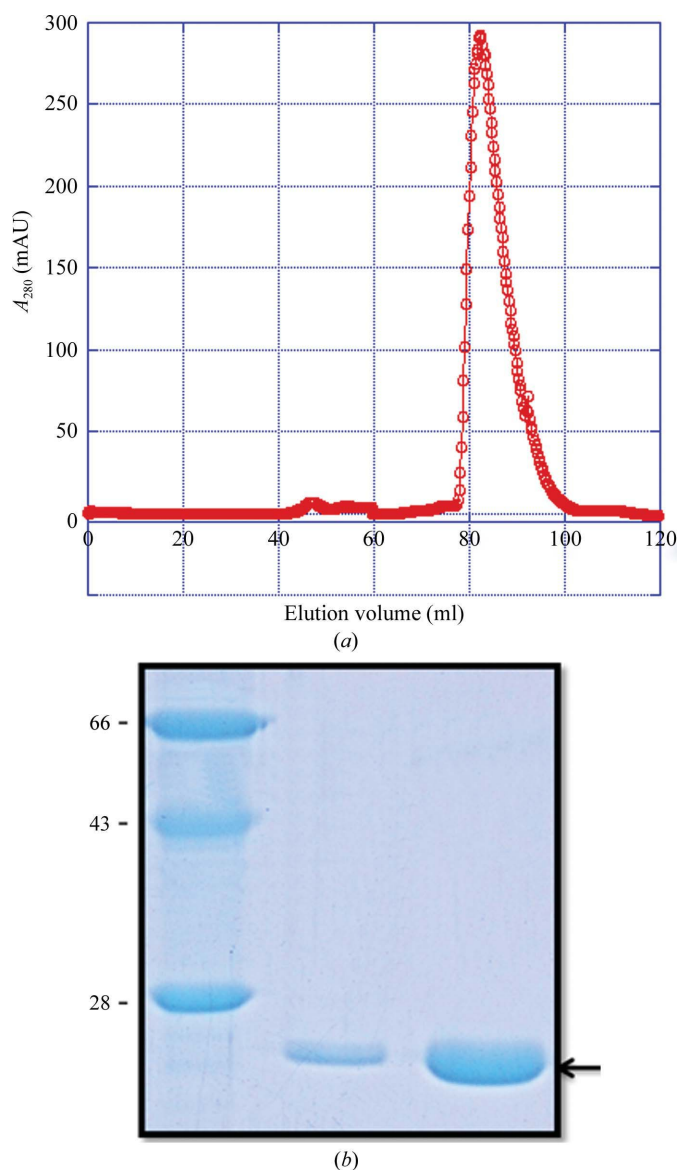


Figure 1
(a) The gel-filtration chromatogram of the BRCA1 BRCT protein eluted at the expected elution volume ~80 ml from a Superdex 200 column. (b) The gel-filtration fractions of purified BRCA1 BRCT protein (~25 kDa) resolved on a 12% SDS–PAGE gel.

Table 1

Crystal and diffraction data statistics.

Values in parentheses are for the outermost resolution shell.

No. of crystals	1
Beamline	BM-14, ESRF
Wavelength (Å)	0.9800
Temperature (K)	100
Crystal-to-detector distance (mm)	350
Rotation range per image (°)	1
Space group	$P4_12_12$ or $P4_32_12$
Unit-cell parameters (Å)	$a = b = 187.18$, $c = 85.31$
Mosaicity (°)	1.23
Resolution limit (Å)	76.47–3.8 (4.01–3.80)
Total No. of reflections	106886
Unique reflections	15491
Multiplicity	6.9
$\langle I/\sigma(I) \rangle$	9.5 (2.0)
Completeness (%)	99.9 (100)
$R_{\text{merge}}^{\dagger}$ (%)	9.1 (87.2)

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $\langle I(hkl) \rangle$ is the mean weighted intensity after rejection of outliers.

crystallization conditions was performed at 295 K by the hanging-drop vapour-diffusion method in 24-well crystallization plates using Crystal Screen and Crystal Screen 2 (Hampton Research). Drops prepared by mixing 2 μl protein solution with an equal volume of reservoir solution were equilibrated against 500 μl reservoir solution.

2.5. X-ray diffraction data collection

The X-ray diffraction data were collected at 100 K on the BM-14 beamline at the ESRF, France. A total of 150 frames, with a 1.0° oscillation of the crystal per frame, were collected at a crystal-to-detector distance of 350 mm and with an exposure time of 30 s.

3. Results and discussion

3.1. Protein preparation

Most of the BRCA1 BRCT domain expressed in *E. coli* accumulated in inclusion bodies. The typical yield from the soluble fraction was about 2.0 mg per litre of bacterial culture. The elution profile from the gel-filtration column (Fig. 1*a*) and the SDS-PAGE (Fig. 1*b*) established the monomeric status, purity and molecular mass of the purified BRCT. The molecular mass of 24 633.21 Da obtained from mass spectrometry was in agreement with expectations.

3.2. Interaction analysis with the ABRAXAS peptide

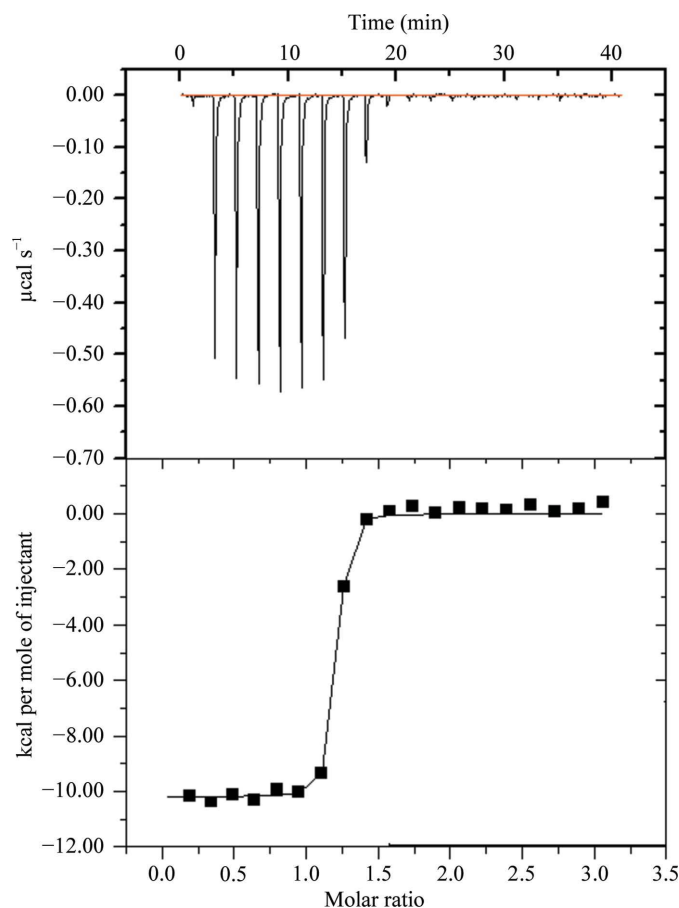
Fig. 2 shows the results of the calorimetric studies. From these observations, the binding affinity, ΔH and stoichiometry of the complex are estimated to be 1.2 μM , $-1.019 \times 10^4 \text{ cal mol}^{-1}$ and 1, respectively. The binding affinity of the ABRAXAS peptide to BRCT is lower than that of the BACH1 peptide (0.9 μM ; Shiozaki *et al.*, 2004), but higher than that of CtIP peptide (3.7 μM ; Varma *et al.*, 2005).

3.3. Crystallization

The BRCA1 BRCT–ABRAXAS peptide complex in 50 mM Tris pH 7.5, 300 mM NaCl crystallized in several conditions from Crystal Screen and Crystal Screen 2. The present diffraction analysis was carried out with crystals grown using 0.1 M MES monohydrate pH 6.5, 0.2 M ammonium sulfate, 30% polyethylene glycol monomethyl ether 5K. The crystal used for diffraction data collection was approximately $0.1 \times 0.08 \times 0.06 \text{ mm}$ in size (Fig. 3).

3.4. Data collection and processing

The crystals were soaked in a cryoprotectant solution obtained by augmenting the well solution with 30% glycerol. The crystals diffracted to 3.8 Å resolution at 100 K. The diffraction data were indexed and integrated using *iMosflm* (Battye *et al.*, 2011). Scaling

**Figure 2**

Representation of the ITC result obtained from the interaction of BRCA1 BRCT domain with the ABRAXAS peptide.

**Figure 3**

Single crystal of the BRCA1 BRCT–ABRAXAS complex.

and merging of the diffraction data were performed using the *CCP4* suite (Winn *et al.*, 2011). A summary of the data-collection and processing statistics is shown in Table 1. The crystal belonged to space group $P4_12_12$ or $P4_32_12$, with unit-cell parameters $a = b = 187.2$, $c = 85.3$ Å. Crystals of BRCT complexed with a variety of oligopeptides reported in the literature belonged to the hexagonal space group $P6_122$. This is the first complex of BRCA1 BRCT to be crystallized in a tetragonal space group. A V_M (Matthews, 1968) value of $2.53 \text{ Å}^3 \text{ Da}^{-1}$ would correspond to the presence of six BRCA1 BRCT molecules in the asymmetric unit with 51% solvent content. Structure solution using the molecular-replacement method is being attempted with PDB entry 1t15 (Clapperton *et al.*, 2004) as the search model. Initial molecular-replacement calculations using *Phaser* (McCoy *et al.*, 2007) suggest space group $P4_12_12$ and the presence of only three molecules in the asymmetric unit, with a solvent content of 76%. Crystals containing a large amount of solvent are generally not well ordered, and the low diffraction resolution observed here might indeed be a consequence of the large solvent content.

We are very grateful to the Indian Consortium for Macromolecular Crystallography, sponsored by the DBT, for giving us beamtime on the BM-14 beamline. DCB is very grateful to DAE and ACTREC for providing a fellowship. MVH thanks DAE for the award of a Raja Ramanna fellowship.

References

- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Campbell, S. J., Edwards, R. A. & Glover, J. N. M. (2010). *Structure*, **18**, 167–176.
- Clapperton, J. A., Manke, I. A., Lowery, D. M., Ho, T., Haire, L. F., Yaffe, M. B. & Smerdon, S. J. (2004). *Nature Struct. Mol. Biol.* **11**, 512–518.
- Eroles, P. (2013). *Transl. Cancer Res.* **2**, 103–105.
- Manke, I. A., Lowery, D. M., Nguyen, A. & Yaffe, M. B. (2003). *Science*, **302**, 636–639.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Rodriguez, M., Yu, X., Chen, J. & Songyang, Z. (2003). *J. Biol. Chem.* **278**, 52914–52918.
- Shao, G., Patterson-Fortin, J., Messick, T. E., Feng, D., Shanbhag, N., Wang, Y. & Greenberg, R. A. (2009). *Genes Dev.* **23**, 740–754.
- Shiozaki, E. N., Gu, L., Yan, N. & Shi, Y. (2004). *Mol. Cell*, **14**, 405–412.
- Solyom, S., Aressy, B., Pylkas, K., Patterson-Fortin, J., Hartikainen, J. M., Kallioniemi, A., Kauppila, S., Nikkila, J., Kosma, V.-M. & Mannermaa, A. (2012). *Sci. Transl. Med.* **4**, 122ra3.
- Varma, A. K., Brown, R. S., Birrane, G. & Ladas, J. A. (2005). *Biochemistry*, **44**, 10941–10946.
- Wang, B., Matsuoka, S., Ballif, B. A., Zhang, D., Smogorzewska, A., Gygi, S. P. & Elledge, S. J. (2007). *Science*, **316**, 1194–1198.
- Williams, R. S., Chasman, D. I., Hau, D. D., Hui, B., Lau, A. Y. & Glover, J. N. M. (2003). *J. Biol. Chem.* **278**, 53007–53016.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yu, X., Chini, C. C., He, M., Mer, G. & Chen, J. (2003). *Science*, **302**, 639–642.



Journal of Biomolecular Structure and Dynamics

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/tbsd20>

Conserved residues at the MAPKs binding interfaces that regulate transcriptional machinery

Bhanu P. Jagilinki^a, Nikhil Gadewal^a, Harshal Mehta^a, Hafiza Mahadik^a, Vikrant Pandey^a, Anamika^a, Ulka Sawant^a, Prasad A. Wadegaonkar^b, Peyush Goyal^c, Satish Kumar^d & Ashok K. Varma^a

^a Tata Memorial Centre, Advanced Centre for Treatment, Research and Education in Cancer, Kharghar, Navi Mumbai 410 210, Maharashtra, India

^b Department of Biotechnology, Sant Gadge Baba Amravati University, Amravati 444602, Maharashtra, India

^c Department of Biotechnology, Block - 2, 6th-8th floor, CGO Complex, Lodhi Road, New Delhi 110 003, India

^d Department of Biochemistry & Bioinformatics Centre, Mahatma Gandhi Institute of Medical Sciences, Sevagram (Wardha) 442102, India

Accepted author version posted online: 16 Apr 2014. Published online: 14 May 2014.

To cite this article: Bhanu P. Jagilinki, Nikhil Gadewal, Harshal Mehta, Hafiza Mahadik, Vikrant Pandey, Anamika, Ulka Sawant, Prasad A. Wadegaonkar, Peyush Goyal, Satish Kumar & Ashok K. Varma (2014): Conserved residues at the MAPKs binding interfaces that regulate transcriptional machinery, Journal of Biomolecular Structure and Dynamics, DOI: [10.1080/07391102.2014.915764](https://doi.org/10.1080/07391102.2014.915764)

To link to this article: <http://dx.doi.org/10.1080/07391102.2014.915764>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Conserved residues at the MAPKs binding interfaces that regulate transcriptional machinery

Bhanu P. Jagilinki^a, Nikhil Gadewal^a, Harshal Mehta^a, Hafiza Mahadik^a, Vikrant^a, Anamika Pandey^a, Ulka Sawant^a, Prasad A. Wadegaonkar^b, Peyush Goyal^c, Satish Kumar^d and Ashok K. Varma^{a*}

^aTata Memorial Centre, Advanced Centre for Treatment, Research and Education in Cancer, Kharghar, Navi Mumbai 410 210, Maharashtra, India; ^bDepartment of Biotechnology, Sant Gadge Baba Amravati University, Amravati 444602, Maharashtra, India; ^cDepartment of Biotechnology, Block – 2, 6th–8th floor, CGO Complex, Lodhi Road, New Delhi 110 003, India; ^dDepartment of Biochemistry & Bioinformatics Centre, Mahatma Gandhi Institute of Medical Sciences, Sevagram (Wardha) 442102, India

Communicated by Ramaswamy H. Sarma

(Received 25 July 2013; accepted 13 April 2014)

Signaling through c-Raf downstream pathways is the crucial subject of extensive studies because over expressed or mutated genes in this pathway lead to a variety of human cancers. On the basis of cellular localization, this pathway has been sub-divided into two cascades. The first RAF1-MEK1-ERK2 cascade which remains in the cytosol, whereas the second MEK1-ERK2-RSKs transduces into the nucleus and regulates the transactivation function. But how a few amino acids critically regulate the transcriptional function remains unclear. In this paper, we have performed *in silico* studies to unravel how atomic complexities at the MEK1-ERK2-RSKs pathways intercedes different functional responses. The secondary structure of the ERK, RSKs have been modeled using Jpred3, PSI-PHRED, protein modeler, and Integrated sequence analyzer from Discovery Studio software. Peptides of RSKs isozymes (RSK1/2/3/4) were built and docked on ERK2 structure using ZDOCK module. The hydropathy index for the RSKs molecules was determined using the KYTE–DOOLITTLE plot. The simulations of complex molecules were carried out using a CHARMM force field. The protein–protein interactions (PPIs) in different cascade of MAP kinase (MAPK) have been shown to be similar to those predicted *in vivo*. PPIs elucidate that the amino acids located at the conserved domains of MAPK pathways are responsible for transactivation functions.

Keywords: *In silico*; MAPKs; Protein–Protein Interactions; transactivation function; RSKs

Introduction

There are several members of the MAP kinase (MAPK) pathway, which feature Ser/Thr family of protein kinases (Hanks, Quinn, & Hunter, 1988; Roux & Blenis, 2004). Every kinase pathway is triggered by a specific activator (Ben-Levy et al., 1995). This activated kinase has long been studied for various responses like gene expression, cell cycle arrest, and cell proliferation (Gong et al., 2008; Igata et al., 2005; Mabuchi et al., 2004; Satomi & Nishino, 2009; Zhang & Kong, 2008). The well studied c-Raf downstream pathway leads to an attractive signaling mechanism which regulates cellular function such as cell growth and motility (Hilger, Scheulen, & Strumberg, 2002; Roux et al., 2007). The general layout of this pathway has been divided into two cascades, the first is RAF1-MEK1-ERK2 which remains in the cytosol, and the second MEK1-ERK2-RSKs is involved in the translocation of this complex into the nucleus. It is difficult to define the distinct function of both the interacting cascades, because the sequential activation of these kinases activates other kinases localized in different parts of cell such as the plasma membrane, cytosol, and nucleus. This

pathway also helps to characterize the directionality of signals in the cellular environment. Despite a large volume of *in vitro* and *in vivo* data available for understanding the signal transduction mechanism of Ras pathways, there are still many unanswered questions. For example, (1) how does RAF1 translocate from the plasma membrane and activate the downstream kinases? (2) how do RAF1-MEK1-ERK2 complexes retain themselves in the cytosol? (3) What are the critical amino acids at the MEK1-ERK2-RSKs cascade that translocate this complex to the nucleus? (4) Is there any cross-talk between with the kinases other than in this pathway? There are contradictory reports about the cytosolic retention of RAF1-MEK1-ERK2 and nuclear translocation of MEK1-ERK2-RSKs complexes. However, our study explores the functional mechanism underlying MEK1-ERK2-RSKs using protein–protein interactions (PPIs).

The MAPK family of kinases is well known for mediating cellular responses to growth signals. The signal is transduced through phosphorylation, which occurs from the upstream kinase to the downstream kinase (Shi et al., 2007). Ras is a member of the G protein family,

*Corresponding author. Email: avarma@actrec.gov.in

while Raf belongs to the MAP3 K family, MEK belongs to the MAP2 K family, and ERK belongs to the MAPK family (Channavajhala et al., 2003; Ellinger-Ziegelbauer, Kelly, & Siebenlist, 1999; He, Cai, Yang, Liu, & Wan, 2009). RSKs are constituents of the AGC subfamily in the human kinome (Nguyen, 2008). The functional motif in MAPK pathways can be divided into two types of domains – the first type is the phosphorylation induced activation domains and the second type are the binding complex induced functional domains. The phosphorylated domain is buried inside the three-dimensional structural packing whereas the binding domain is located on the surface of the protein (Figure 1(A)).

Activated c-Ras interacts with the c-Raf at the plasma membrane which undergoes through conformational changes. Amino acids critical for the activation of Raf have been identified earlier (Terai & Matsuda, 2005). The conformational changes on c-Raf induce its binding to MEK1 kinases. The c-Raf-MEK1 complex activates MEK1 and phosphorylates it. This activated MEK1 binds to its downstream kinase ERK2 and phosphorylates it. The phosphorylated ERK2 binds to RSKs which activates the transcriptional and translational machinery by means of phosphorylation.

The extracellular signal regulated kinase ERK-2 is a 42 kDa protein which transduces the signal through the activation of MEK1/2 (Tarrega et al., 2005). MEK1/2 is an upstream signal transducer for ERK2 while ERK2 is an upstream signal transducer for RSKs. The RSKs (p90 ribosomal S6 family of kinases) are located downstream of the Ras/mitogen-activated protein kinase (MAPK)

pathway (Poteet-Smith, Smith, Lannigan, Freed, & Sturgill, 1999). RSKs have two sequentially diverse and functionally distinct C-terminal (CTD) and N-terminal (NTD) protein kinase domains, with a linker at the interface, and tails of a few amino acids at the N and C-terminals. Sequentially, the N-terminal kinase domain bears a resemblance to the AGC family of kinases, whereas the C-terminal kinase domain resembles the calcium/calmodulin kinase family (Hanks et al., 1988; Jones, Erikson, Blenis, Maller, & Erikson, 1988). Functionally, the N-terminal kinase domain mediates substrate phosphorylation whereas C-terminal kinase domain is involved in autophosphorylation (Frödin & Gammeltoft, 1999; Poteet-Smith et al., 1999). The activation of RSKs is a complex process which involves multiple phosphorylations at various conserved amino acids (Dalby, Morrice, Caudwell, Avruch, & Cohen, 1998). These phosphorylation sites are highly conserved in RSK isoforms. The extended C-terminal kinase domains contain highly conserved phosphorylated serine, which is known to be one of the binding domains of ERK2. The fully activated P90RSK has four major phosphorylation sites (Ser 221, Ser 363, Ser 380, and Thr 573) and two additional sites (Thr 359 and Ser 732) with specific functions. The atomic association between different MAPKs pocket will help in finding the critical amino acids responsible for signal transduction and transactivation functions. An efficient signal transducer needs to avoid the cross-talk between the genes scaffold. Earlier data predicted that kinases in these pathways form robust complexes, but the molecular association of the proteins

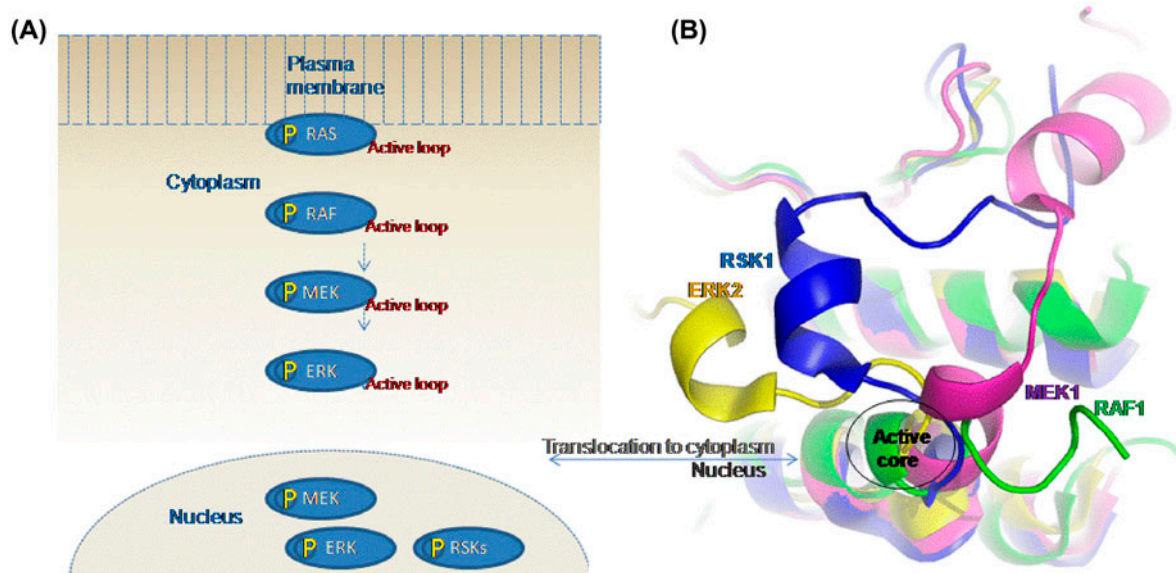


Figure 1. (A) The schematic representation of c-Raf downstream kinase pathway, the phosphorylation site (P) is buried inside the kinase domains and active binding domain is located at the molecular surface. (B) The superimposed structures of RAF1, MEK1, ERK2, and RSK which shows good structural similarity, however the bio-active core is conformationally diverse.

at the residual level is still implicit (Tian & Song, 2012). This paper provides structural insights due to conformational changes occurred in the pathways, and furthermore their effects on functional mechanism.

Results and discussion

It is well known that PPIs play an important role in coordinated signaling process of MAPK pathways. It has been reported that ERK2 has essential hydrophobic and charged amino acids that help to activate RSKs. But these sequences have not demonstrated a role to establish the signaling between the MAPKs pathways, which eventually leads to activation of the downstream protein kinase. Considering this fact, we looked at the upstream protein kinases in MAPK and found a docking site which is away from the enzymatic active site. The molecular *in silico* based approach has been applied to discover sequences of specific amino acid at the different MAPKs binding cascade which has not been evaluated extensively (Figure 1(B)).

MEK1-ERK2 binding interface

It has been reported that the MEK1-ERK2 binding interface is essential for the activation of RSKs. MEK1/2 in this pathway interact with ERK1/2 and transmit the signals to its downstream target. The N-terminus of MEK1 has mostly hydrophobic and basic residues which are essential for binding to ERK2 (Fukuda, Gotoh, & Nishida, 1997). Several lines of investigations have suggested that proper folded ERK2 dimer accumulate in the nucleus, whereas MEK1 associated ERK2 retained in the cytoplasm (Fukuda et al., 1997; Rubinfeld, Hanoch, & Seger, 1999; Xu, Stippec, Robinson, & Cobb, 2001). Hence, ERK2 is shuttling between cytoplasm and nucleus. To predict the atomic orientations of amino acids at the MEK1-ERK2 binding interface responsible for translocation of proper folded complex into nucleus, crystal structure of phospho-ERK2 which is found to be in twofold symmetry has been visualized. The sequence alignment of MEK1 (PDB ID: 3E8 N) and ERK2 (PDB ID: 1TVO) using Clustalw (Larkin et al., 2007) scores very less 18.0 while the overall three-dimensional folding and ATP binding site is similar except distortion at N and C-terminus. Few important domains of ERK2 have been characterized as binding domains of MEK1. The molecular docking between MEK1 and ERK2 has been performed on structurally determined molecule of ERK2 as well as symmetry related molecules (Figure 2(A) and (B)). It has been observed that ERK2 residues Tyr 316, Tyr 317, Asp 318, Asp 321, and Glu 322 are at the binding interface of MEK1 (Figure 2(C)) (Table 1, Supplementary material). However, MEK1 docked on the dimer structure of ERK2 is forming steric clashes (Figure 2(D))

which indicates the dimer form of ERK2 may not form complex with MEK1. This further elucidates that the MEK1 is not in a position to form a trimeric complex with ERK2 dimer and translocate into nucleus but rather remains in the cytosol. However, monomeric form of ERK2 is associated with the RSK1 and translocate to the nucleus.

Interactions between ERK2 and RSKs isoforms

The sequences of all RSK isoforms flanking the docking site were evaluated on the basis of secondary structure, hydrophobicity/hydrophilicity characteristics, flexibility, and accessibility of the sequences for ERK2 interactions. The homology model building, considering the appropriate sequence similarity (>30%) with the template, is critically important in analyzing the bioactive core of the protein structure. However, our initial search yielded ~30% sequence homology for the entire C-terminal tail region and significantly lower homology for the RSK CTD Docking sequences. To construct the sequences to be docked onto ERK2, a build protein tool integrated in the Discovery Studio2.1 was used. Due to lower sequence homology, an alternate method, Jpred3, was also tried for protein structure construction (Cole, Barber, & Barton, 2008). Jpred3 predicts secondary structure at an accuracy of 81.5%. Jpred3 derives alignment profiles and predicts secondary structure and solvent accessibility based on a Jnet algorithm. Both the programs observed similar predictions for the RSK isoforms docking sequences. The primary sequence was then modeled into a tertiary structure using *ab initio* based protein modeling. The RSK isoforms of 14 amino acid docking sequences were built into a model having minimum energy conformation. Since energy-based modeling works better for a smaller number of amino acids than a long peptide chain (Petsalaki, Stark, García-Urdiales, & Russell, 2009). The conserved docking sequence was predicted to be helical in structure and the flanking sequences were seen to be consisting of helices and coils. We have selected 14 amino acid sequences through observation of docking physicochemical attributes which includes conserved docking sequence as well as the flanking sequences. This selection has been analyzed using the Discovery Studio module for analyzing docked structures (Chen, Lyne, Giordanetto, Lovell, & Li, 2006). RSKs binding site has been found on the surface groove of ERK2, and far away (>20 Å) from the inhibitor binding sites, speculating inhibitor independent binding sites for each RSKs (Figure 3(A)). However, ERK2 residues from 323 to 329 responsible for nuclear translocation are located at the binding interface of RSK (Figure 3(B)) (Table 2), Supplementary material. It can be concluded here that the enzymatic site of ERK2 is buried inside of the three-dimensional structure, however

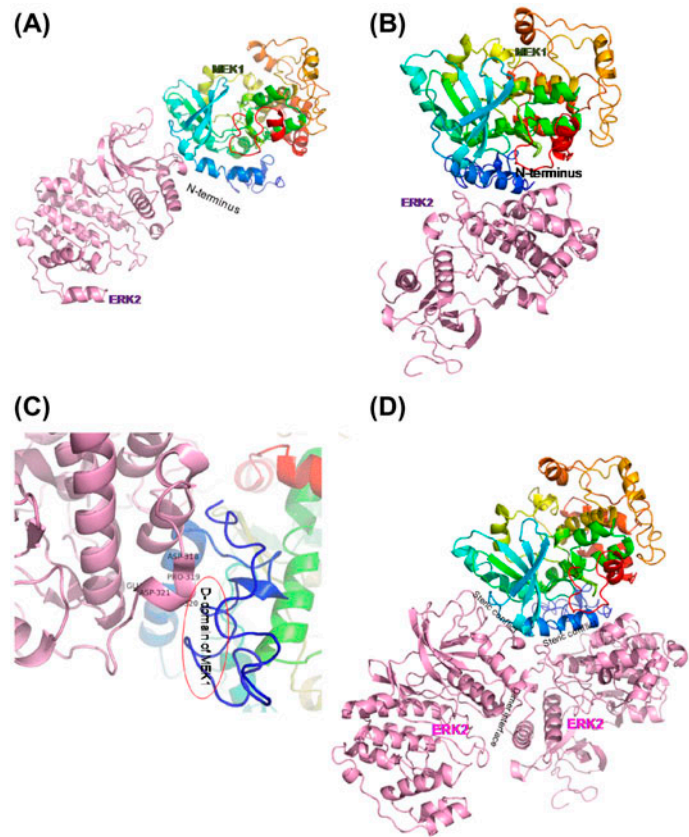


Figure 2. (A) Docking of MEK1 model on ERK2, N-terminus of MEK1 is forming complex with ERK2 molecule. (B) Docked structure of symmetry related molecules of ERK2 with MEK1. (C) The important residue at the docking interface of ERK2 is labeled and the docking domain of MEK1 is encircled. (D) There are steric clashes between the docked structure of p-ERK2 dimer and MEK1. ERK2 is colored in pink and MEK1 in rainbow.

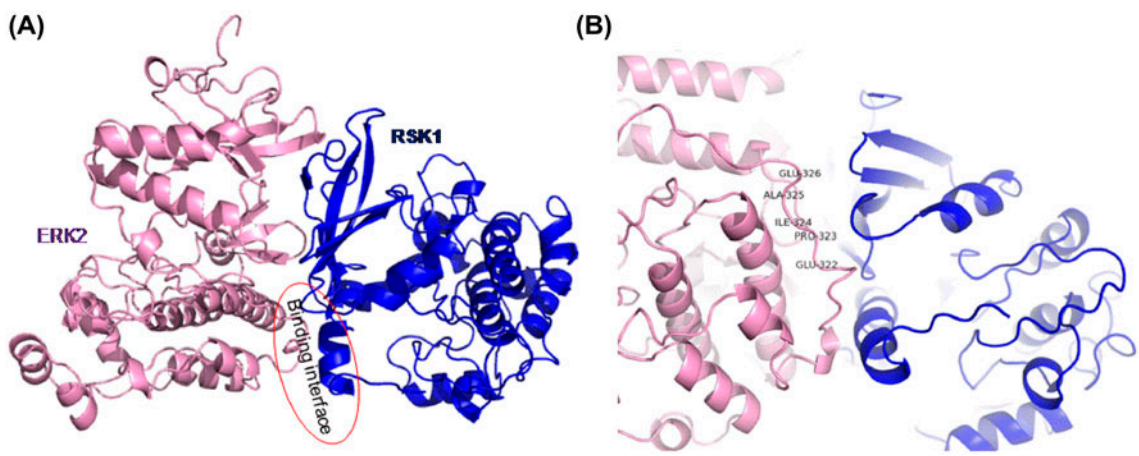


Figure 3. Complex model of ERK2-RSK1. (A) The ERK2 has a docking site at the C-terminus of RSK1. (B) The important residues at the binding interface has been shown.

binding domain being far from enzymatic sites helps in the translocation of this complex into the nucleus.

RSK isoforms at the C-terminal tail have conserved “LAQRR” amino acids. These conserved amino acids at

docking region, along with its flanking sequences, has been identified as interacting amino acids for the specific residue of an ERK2 molecule (Gavin & Nebreda, 1999; Smith, Poteet-Smith, Malarkey, & Sturgill, 1999). The sequences flanking around conserved “LAQRR” motif are variable in different RSKs isoforms which interacts with ERK2 with a specific docking orientation. These docking orientations determine the extent of association between the two molecules as well as variable functionalities. There are differences in the weak intermolecular interactions observed between ERK2-RSKs complexes (Figure 4(A)–(D)) (Tables 3–6, Supplementary material). It concludes that the differences of amino acids at the C-terminus and N-terminus of this conserved sequence are responsible for the change in the conformation and binding affinity. So within the limits of the docking software, to our conclusion the hydrophobic interactions and hydrogen bonding observed between ERK2-RSKs predict the folding behavior of complex molecules and also the directionality of signal transduction. It has been analyzed that the RSK1 has less binding than RSK2, RSK3, and RSK4. The *in silico* based results are in full agreement with what has already been observed *in vivo* (Smith et al., 1999).

RSK docking region is hydrophilic in nature

The RSKs isoforms sequences were tested for their hydrophobic/hydrophilic characteristics in order to predispose docking site for upstream kinase. For any residue to interact with a protein molecule, it has to be sufficiently exposed, i.e. accessible to the docking sequence of the approaching protein. The KYTE–DOOLITTLE (Kyte & Doolittle, 1982) plot in the ProtScale tool describes the sequences flanking at the end of the C-terminal of RSK isoforms to be particularly hydrophilic in nature. This is one of the rationale these sequences containing motifs are adopting the docking site at the surface of protein structures. The conserved residues within the carboxyl-terminal end of the RSK isoforms provides a specific ERK docking site (Biondi & Nebreda, 2003). The presence of an Arginine or a Lysine residue at the docking regions in all RSK isoforms raises the possibilities for such residues playing an important role in the binding of specific amino acid residues of ERK2 (Roux, Richards, & Blenis, 2003).

Transcriptional machinery in the nucleus at atomic level

Aspartic acid at positions 318 and 321, and residues from 323 to 329 in ERK2 makes prominent contacts with the MEK1 and RSKs, respectively. Therefore, it can be concluded that regions between 314 and 329 of ERK2 is a common docking domain for MEK1 and RSKs. The primary sequences of RSKs involved in

binding are characterized by a cluster of charged residues surrounded by hydrophobic residues (Figure 5(A)). The PPIs between MEK1-ERK2-RSKs cascade are very robust which make this complex very intact (Figure 5(B)). However, there is conflict in the results on translocation of the MEK1-ERK2-RSKs complex into the nucleus. A large amount of data predicts that the properly folded MEK1-ERK2-RSKs transduce into the nucleus (Skarpen et al., 2008; Smith et al., 1999) and MEK1 return back to cytosol leaving ERK2-RSKs in the nucleus, but how long the MEK1-ERK2-RSKs complex remains there in the nucleus is vague. The results presented in this paper unravel the molecular association between ERK2-RSKs that may help in predicting the amino acids responsible for cellular translocation. This difference in the position of the inhibitor site and the RSKs docking site anticipates the dual function of ERK2, first activation of ERK2 due to an enzymatic site present in the core of the structure and the second translocation of the ERK2-RSKs complex into the nucleus present at the protein surface. The RSKs are large isozymes which form complexes with different components of transactivation domain. Within the limit of the *in silico* docking, it has been that the amino acids Asp 318, Asp 321 are critical for docking between the ERK2-RSKs interface and may play a role in regulating the transcription function.

Conclusion

Comparing the PPIs between different cascades of MEK1-ERK2-RSK1/2/3/4 complexes, it has been observed that MEK1 and ERK2 are key molecules regulating the functional activity in the MAPK pathway. However, residues from 314 to 329 of ERK2 interact with MEK1 and RSKs in two pronged manner. Binding association between MEK1-ERK2-RSKs complex has revealed that there are steric hindrances between RSKs isozymes in the complex which in turn disassociates MEK1 from ERK2 and RSK complex. The differences observed in docking orientations for each isozyme of RSKs indicate that the sequence present at the N and C-terminus of conserved “LAQRR” sequence motif of RSKs plays a major role in predicting the binding affinity between ERK2-RSKs molecule. The PPIs between ERK2-RSKs molecules unravel the role of the ERK2 molecule for different RSKs isozymes, and eventually conclude the catalog of important differences among the closely related RSK family of kinases. In our studies we have showed the macromolecular docking and interactions of amino acids at atomic level. We also showed the importance of steric hindrance which plays a major role in dissociation of upstream kinases from the signaling complex in the MAPK pathway. The differences in binding affinities for various isozymes are essential for triggering a specific

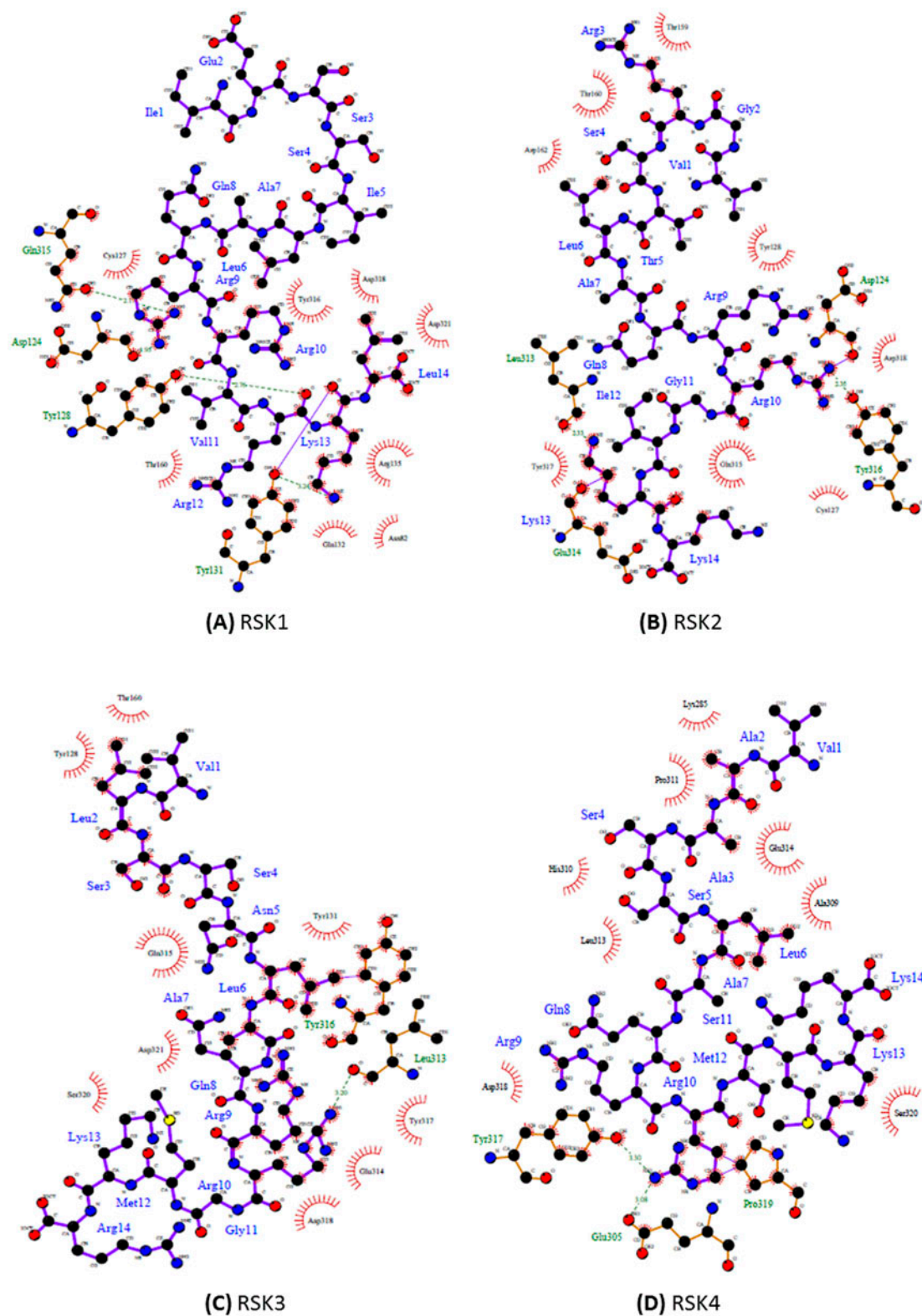


Figure 4. Binding association between ERK2 and C-terminus peptide of (A) RSK1, (B) RSK2, (C) RSK3, (D) RSK4.

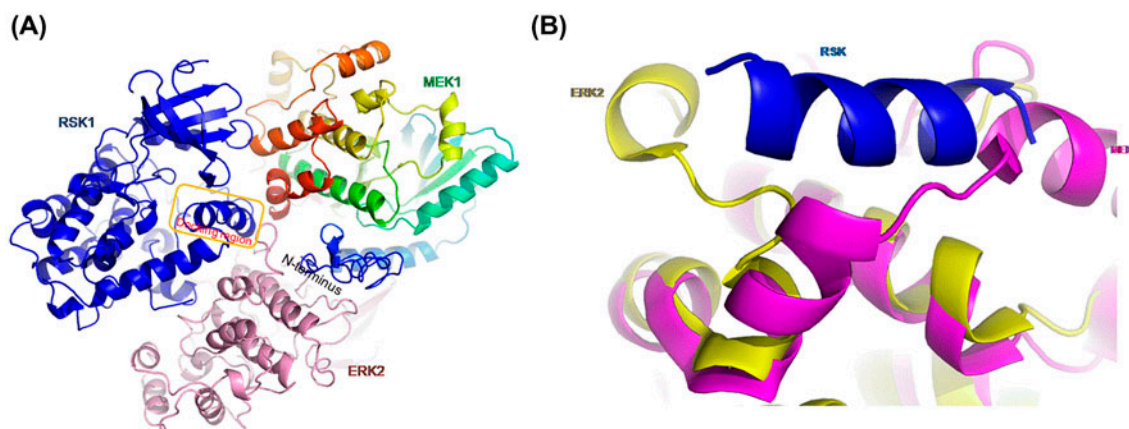


Figure 5. (A) Trimeric complex model of MEK1-ERK2-RSK1, N-terminus of MEK1 and C-terminus of RSK1 is interacting with interface of ERK2. The RSK1 is colored in blue. (B) N-terminus peptide motif of RSK1 docked to the model interface of ERK2-MEK1, it has been observed that peptide motif is also finding the same conformation what has been observed in *in vitro*.

type of transcription factors depending on the requirement of the cell. Atomic association between different MEK1-ERK2-RSKs complexes will be helpful in designing the small molecule inhibitors, which may regulate the structural basis of transactivation function.

Methodology

Model building of MEK1, ERK2, and RSK1

The MEK1 N-terminus contains the ERK docking site. Since the PDB structure of MEK1 (PDB ID: 3E8N) does not contain 60 residues at N-terminal, I-tasser (Zhang, 2008) server was used to model the full length structure. At C-terminal end of RSK1 (PDB ID: 2WNT) the ERK docking site was missing, therefore C-terminal domain of RSK1 was also modeled using the I-tasser server. The models were then evaluated using Procheck (Laskowski, Rullmannn, MacArthur, Kaptein, & Thornton, 1996) and Verify_3D server (Lüthy, Bowie, & Eisenberg, 1992). The missing loops of ERK2 (PDB ID: 1TVO) were modeled and minimized using the Prime tool of Maestro 9.2 (Schrodinger).

Docking of MEK1 and ERK2

Docking of MEK1 and ERK2 was performed using ZDock from Discovery Studio 2.5. The hydrophobic (⁹IQL¹¹) and basic (³KKK⁵) residues at N-terminal of MEK1 and acidic residues (Asp318 and Asp321) at C-terminal L16 segment of ERK2 are selected to filter the docking poses. In addition to the shape complementarity, the desolvation and electrostatic energy were used for ranking of the docked poses. Based on the docking scores and evaluation of binding interface residue the best docked pose was selected.

Docking of RSK1 to docked complex of MEK1 and ERK2

The docked complex of MEK1 and ERK2 was considered as a receptor and docked with RSK1. The poses were filtered by specifying the C-terminal residues (⁷²²LAQRRVRKL⁷³⁰) in RSK1 and acidic residues (Asp318 and Asp321) at C-terminal L16 segment of ERK2. Based on the docking scores the best docked pose of RSK1 with MEK1 and ERK2 was selected.

Modeling of RSK1/2/3/4 peptides

RSKs sequences (RSK1: IESSILAQRRVRKL, RSK2: VGRSTLAQRRGIKK, RSK3: VLSSNLAQRRGMKR, RSK4: VAASSLAQRRSMKK) subjected to BLAST (Müller, MacCallum, & Sternberg, 1999) in order to find a suitable template of known structures from PDB. PHYRE web-based program (Bennett-Lovsey, Herbert, Sternberg, & Kelley, 2008) was used for protein fold recognition using 1D and 3D sequence profiles coupled with secondary structure and solvation potential information. A build protein tool from Discovery studio 2.5.5 was used to construct the structures of RSK1/2/3/4 in alpha helical form. The side chains were then subjected to refinement using CHARMM force field (Brooks et al., 2009). The hydrophilic nature of RSK was calculated by KYTE-DOOLITTLE scale in ProtScale. Window size of 7 was kept and is good for finding hydrophilic regions that are likely exposed on the surface.

Docking of RSK1/2/3/4 peptides on ERK2

The Modeled RSKs peptides were docked to ERK2 molecule using the ZDock program keeping angular step size as 15. Out of the 2000 poses generated, only top poses that showed optimal hydrogen bonding were

selected. The molecular interactions between RSK1/2/3/4 and ERK2 were plotted using Ligplot tool (Laskowski & Swindells, 2011).

Supplementary material

The supplementary material for this paper is available online at <http://dx.doi.org/10.1080/07391102.2014.915764>.

Acknowledgments

We thank DBT-BTIS facility at ACTREC for providing necessary software for this study.

References

- Ben-Levy, R., Leighton, I. A., Doza, Y. N., Attwood, P., Morrice, N., Marshall, C. J., & Cohen, P. (1995). Identification of novel phosphorylation sites required for activation of MAPKAP kinase-2. *EMBO Journal*, 14, 5920–5930.
- Bennett-Lovsey, R. M., Herbert, A. D., Sternberg, M. J., & Kelley, L. A. (2008). Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins*, 70, 611–625.
- Biondi, R. M., & Nebreda, A. R. (2003). Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions. *Biochemical Journal*, 372, 1–13.
- Brooks, B. R., Brooks, C. L., 3rd, Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., ... Karplus, M. (2009). CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry*, 30, 1545–1614.
- Channavajhala, P. L., Wu, L., Cuozzo, J. W., Hall, J. P., Liu, W., Lin, L. L., & Zhang, Y. (2003). Identification of a novel human kinase supporter of Ras (hKSR-2) that functions as a negative regulator of Cot (Tpl2) signaling. *Journal of Biological Chemistry*, 278, 47089–47097.
- Chen, H., Lyne, P. D., Giordanetto, F., Lovell, T., & Li, J. (2006). On evaluating molecular-docking methods for pose prediction and enrichment factors. *Journal of Chemical Information and Modeling*, 46, 401–415.
- Cole, C., Barber, J. D., & Barton, G. J. (2008). The Jpred 3 secondary structure prediction server. *Nucleic Acids Research*, 36, W197–W201.
- Dalby, K. N., Morrice, N., Caudwell, F. B., Avruch, J., & Cohen, P. (1998). Identification of regulatory phosphorylation sites in mitogen-activated protein kinase (MAPK)-activated protein kinase-1a/p90rsk that are inducible by MAPK. *Journal of Biological Chemistry*, 273, 1496–1505.
- Ellinger-Ziegelbauer, H., Kelly, K., & Siebenlist, U. (1999). Cell cycle arrest and reversion of Ras-induced transformation by a conditionally activated form of mitogen-activated protein kinase kinase 3. *Molecular Cell Biology*, 19, 3857–3868.
- Frödin, M., & Gammeltoft, S. (1999). Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction. *Molecular and Cellular Endocrinology*, 151, 65–77.
- Fukuda, M., Gotoh, Y., & Nishida, E. (1997). Interaction of MAP kinase with MAP kinase kinase: Its possible role in the control of nucleocytoplasmic transport of MAP kinase. *The EMBO Journal*, 16, 1901–1908.
- Gavin, A. C., & Nebreda, A. R. (1999). A MAP kinase docking site is required for phosphorylation and activation of p90rsk/MAPKAP kinase-1. *Current Biology*, 9, 281–286.
- Gong, X. W., Wei, J., Li, Y. S., Cheng, W. W., Deng, P., & Jiang, Y. (2008). Effect of p38 mitogen-activated protein kinase gene knockout on cell proliferation of embryonic fibroblasts in mice. *Zhongguo Wei Zhong Bing Ji Jiu Yi Xue*, 20, 527–529.
- Hanks, S. K., Quinn, A. M., & Hunter, T. (1988). The protein kinase family: Conserved features and deduced phylogeny of the catalytic domains. *Science*, 241, 42–52.
- He, Y. Y., Cai, B., Yang, Y. X., Liu, X. L., & Wan, X. P. (2009). Estrogenic G protein-coupled receptor 30 signaling is involved in regulation of endometrial carcinoma by promoting proliferation, invasion potential, and interleukin-6 secretion via the MEK/ERK mitogen-activated protein kinase pathway. *Cancer Science*, 100, 1051–1061.
- Hilger, R. A., Scheulen, M. E., & Strumberg, D. (2002). The Ras-Raf-MEK-ERK pathway in the treatment of cancer. *Onkologie*, 25, 511–518.
- Igata, M., Motoshima, H., Tsuruzoe, K., Kojima, K., Matsumura, T., Kondo, T., ... Araki, E. (2005). Adenosine monophosphate-activated protein kinase suppresses vascular smooth muscle cell proliferation through the inhibition of cell cycle progression. *Circulation Research*, 97, 837–844.
- Jones, S. W., Erikson, E., Blenis, J., Maller, J. L., & Erikson, R. L. (1988). A Xenopus ribosomal protein S6 kinase has two apparent kinase domains that are each similar to distinct protein kinases. *Proceedings of the National Academy of Sciences USA*, 85, 3377–3381.
- Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157, 105–132.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., ... Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23, 2947–2948.
- Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR*, 8, 477–486.
- Laskowski, R. A., & Swindells, M. B. (2011). LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling*, 51, 2778–2786.
- Lüthy, R., Bowie, J. U., & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*, 356, 83–85.
- Mabuchi, S., Ohmichi, M., Kimura, A., Ikebuchi, Y., Hisamoto, K., Arimoto-Ishida, E., ... Murata, Y. (2004). Tamoxifen inhibits cell proliferation via mitogen-activated protein kinase cascades in human ovarian cancer cell lines in a manner not dependent on the expression of estrogen receptor or the sensitivity to cisplatin. *Endocrinology*, 145, 1302–1313.
- Müller, A., MacCallum, R. M., & Sternberg, M. J. (1999). Benchmarking PSI-BLAST in genome annotation. *Journal of Molecular Biology*, 293, 1257–1271.
- Nguyen, T. L. (2008). Targeting RSK: An overview of small molecule inhibitors. *Anti-Cancer Agents in Medicinal Chemistry*, 8, 710–716.
- Petsalaki, E., Stark, A., García-Urdiales, E., & Russell, R. B. (2009). Accurate prediction of peptide binding sites on protein surfaces. *PLoS Computational Biology*, 5, e1000335.
- Poteet-Smith, C. E., Smith, J. A., Lannigan, D. A., Freed, T. A., & Sturgill, T. W. (1999). Generation of constitutively active p90 ribosomal S6 kinase *in vivo*. Implications for the mitogen-activated protein kinase-activated protein

- kinase family. *Journal of Biological Chemistry*, 274, 22135–22138.
- Roux, P. P., & Blenis, J. (2004). ERK and p38 MAPK-activated protein kinases: A family of protein kinases with diverse biological functions. *Microbiology and Molecular Biology Reviews*, 68, 320–344.
- Roux, P. P., Richards, S. A., & Blenis, J. (2003). Phosphorylation of p90 ribosomal S6 kinase (RSK) regulates extracellular signal-regulated kinase docking and RSK activity. *Journal of Molecular Cell Biology*, 23, 4796–4804.
- Roux, P. P., Shahbazian, D., Vu, H., Holz, M. K., Cohen, M. S., Taunton, J., ... Blenis, J. (2007). RAS/ERK signaling promotes site-specific ribosomal protein S6 phosphorylation via RSK and stimulates cap-dependent translation. *Journal of Biological Chemistry*, 282, 14056–14064.
- Rubinfeld, H., Hanoch, T., & Seger, R. (1999). Identification of a cytoplasmic-retention sequence in ERK2. *Journal of Biological Chemistry*, 274, 30349–30352.
- Satomi, Y., & Nishino, H. (2009). Implication of mitogen-activated protein kinase in the induction of G1 cell cycle arrest and gadd45 expression by the carotenoid fucoxanthin in human cancer cells. *Biochimica et Biophysica Acta*, 1790, 260–266.
- Shi, F., Chiu, Y. J., Cho, Y., Bullard, T. A., Sokabe, M., & Fujiwara, K. (2007). Down-regulation of ERK but not MEK phosphorylation in cultured endothelial cells by repeated changes in cyclic stretch. *Cardiovascular Research*, 73, 813–822.
- Skarpen, E., Flinder, L. I., Rosseland, C. M., Orstavik, S., Wi-erod, L., Oksvold, M. P., ... Huitfeldt, H. S. (2008). MEK1 and MEK2 regulate distinct functions by sorting ERK2 to different intracellular compartments. *FASEB Journal*, 22, 466–476.
- Smith, J. A., Poteet-Smith, C. E., Malarkey, K., & Sturgill, T. W. (1999). Identification of an extracellular signal-regulated kinase (ERK) docking site in ribosomal S6 kinase, a sequence critical for activation by ERK *in vivo*. *Journal of Biological Chemistry*, 274, 2893–2898.
- Tarrega, C., Rios, P., Cejudo-Marin, R., Blanco-Aparicio, C., van den Berk, L., Schepens, J., ... Pulido, R. (2005). ERK2 shows a restrictive and locally selective mechanism of recognition by its tyrosine phosphatase inactivators not shared by its activator MEK1. *Journal of Biological Chemistry*, 280, 37885–37894.
- Terai, K., & Matsuda, M. (2005). Ras binding opens c-Raf to expose the docking site for mitogen-activated protein kinase kinase. *EMBO Reports*, 6, 251–255.
- Tian, T., & Song, J. (2012). Mathematical modelling of the MAP kinase pathway using proteomic datasets. *PLoS One*, 7, e42230.
- Xu, B., Stippec, S., Robinson, F. L., & Cobb, M. H. (2001). Hydrophobic as well as charged residues in both MEK1 and ERK2 are important for their proper docking. *Journal of Biological Chemistry*, 276, 26509–26515.
- Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, 9, 40.
- Zhang, Y. X., & Kong, C. Z. (2008). The role of mitogen-activated protein kinase cascades in inhibition of proliferation in human prostate carcinoma cells by raloxifene: An *in vitro* experiment. *Zhonghua Yi Xue Za Zhi*, 88, 271–275.

Over-expression, purification and isotopic labeling of a tag-less human glucose-dependent insulintropic polypeptide (hGIP)

Rakesh C. Chandarana · Vikrant · Ashok K. Varma ·
Anil Saran · Evans C. Coutinho · Jacinta S. D'Souza

Received: 10 July 2013 / Accepted: 8 October 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract Glucose-dependent insulintropic polypeptide (GIP), a gut peptide released in response to food intake brings about secretion of insulin in a glucose-dependent manner upon binding to its receptor, GIPR. GIP–GIPR has emerged as a new vista for anti-diabetic drug discovery and their interaction is being probed at the atomic level to aid rational drug design. In order to probe this interaction on cells, the current study attempts towards expressing ^{15}N -labeled GIP using classical molecular biology tools. We have developed a methodology to obtain GIP devoid of extra amino acid(s); a prerequisite to the intended interaction study. The synthetic GIP cDNA with a Factor Xa protease site at the N-terminus of GIP was inserted in the vector pET32a(+); the fusion protein thus expressed was eventually cleaved to obtain GIP. After successful Factor Xa cleavage, the cleaved GIP was confirmed by western blot. Subsequently, the (^{15}N)GIP was obtained using the aforementioned procedure and confirmed by MALDI-TOF.

Keywords Glucose-dependent insulintropic polypeptide · Affinity chromatography · Isotopic labeling · Recombinant fusion protein · Factor Xa protease cleavage site · Diabetes mellitus

R. C. Chandarana · A. Saran · E. C. Coutinho
Department of Pharmaceutical Chemistry, Bombay College of
Pharmacy, Kalina, Santacruz (E), Mumbai 400098, India

Vikrant · A. K. Varma
Structural and Molecular Biology Laboratory, Tata Memorial
Centre, Advanced Centre for Treatment, Research and Education
in Cancer, Kharghar, Navi Mumbai 410 210, India

J. S. D'Souza (✉)
UM-DAE Centre for Excellence in Basic Sciences, Kalina
Campus, Santacruz (E), Mumbai 400098, India
e-mail: jacinta@cbs.ac.in

Introduction

Glucose-dependent insulintropic polypeptide (GIP) is a peptide hormone released into the blood stream in response to glucose and nutrients absorption from the intestine along with glucagon like polypeptide-1 (GLP-1) (Drucker 2006; McIntosh et al. 2009). Both peptides are known to exert varied physiological actions on different body tissues with the pancreas being the major effector. These peptides interact with their respective G-protein-coupled receptors present on the cell surface and activate the adenylate cyclase signaling pathway. It brings about insulin secretion from the pancreas in a glucose-dependent manner and thus the name GIP. This pathway of insulin secretion is known as entero-insular axis and the peptides as incretins. Due to the glucose-dependent action, both GIP and GLP-1 have attracted immense attention for the design of novel anti-diabetics. In order to identify the key determinants of their interaction, attempts have been made to study their structures (Parthier et al. 2007; Runge et al. 2008; Underwood et al. 2010). In particular, studies aimed at probing the interaction of the peptides with the N-terminus domain of their receptors is of interest. Currently, nuclear magnetic resonance (NMR) is the only technique that permits determination of the structure of biomolecules at atomic-level resolution in near-physiological conditions (solution state) (Billeter et al. 2008; Banci et al. 2010). With isotopic labeling, determining complex structures and studying the interaction dynamics of biomolecules in the presence of other biomolecules and within the cell is now feasible (Stockman 2002; Takeuchi and Wagner 2006; Banci et al. 2010). However, much of the success of structural biology depends on the availability of biomolecules in its utmost pure and native state. Moreover, techniques such as X-ray crystallography and NMR require proteins in mM

concentration. With the advent of recombinant DNA technology (rDNA) and techniques of gene synthesis, affinity tag-based purification and subsequent removal of the tag by protease cleavage, it is now possible to obtain proteins in their native states with higher yields (Arnau et al. 2006).

Expression of isotopic hGIP had not been reported till date. The current study reports a method of isotopically (^{15}N)-labeled and unlabeled recombinant over-expression of hGIP in pET32a(+) vector, with complete tag removal by Factor Xa cleavage and purification of the native hGIP. The peptide obtained here has been characterized by western blotting using hGIP-specific antibody and the labeling has been confirmed by MALDI.

Materials and methods

All the materials were obtained from Sigma-Aldrich India Ltd., Merck-Millipore India, Himedia (India) and were of molecular biology grade.

Designing the DNA construct for *hgip* gene

To overcome the problem of codon bias and facilitate complete removal of the tags, the cDNA encoding hGIP was commercially gene synthesized (First Base, Singapore) with a nucleotide sequence coding for an Factor Xa protease cleavage at 5' and a stop codon at the 3' ends. The cDNA was cloned into pET32a(+) and subsequently transformed into *Escherichia coli* BL21(DE3) and *E. coli* DH5 α strains along with appropriate positive and negative controls. The colonies grown in Luria–Bertani (LB) agar plates containing ampicillin was checked for the presence of the insert by colony PCR and the amplicons were electrophoresed on a 2 % agarose gel containing ethidium bromide. Glycerol stocks of the positive clones were prepared and cryopreserved at $-80\text{ }^{\circ}\text{C}$. Plasmids were isolated from the cells containing the construct and sequenced using pET32a(+)-based primers.

Over-expression and purification of pET-GIP fusion protein

The *E. coli* BL21(DE3) cells harboring the construct were grown to an OD of 0.6 in 10 ml LB medium containing ampicillin and induced using 1 mM IPTG. In order to determine an optimum time of expression, induction was carried out for 1, 3 and 18 h. The cells collected at different time-points were harvested by centrifugation at 5,000g/4 $^{\circ}\text{C}$ and the expression of pET-GIP fusion protein was checked by electrophoresing on a denaturing gel. The solubility of the fusion protein was ascertained by

disrupting the cell lysate under non-denaturing conditions and electrophoresing the supernatant ($\sim 100\text{ }\mu\text{g}$ total protein/lane) on a denaturing gel. The pET-GIP fusion protein was confirmed using an antibody to the hexa-histidine portion of the tag. Upon observing over-expression of the 23 kDa pET-GIP polypeptide on SDS-PAGE, the conditions were scaled up to 1 l (culture volume) for 3 h and the cells harvested by centrifugation at 5,000g for 10 min at 4 $^{\circ}\text{C}$. The cell pellets were stored at $-80\text{ }^{\circ}\text{C}$ until the next step of purification. At the time of purification, the pellet was thawed and the cells were re-suspended in lysis buffer (20 mM Tris–HCl buffer pH 7.5, 300 mM NaCl, 1 mM PMSF, 10 % glycerol and 10 mM imidazole) and ruptured by sonication. The supernatant was collected after centrifuging the cell lysate at 10,000g for 30 min at 4 $^{\circ}\text{C}$ and mixed with Ni–NTA beads that were pre-washed with the lysis buffer, in a 50-ml centrifuge tube. The tube was kept on a cell mixer and the binding of the fusion protein to the beads was carried out for 2 h at 4 $^{\circ}\text{C}$. After binding, the slurry was loaded on a BioRad elution column and the flow-through was collected. The column was washed with lysis buffer containing 20 mM imidazole to remove the non-specific binding of proteins to the column. Elution of the pET-GIP fusion protein was then carried out using imidazole gradient from 20 mM to 2 M in the same buffer and the various fractions were electrophoresed on a denaturing gel. The purity of the fusion protein was assessed by silver staining of the gel. Subsequent purification of pET-GIP fusion protein was carried out by the batch elution method where the protein was eluted using 5 ml of lysis buffer containing 100 mM, 500 mM and 1 M imidazole. The fraction containing the protein was dialyzed against the Factor Xa digestion buffer –20 mM Tris–HCl pH 6.5, 50 mM NaCl, 1 mM CaCl_2 at 4 $^{\circ}\text{C}$.

Factor Xa cleavage of pET-GIP fusion protein

The cleavage of the purified pET-GIP fusion protein obtained in the Factor Xa digestion buffer after dialysis was optimized by varying parameters such as temperature and enzyme concentration. The cleaved protein was electrophoresed on a denaturing gel along with the uncleaved control and molecular weight marker and different fragments were observed. The fragments obtained upon cleavage were also characterized by western blot using hGIP-specific antibody to confirm the presence of the peptide.

Purification of GIP from the mixture by gel filtration

The 5-kDa hGIP was purified at room temperature by separating it from the mixture containing the residual tag (18 kDa), other contaminants ($>11\text{ kDa}$) and Factor Xa

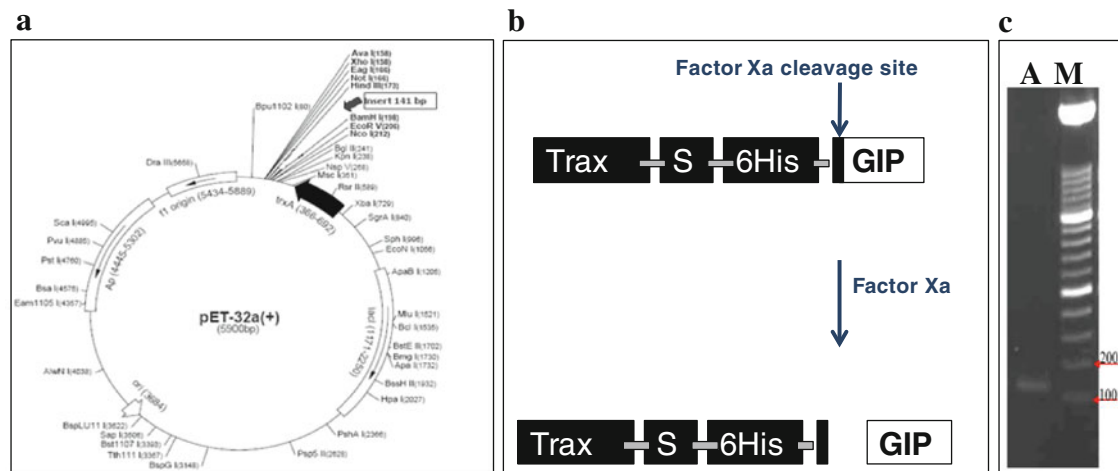


Fig. 1 Design of the construct resulting in the cloning of the GIP ORF. **a** pET32a(+) plasmid (obtained from First Base, Singapore) showing the position (between *Bam*HI and *Hind*III) at which the *hgp* insert was cloned. **b** A cartoon depiction of the Factor Xa cleavage

site. Note that upon cleavage the released hGIP is untagged. **c** PCR of the *hgp* amplicon of ~125 bp; lane labeled *M* is the standard molecular weight marker, lane labeled *A* is the *hgp* amplicon

enzyme (55 kDa) by gel filtration using a superdex 75 column on a GE AKTA FPLC system in 20 mM Tris–HCl buffer (pH 7) 150 mM NaCl. The elution was monitored by UV detector (280 nm) and elutes were collected as 2 ml fractions. The fractions showing the protein were electrophoresed on 15 % SDS-PAGE and the separation was checked.

Isotopic (^{15}N) labeling and characterization of hGIP by MALDI

Upon optimizing the basic strategies for over-expression, tag removal and purification of hGIP, isotopic labeling of the peptide was carried out by growing the *E. coli* BL21(DE3) cells up to OD₆₀₀ of 0.6 in the minimal medium (M9) containing isotopic (^{15}N) ammonium chloride as the nitrogen source and induced using 1 mM IPTG. Subsequent steps of purification were carried out in the manner as reported for unlabeled hGIP.

The isotopic peptide (0.5 μl corresponding to 0.25 μg concentration) was mixed with the non-isotopic hGIP (0.5 μl corresponding to 0.4 μg concentration) in 1 μl of Sinapinic acid as matrix and spotted on a ground steel MALDI-TOF plate (Bruker Daltonics, USA). This mixture was allowed to dry at room temperature and the plate was then loaded into the instrument. Subsequent to adjusting the appropriate parameters, the spot was fired using the linear mode at 22 ± 2 °C, 27 kV and a microchannel plate detector. The molecular weight was determined using an external standard (insulin, ubiquitin i, cytochrome *c*, myoglobin that covered a mass range of ~5,000–17,500 Da) at an error of ± 100 ppm, the software for deconvolution being Flex analysis, Bruker Daltonics.

Results and discussion

The need to study the structure of biomolecules at near-physiological conditions is mandatory for rational drug design; hGIP being one of them. For this purpose, cellular and molecular biologists as well as pharmacologists have felt the need to obtain decent concentrations and yields of hGIP in their studies. For long, such experimentalists have been using either the synthetic peptide or purifying it in bulk from natural sources (Pederson and Brown 1976; Alaña et al. 2004). Besides, acquiring (^{15}N)hGIP or doubly labeled (^{15}N , ^{13}C)hGIP synthetically for structural and/or receptor-binding studies has proven to be cost intensive. The aim of the present study was to obtain an economic and constant source of unlabeled and labeled hGIP in reasonable yields using recombinant DNA technology. The pET series of vectors have T7 promoters that over-express proteins in *E. coli* BL21(DE3) cells (Studier and Moffatt 1986). Apart from the hexa-histidine affinity tag, pET32a(+) also has a localization tag (S-tag) and a solubilization tag (ThioredoxinA, TrxA) (Tsunoda et al. 2005). To prevent the possibility of the protein routing into inclusion bodies, the idea of tagging it with a TrxA tag was sought after. The S-tag would aid in future receptor-binding studies under physiological conditions. The current study has directed efforts towards obtaining a constant source of unlabeled as well as labeled hGIP. For the ultimate need of obtaining a tag-less GIP, a construct was designed suitably using pET32a(+) as the backbone vector. The ThioredoxinA (TrxA) tag was followed by the Streptavidin (S) tag, the hexa-histidine (6XHis) tag and a Factor Xa cleavage site (IEGR). The GIP sequence was

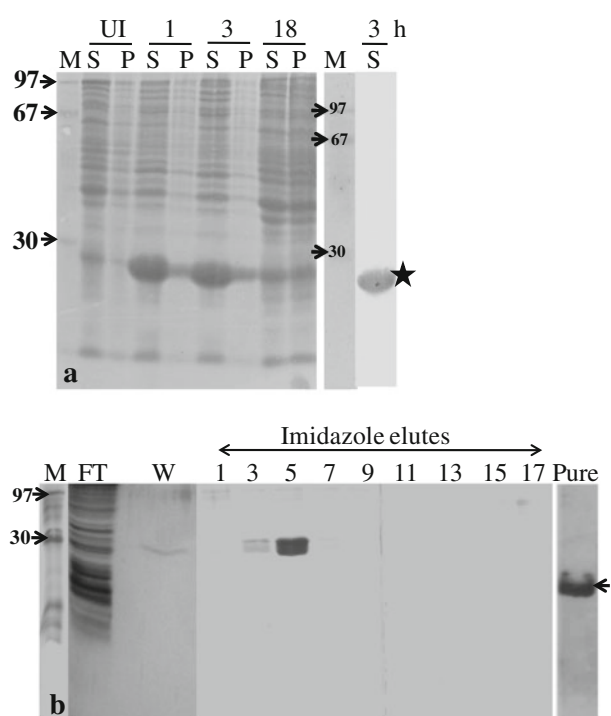


Fig. 2 Induction of the fusion protein (Trx-S-6XHis-fXa-GIP). **a** Kinetics of the over-expression of the fusion protein, induction with IPTG and detection of the 6XHis tag using an anti-hexa-histidine antibody (denoted by a star in the last lane of this figure). **b** Affinity (IMaC) based purification of the fusion protein using gradient elution (20–2,000 mM imidazole). Note the purification of the fusion protein in elutes 3 and 5. A silver stained gel of the purified protein (see arrowhead). Lanes labeled *M* protein molecular weight standards, *P* pellets, *S* supernatants obtained after extraction, *UI* uninduced control, *numbers* time-points (in hours) of induction, *FT* flow-through, *W* wash obtained after the proteins have been bound to the column; *numbers* against arrows, molecular weights in kDa; *pure*, purified protein (~5 µg)

introduced after this cleavage site so that the tag-less GIP alone is released upon Factor Xa cleavage. The synthetic sequence (Fig. 1a, b) so designed was then synthesized commercially by First Base, Singapore. The resultant recombinant vector (pET32a-fXa-gip) was used to transform *E. coli* BL21(DE3) competent cells so as to yield colonies on LB agar plates containing ampicillin as the antibiotic of selection. Subsequent characterization by colony PCR using insert-specific and vector-specific primers showed desired amplicons on 2 % agarose gel indicating the presence of the appropriate insert in all the colonies (Fig. 1c). The sequencing of the plasmids isolated from three such positive clones confirmed the presence of the correct sequence of the insert.

Over-expression and purification of pET-GIP fusion protein

Escherichia coli BL21(DE3) strains containing the plasmid (pET32a-fXa-gip) were induced with 1 mM IPTG. An intense band of fusion protein (Trx-S-6XHis-fXa-GIP) was observed at ~23 kDa; the intensity of this band increased with the duration of induction (Fig. 2a). The optimum time for induction was found to be 3 h, beyond which the protein degraded. As expected, due entirely to the TrxA tag, the fusion protein was found to be soluble as it was released into the aqueous medium by non-denaturing cell disruption (Fig. 2a; S for supernatant fraction that contained the fusion protein and not in the pellet P). The presence of the 6XHis tag, as confirmed using an anti-6XHis antibody on a western blot of the fusion protein (Fig. 2a; last lane) permitted affinity-based purification by IMaC using Ni-NTA matrix.

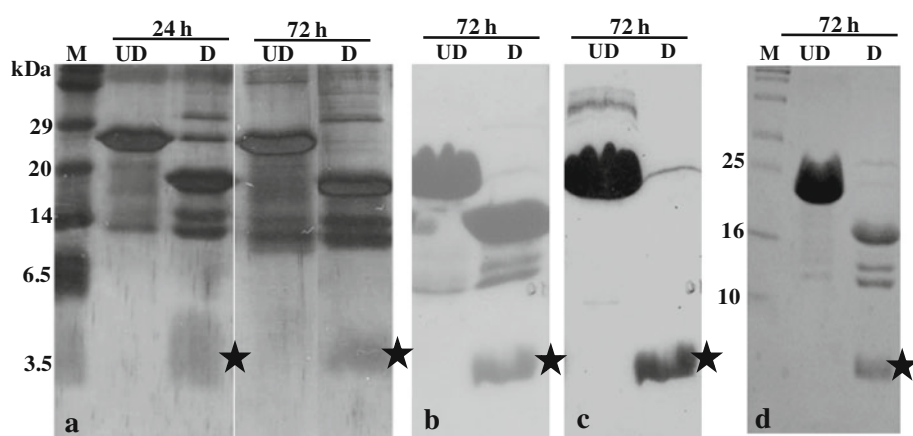


Fig. 3 Factor Xa digestion and characterization by western blotting. **a** The fusion protein (75 µg) was digested (lanes labeled *D*) with Factor Xa and the digested products were electrophoresed on a denaturing gel and silver stained; **b** Ponceau S staining of the nitrocellulose membrane after transferring the digested products (starting amount used was 200 µg) by the western blotting procedure;

c immunoblot for the detection of unlabeled hGIP (starting amount used was 150 µg), and **d** labeled hGIP, using anti-hGIP antibodies. *UD* and *D* labels in all the figures refer to undigested and digested, respectively. The stars in all the figures show the position of the digested product, viz. hGIP

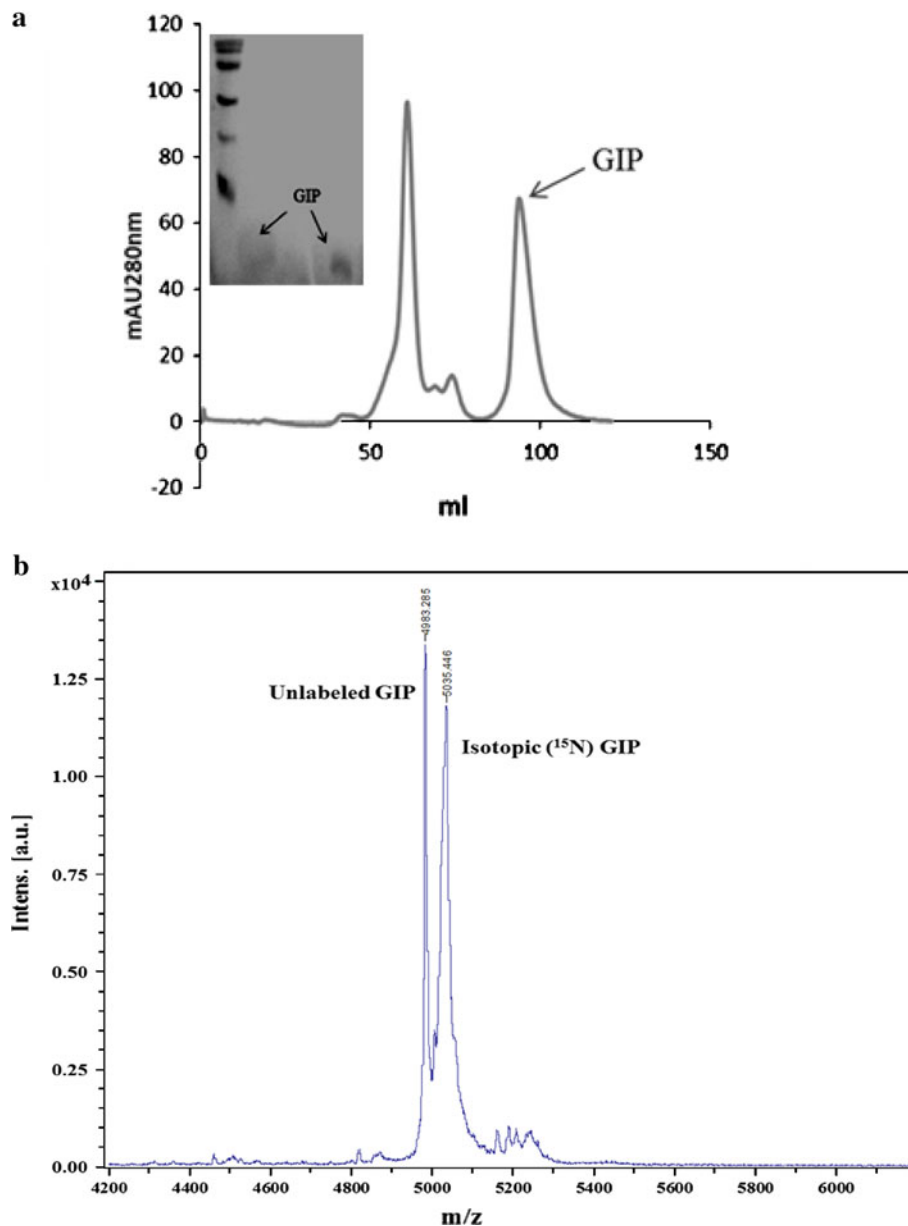


Fig. 4 **a** Isotopically labeled fusion protein was digested with Factor Xa and the GIP was purified using gel filtration chromatography on Superdex 75 column on an FPLC system (GE, AKTA). The *inset* is

the purified protein that was obtained after digestion. **b** Molecular weight determination of unlabeled and isotopic (^{15}N)GIP by MALDI of unlabeled and isotopic (^{15}N)GIP

The purification of pET-hGIP fusion protein carried out using Ni-NTA matrix and an imidazole gradient (20 mM–2 M) showed the presence of pure 23 kDa fusion protein in the elutes when electrophoresed on 15 % SDS-PAGE (Fig. 2b). The protein eluted in the fraction corresponding to ~100–150 mM imidazole and subsequent batch purification of the ^{15}N -labeled fusion protein at 100, 250 and 1,000 mM imidazole resulted in elution of maximum protein at 100 mM (Fig. 2b). The silver stained gel showed ~99 % purity and absence of any contaminants. The yield of the fusion protein as estimated by Bradford was found to be ~30 mg/l culture volume.

Factor Xa cleavage of the fusion protein, purification by gel filtration and MALDI-TOF analysis

After obtaining the purified fusion protein, the conditions for release of the GIP peptide after Factor Xa cleavage were explored. The digestion with Factor Xa was found to be optimum at pH 7.5, 4 °C, enzyme concentration of 1 unit of protease and digestion time being 72 h (Fig. 3a). Upon electrophoreses of the cleaved fragments on a 15 % SDS-PAGE, distinct bands of M_r ~18 kDa (corresponding to the residual tags) and ~5 kDa (corresponding to hGIP; Fig. 3a) were observed. Another doublet of M_r ~14 kDa

was also seen and possibly is a result of degradation of the fusion protein. In order to ascertain the digested product as hGIP (~5 kDa), a western blot of the digested products was probed with the anti-hGIP antibody (Fig. 3b, c; hGIP is marked with a star). This confirmed that the band was of hGIP and also permitted for further purification of the peptide.

The over-expression of the pET-hGIP fusion protein was reproducible when the cells harboring the construct were cultured in minimal media (M9) containing isotopic (^{15}N) ammonium chloride and induced using 1 mM IPTG. Subsequently, the purification of pET-GIP fusion protein was achieved successfully by affinity chromatography on a Ni-NTA column. Although the labeled protein eluted at a higher imidazole concentration (250 mM imidazole), the conditions of digestion with Factor Xa were reproducible. Further purification of hGIP was evident from the chromatogram where different fragments of the digested pET-GIP protein separated differentially by gel filtration on a Superdex 75 column, the tag and the doublet eluted in the 22nd–38th fraction (44–74 ml) and the hGIP at 41st–54th fraction (82–108 ml) showing a peak in the 47th fraction (94 ml). When the fraction corresponding to the hGIP peak was loaded on 15 % SDS-PAGE it showed the hGIP fragment in purified form (inset of Fig. 4a). The peptide was concentrated by ultrafiltration using amicon column mwco 3 kDa. The characterization of the labeled hGIP by MALDI showed the desired peak precisely at 4.98 kDa corresponding to the unlabeled peptide and also a peak at 5.44 kDa corresponding to the molecular weight of the ^{15}N labeled peptide. This confirmed the isotopic labeling of the hGIP. The yield of the hGIP peptide was ~1 mg/l of the culture.

It may be noted that GIP (full-length) or truncated versions have thus far been produced synthetically (Fehmann and Göke 1995). Such synthetic peptides have been used for various binding as well as physiological studies. However, these have been useful in proton-NMR studies and thus far no such method has been developed to obtain isotopic (^{15}N)hGIP. Hence, this is the first report of ^{15}N -labeled hGIP and we present a clone that harbors a fusion gene containing the *hgip* gene that can be reproducibly used to produce the fusion protein at ~30 mg/l and upon digestion and further purification, a peptide of almost 99 % purity with a yield of ~1 mg/l.

Acknowledgments R.C.C. and E.C. wishes to thank the Department of Science and Technology, New Delhi, India, for the grant SR/SO/BB-10/2005 dtd. 8-9-2006. A.S. wishes to thank the Indian National Science Academy, India, for financial support. J.S.D. wishes to thank the Department of Atomic Energy, India for financial support. E.C, J.D and R.C.C wish to thank Prof. B. J. Rao, Department of Biological Sciences, Tata Institute of Fundamental Research, India

for permitting the use of his laboratory during the initial part of this work.

Conflict of interest None.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Alañia I, Hewage CM, Malthouse JPG et al (2004) NMR structure of the glucose-dependent insulinotropic polypeptide fragment, GIP(1–30)amide. *Biochem Biophys Res Commun* 325:281–286. doi:[10.1016/j.bbrc.2004.10.033](https://doi.org/10.1016/j.bbrc.2004.10.033)
- Arnau J, Lauritzen C, Petersen GE, Pedersen J (2006) Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expr Purif* 48:1–13
- Banci L, Bertini I, Luchinat C, Mori M (2010) NMR in structural proteomics and beyond. *Prog Nucl Magn Reson Spectrosc* 56:247–266. doi:[10.1016/j.pnmrs.2009.12.003](https://doi.org/10.1016/j.pnmrs.2009.12.003)
- Billeter M, Wagner G, Wüthrich K (2008) Solution NMR structure determination of proteins revisited. *J Biomol NMR* 42:155–158. doi:[10.1007/s10858-008-9277-8](https://doi.org/10.1007/s10858-008-9277-8)
- Drucker DJ (2006) The biology of incretin hormones. *Cell Metab* 3:153–165. doi:[10.1016/j.cmet.2006.01.004](https://doi.org/10.1016/j.cmet.2006.01.004)
- Fehmann HC, Göke B (1995) Characterization of GIP(1–30) and GIP(1–42) as stimulators of proinsulin gene transcription. *Peptides* 16:1149–1152
- McIntosh CHS, Widenmaier S, Kim SJ (2009) Glucose-dependent insulinotropic polypeptide (gastric inhibitory polypeptide; GIP). *Vitam Horm* 80:409–471. doi:[10.1016/S0083-6729\(08\)00615-8](https://doi.org/10.1016/S0083-6729(08)00615-8)
- Parthier C, Kleinschmidt M, Neumann P et al (2007) Crystal structure of the incretin-bound extracellular domain of a G protein-coupled receptor. *Proc Natl Acad Sci USA* 104:13942–13947. doi:[10.1073/pnas.0706404104](https://doi.org/10.1073/pnas.0706404104)
- Pederson RA, Brown JC (1976) The insulinotropic action of gastric inhibitory polypeptide in the perfused isolated rat pancreas. *Endocrinology* 99:780–785
- Runge S, Thøgersen H, Madsen K et al (2008) Crystal structure of the ligand-bound glucagon-like peptide-1 receptor extracellular domain. *J Biol Chem* 283:11340–11347. doi:[10.1074/jbc.M708740200](https://doi.org/10.1074/jbc.M708740200)
- Stockman B (2002) NMR screening techniques in drug discovery and drug design. *Prog Nucl Magn Reson* 41:187–231
- Studier FW, Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* 189:113–130
- Takeuchi K, Wagner G (2006) NMR studies of protein interactions. *Curr Opin Struct Biol* 16:109–117. doi:[10.1016/j.sbi.2006.01.006](https://doi.org/10.1016/j.sbi.2006.01.006)
- Tsunoda Y, Sakai N, Kikuchi K et al (2005) Improving expression and solubility of rice proteins produced as fusion proteins in *Escherichia coli*. *Protein Expr Purif* 42:268–277. doi:[10.1016/j.pep.2005.04.002](https://doi.org/10.1016/j.pep.2005.04.002)
- Underwood CR, Garibay P, Knudsen LB et al (2010) Crystal structure of glucagon-like peptide-1 in complex with the extracellular domain of the glucagon-like peptide-1 receptor. *J Biol Chem* 285:723–730. doi:[10.1074/jbc.M109.033829](https://doi.org/10.1074/jbc.M109.033829)



Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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



Tetrameric ZBRK1 DNA binding domain has affinity towards cognate DNA in absence of zinc ions



Lumbini R. Yadav, Mahamaya N. Biswal, Vikrant, M.V. Hosur, Ashok K. Varma^{*}

Advanced Centre for Treatment, Research and Education in Cancer, Kharghar, Navi Mumbai, Maharashtra 410 210, India

ARTICLE INFO

Article history:

Received 12 May 2014

Available online 9 June 2014

Keywords:

ZBRK1

Zinc-finger

Oligomer

Secondary structure

ABSTRACT

Zinc finger transcription regulatory proteins play crucial roles in cell-cycle regulation, DNA damage response and tumor genesis. Human ZBRK1 is a zinc-finger transcription repressor protein, which recognizes double helical DNA containing consensus sequences of 5'GGGXXXCAGXXTTT3'. In the present study, we have purified recombinant DNA binding domain of ZBRK1, and studied binding with zinc ions and DNA, using biophysical techniques. The elution profile of the purified protein suggests that this ZBRK1 forms a homotetramer in solution. Dissociation and pull down assays also suggest that this domain forms a higher order oligomer. The ZBRK1-DNA binding domain acquires higher stability in the presence of zinc ions and DNA. The secondary structure of the ZBRK1-DNA complex is found to be significantly altered from the standard B-DNA conformation.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Zinc finger proteins, are known to play pivotal roles, through transcriptional regulation in DNA replication and repair, protein translation, cell proliferation and apoptosis. Regulation of gene expression using customized transcription factors has been reported recently [1,2]. Chimeric ZF proteins have been engineered to obtain customized activities, such as restriction, methylation and integration [3,4]. The erb2 oncogene and peripheral arterial obstructive diseases have been repressed by customized ZF proteins [5]. ZF proteins contain more than one ZF in tandem, and are considered potential targets for drugs. They can also make the genome a potential drug target.

There are different types of ZF proteins, and these are classified based on the environment around zinc ions. The major class is the C2H2 class, which contains a consensus amino acid sequence ψ -X-Cys-X(2-4)-Cys-X3- ψ -X5- ψ -X2-His-X(3,4)-His, where X is any amino acid and ψ is any hydrophobic residue. C2H2 type containing the KRAB (Kruppel-associated box) domain forms the largest subfamily (KRAB-ZFP) of ZF proteins [6] and, ZBRK1 is a member of this family. Human ZBRK1 performs a variety of biological functions through its various domains: the N-terminal KRAB, eight

consecutive central C2H2 zinc fingers and a BRCA1-dependent C-terminal transcriptional repression domain (CTRD) [7]. ZBRK1 consists of 532 amino acids, and the region from residues 206 to 424 is the DNA-binding domain (DBD). The N-terminal KRAB domain also performs repressor function by interacting with other proteins, such as KAP1. ZBRK1 represses cellular invasion and metastasis and, its loss enhances MMP9 transcription in cervical cancer [8]. ZBRK1 represses a high mobility group AT-hook 2 (HMG2), a DNA architectural protein, activation of which plays a significant role in tumor genesis and metastasis [9]. The repressor ZBRK1 binds to a 15 bp consensus nucleotide sequence, GGGXXXCAGXXTTT (where X is any nucleotide), which is found near promoter regions of many DNA damage inducible genes like p21, ANG1 [10] GADD153 and GADD45 [11,12]. Exposure of cells to DNA damaging agents leads to degradation of ZBRK1 via the BRCA1-independent ubiquitin-proteasome pathway leading to de-repression [13]. Interaction of ZBRK1 with Trim28 and histone deacetylase, results in repression of HIV-1 LTR acting as an intrinsic retroviral defense system [14]. In order to understand the various functions of ZBRK1, in details, we have initiated to study the structure and interactions of ZBRK1 with different ligands.

Here, we report purification and biophysical characterization of the DNA binding domain of ZBRK1 (ZBRK1-DBD), and an analysis of its binding with zinc and the cognate oligonucleotide. It is found that this domain exists as a homotetramer in solution, in contradiction to earlier report that C-terminal residues were essential for tetramer formation [15]. ZBRK1-DBD binds only to double

Abbreviations: CD, circular dichroism spectroscopy; ZBRK1, zinc finger and BRCA1-interacting protein with a KRAB domain 1; ZF, zinc-finger.

^{*} Corresponding author. Fax: +91 22 2740 5085.

E-mail address: avarma@actrec.gov.in (A.K. Varma).

stranded DNA, even in the absence of zinc ions. The DNA is found to be overwound on binding, while the conformation of the protein becomes less compact.

2. Materials and methods

2.1. Reagents

All the enzymes used for cloning were procured from NEB (US). Oligonucleotides were ordered from Sigma Aldrich (US). Chemicals used were from HiMedia, Qualigens and Sigma Aldrich. Bacterial culture media, IPTG and ampicillin were from HiMedia (India).

2.2. Cloning of ZBRK1-DNA binding domain

ZBRK1-DBD from amino acid 206 to 424 was sub-cloned at the BamHI and HindIII site in pRSET-A vector (Invitrogen). Forward primer 5'-GTCTCCGAGAACCTGTACTTTCAG GGTTGTGTGCAGTGAA TGTGGGAA-3' and Reverse primer 5'-GTCAAGCTTCT ATTAGTGAT TCTCTTATGCTTAAC-3' were used for polymerase chain reaction. The forward primer has TEV site incorporated in it so as to enable cleavage of the native protein from histidine tag. The ligated samples were transformed into *Escherichia coli* DH5 α cells and plated on ampicillin plates. Ampicillin resistant colonies were screened for positive clones and correct DNA sequence (Genetic Analyser-Applied Bio Systems).

2.3. Protein expression and purification

E. coli bacterial strain Rosetta 2(DE3) (Invitrogen) was transformed with ZBRK1-DBD construct. Single colony was inoculated in Luria Bertani broth containing ampicillin (100 μ g/ml) and chloramphenicol (34 μ g/ml), and the culture was allowed to grow for 12–14 h. Cells were then diluted 100 fold and grown till the OD₆₀₀ reached a value of 0.7. This culture was then induced with 0.4 mM IPTG for 16 h at 18 °C. Cells were harvested and resuspended in buffer-A containing 10 mM HEPES pH 7.0, 500 mM NaCl, 0.1% Triton X-100, 20 mM β -ME, 1 mM EDTA, 2.5% glycerol, and 10 mM imidazole. Cells were lysed by sonication at a pulse rate of 50, three times with 1 min duty cycle. Cell debris was removed by centrifugation. Cell lysis was allowed to bind on Ni-NTA resin for 1–2 h at 25 °C. To remove the nonspecifically bound protein, resin was washed with buffer-A addition containing 50 mM imidazole. Fusion protein was eluted with imidazole (500 and 750 mM) and further treated with TEV protease to remove histidine tag.

2.4. Gel filtration chromatography

ZBRK1-DBD was further purified to homogeneity using AKTA explorer system (GE Healthcare). Concentrated protein was chromatographed using Superdex-75 column pre equilibrated with buffer (5 mM HEPES pH 7.0, 150 mM NaCl, and 20 mM β -mercaptoethanol). Four proteins (BSA, lysozyme, phosphorylase B, and carbonic anhydrase) were eluted from the same column under similar conditions to prepare the standard plot for molecular weight estimation. The sample was run at a flow rate of 0.5 ml/min, and 1 ml fractions were collected. FPLC purified protein was further passed through buffer exchange column to remove β -mercaptoethanol. Peptide mass fingerprinting was done for protein confirmation using MALDI-MS (Ultraflex Bruker Daltonics system).

2.5. MBP pull down assay

MBP-ZBRK1 Δ K and MBP were expressed and affinity purified using amylose resin. Purified ZBRK1-DBD was incubated with

MBP-ZBRK1 Δ K and MBP for 3 h on ice with gentle mixing. Amylose resin was washed with buffer-A to remove non-specifically bound ZBRK1-DBD. The resin was then loaded on 12% SDS gel to check for the pull down.

2.6. Circular dichroism spectroscopy

Far UV CD scan of ZBRK1-DBD (7 μ M) was monitored to study secondary structural characteristics. Scan was recorded from 240 to 190 nm at a scanning speed of 20 nm/min, data interval of 0.1 nm and 3 accumulations. To study the effect of zinc sulfate and DNA on secondary structural characteristics, ZBRK1-DBD was titrated with zinc sulfate (10–180 μ M) and DNA (5–30 μ M). To reveal the structural alteration in DNA due to ZBRK1 binding, CD scan was obtained from 300 to 200 nm at a scanning speed of 20 nm/min. Protein concentration was estimated directly by Bradford method assuming a molar extinction coefficient of 12,950 M⁻¹ cm⁻¹. All the spectra were measured at 25 °C using (Jasco J-815) spectropolarimeter, and were baseline corrected. α -Helical content of the protein was estimated using ellipticity signals at 208 and 222 nm and the standard equation [16].

2.7. Thermal stability

To measure the thermal stability of ZBRK1-DBD (5 μ M), CD spectra in the range 250–200 nm was recorded as a function of temperature. Far UV CD scan was monitored at every 5 °C increase in temperature, with an equilibration time of 3 min and a scanning speed of 50 nm/min. Thermal unfolding pathway was also studied to reveal the effect of ZnSO₄ (0.1 mM) and DNA (30 μ M) on the stability of ZBRK1-DBD. Ellipticity at 222 nm was plotted as a function of temperature to understand the course of change in secondary structure elements.

2.8. Fluorescence spectroscopy

Spectrofluorometer (JOBIN YVON Horiba Fluorolog 3) was used for fluorescence measurements. The emission slit width was 5 nm. The fluorescence scan was measured in the emission wavelength range of λ_{em} 310–500 nm, and the excitation was at 295 nm. The spectral data were collected using FluorESSENCE software. Protein concentration was measured using molar extinction coefficient as described before. Titration experiments with ZBRK1-DBD and different concentrations of DNA (4–60 μ M) and ZnSO₄ (5–110 μ M) were performed to study the tertiary structural alterations.

2.9. Isothermal titration calorimetry (ITC)

ZBRK1-DBD (50 μ M) was injected into calorimetric cell containing buffer to study the heat effects from dissociation. Thirteen injections each 3 μ l, were injected at 25 °C, allowing an equilibration time of 150 s between injections. Interaction studies of ZBRK1-DBD with ZnSO₄ and consensus double stranded DNA oligonucleotide (5'GGGACGCGAGTTTAT3') were performed using MicroCal iTC200 calorimeter. (GE Healthcare) The titration experiments were done under following conditions: (1) 5.2 μ M of ZBRK1-DBD in cell and 1 mM of ZnSO₄ in syringe, (2) 20 μ M of ZBRK1-DBD in cell and 200 μ M of DNA in syringe. Protein concentration was estimated, using BSA as standard.

3. Results

3.1. Expression and purification of ZBRK1-DBD

Two liters of bacterial culture yielded 0.4 mg of >95% of pure ZBRK1-DBD protein (Fig. 1A). A single band corresponding to a

molecular weight of 24 kDa suggests that the preparation is very homogeneous. Peptide mass fingerprinting further confirmed the sample to be ZBRK1-DBD (data not shown).

3.2. ZBRK1-DBD forms a tetramer

ZBRK1-DBD eluted at approximately 58 ml from 16/60 Superdex75 size exclusion column (Fig. 1B). This indicates that the protein forms a higher order oligomer. From the standard linear graph of \log_{10} (molecular weight) versus (elution volume), the size of the eluted sample was estimated to be approximately 94 kDa (Fig. 1B). The partition coefficient K_{av} for ZBRK1-DBD was calculated from the elution volume [17,18]. Stoke's radius of ZBRK1-DBD was interpolated from a plot of partition coefficient (K_{av}) versus Stoke's radii for standard markers. The calculated Stoke's radius of ZBRK1-DBD was 5.6 nm which is closer to Stoke's radius of standard of 97 kDa. To explore the ability of self-association of ZBRK1-DBD, MBP pull down assay with MBP tagged ZBRK1 Δ K construct was performed. ZBRK1-DBD showed interaction with ZBRK1 Δ K construct. No interaction was visible with MBP (Fig. 1C).

Thus, based on size exclusion chromatography and pull down assay, it is proposed that the ZBRK1-DBD forms a homotetramer in solution.

To confirm the formation of an oligomer, next we performed dissociation study using ITC. Addition of concentrated protein from the syringe into calorimeter cell containing the buffer resulted in a heat change, which gradually decreased as concentration of the protein inside the cell increased (Fig. 1D). This observation suggests that the heat change was due to dissociation of an oligomeric species. The dissociation curve yielded enthalpy of dissociation as -482 kcal/mol with a dissociation constant of 0.125 μ M.

3.3. Secondary structure alterations on ligand binding

Far UV CD spectrum of ZBRK1-DBD is shown in (Fig. 2A). The scan showed peaks at approximately 222, 208 and 195 nm, which are characteristic of alpha-helical conformation [16,19]. The alpha-helical content of ZBRK1-DBD, estimated using measured ellipticities at 222 and 208 nm, was 12% and 15% respectively. The CD spectra of ZBRK1-DBD were recorded at different concentrations of zinc (Fig. 2B) and DNA (Fig. 2C). The $\theta_{222}/\theta_{208\text{ nm}}$ ratio of only protein is 1.0, and this ratio increases with further addition of ZnSO_4 , indicating incorporation of quaternary structure. In the case of interaction with DNA, this ratio decreased, indicating reduction in coiled coil structure [16,20]. Fig. 2D shows the CD spectrum over the λ range of 320–230 nm, when DNA was titrated with varying

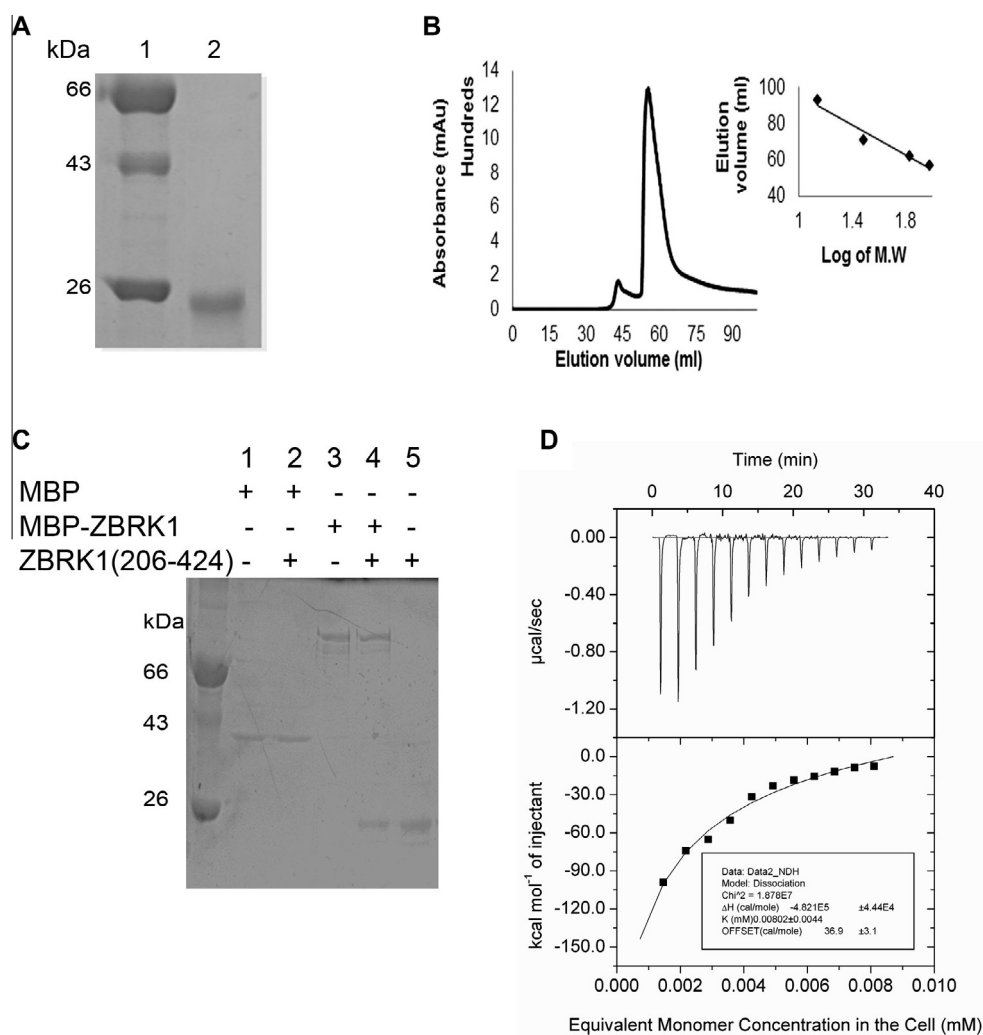


Fig. 1. Oligomeric characterization of ZBRK1-DBD: (A) size exclusion chromatography profile of purified ZBRK1-DBD protein. Inset is a plot of log of molecular weight standards (carbonic anhydrase, lysozyme, phosphorylase B and BSA) vs. elution volumes. (B) SDS electrophoresis showing sample purity. (C) MBP pull-down (lanes 1, 3, 5 – show input), (lane 2, 4 – pull down samples). (D) Exothermic heat pulses for 3 μ l injections into buffer of ZBRK1-DBD (50 μ M). Lower panel shows integrated heat data after blank correction.

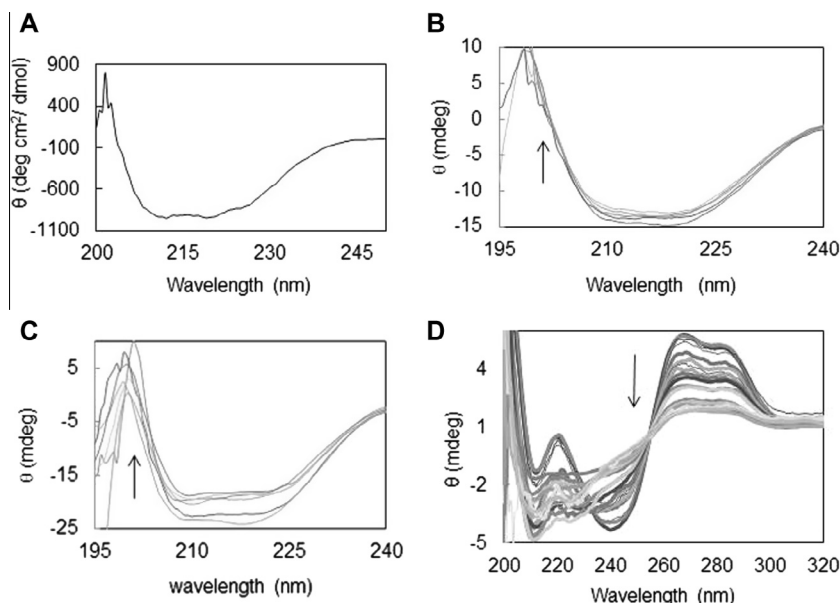


Fig. 2. CD spectra: (A) ZBRK1-DBD, ZBRK1-DBD titrated with (B) DNA, (C) zinc sulfate, (D) DNA titrated with ZBRK1-DBD. Arrows indicate directions of decrease in ellipticity.

amounts of ZBRK1-DBD. The CD spectrum of B-form DNA is composed of four major peaks: negative peaks around 214 and 245 nm and positive peaks around 225 and 280 nm [21,22]. The positive maximum around 280 nm is attributed to base stacking and the negative peak at 245 nm is attributed to DNA helicity [23]. The changes in ellipticities at 275 nm and 245 nm, on addition of ZBRK1-DBD, indicate that the standard B-geometry of the DNA is being altered by protein binding.

3.4. Intrinsic tryptophan fluorescence to elucidate local tryptophan environment

Tryptophan emission spectrum recorded from 310 to 500 nm, upon excitation at 295 nm (supplemental Fig. 1A). Emission maximum observed at 336 nm indicates that the single tryptophan residue is in hydrophobic environment [24]. Titration of ZBRK1-DBD with DNA (supplemental Fig. 1B) and zinc sulfate (supplemental Fig. 1C) showed a decrease in fluorescence intensity without any shift in the position of the maxima. This suggests that interaction of ZBRK1-DBD with either zinc sulfate or with DNA does not have any effect on local tryptophan environment.

3.5. Thermal stability

To evaluate the nature of folding and unfolding transitions, and to measure the structural stability of ZBRK1-DBD in presence of DNA and ZnSO_4 the far UV CD spectra, were recorded at different temperatures (Fig. 3). Variation in the ellipticity at 222 nm as a function of temperature suggests that ZBRK1-DBD unfolds via three state unfolding transitions with T_m values of 58 and 69 °C. Presence of zinc sulfate or DNA with ZBRK1 increased thermal stability of protein as suggested by increases in melting temperatures from 57 °C to 62 °C and 68 °C, respectively. Interestingly, in the presence of either DNA or zinc sulfate, ZBRK1 demonstrated simply two state unfolding transitions.

3.6. Thermodynamics of ligand binding to ZBRK1-DBD

Fig. 4 shows the results of calorimetry experiments to study the interactions of ZBRK1 with zinc sulfate and DNA. Model fitting for ZBRK1-DBD with zinc sulfate with one set of sites showed

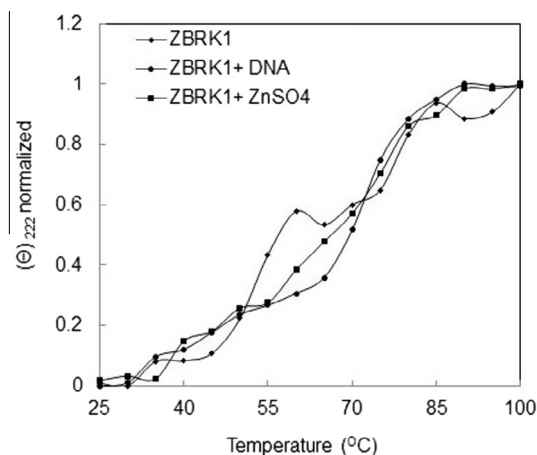


Fig. 3. Thermal denaturation profile of ZBRK1-DBD. Plot of fraction unfolded versus temperature at 5 °C intervals.

exothermic reaction with a stoichiometry of 1:42. It demonstrated an affinity of 1.4 μM under these conditions (Fig. 4A). Nonlinear curve fitting done using one set of sites for ZBRK1-DBD with DNA demonstrated a stoichiometry of 1:2 with binding affinity of 0.29 μM (Fig. 4B). ZBRK1-DBD interaction with zinc sulfate and DNA is an enthalpically driven reaction. Thermodynamic parameters of protein interaction with zinc and DNA, derived from ITC experiments are tabulated (Table 1).

4. Discussion

4.1. Oligomeric DBD

The elution profile of ZBRK1-DBD from the gel filtration column clearly indicates that ZBRK1-DBD was a homotetramer in solution. The absence of any substantial monomeric peak implies a very high affinity for tetramer formation. MBP pull down assay also revealed the ability of ZBRK1-DBD to oligomerize with ZBRK1ΔK in the MBP tagged ZBRK1ΔK construct. The ITC titration also suggests that dissociation of ZBRK1-DBD is enthalpically favored.

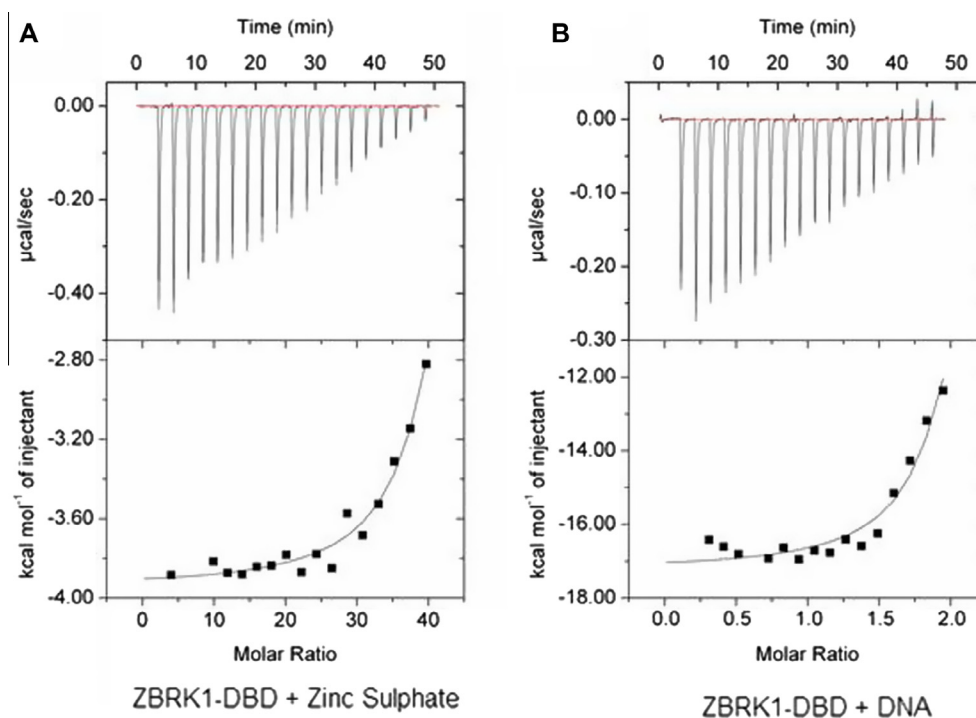


Fig. 4. Interaction of ZBRK1-DBD with: (A) zinc sulfate (B) DNA.

Table 1

Thermodynamic parameters derived from heat change: ZBRK1-DBD interaction with (1) zinc sulfate, (2) DNA.

	ZnSO ₄	DNA
Interaction model	One set of sites	One set of sites
N sites	42.9 ± 0.0817	2.1 ± 0.0538
K _d (μM)	1.4	0.29
K (M ⁻¹)	7.06E5 ± 1.91E5	3.44E6 ± 1.26E6
ΔH (cal/mol)	−3927 ± 29.88	−1.714E4 ± 192.6
ΔS (cal/mol/deg)	13.4	−27.6
Chi ² value	3510	1.34E + 05

The negative enthalpy change of dissociation implies that association process will be endothermic and entropically driven. The presence of two transitions in thermal unfolding pathway is also consistent with presence of tetrameric species. The first transition with a T_m of 58 °C would be for the disassembly of the tetramer, while the second transition with a T_m of 68 °C would be for denaturation of the monomers. The stoichiometry of 1:42 for zinc binding of ZBRK1-DBD, containing eight zinc fingers is also supportive for tetrameric oligomerisation of ZBRK1-DBD (Table 1). Tetramer formation has also been established by chemical cross-linking experiments [15]. However, in these studies, the molecule was full length ZBRK1ΔK construct, and not just the DNA binding domain used here. In fact experiments conducted by these authors with a variety of C-terminal deletion constructs show that, C-terminal nine residues are absolutely essential for formation of stable tetramers. Against this background our observation that ZBRK1-DBD, which lacks 108 residues from the C-terminal region, is a homo tetramer in solution, is a very novel finding. It is conceivable that the nature of tetramers formed by full length ZBRK1 and ZBRK1-DBD are different. The biological implications of this observation are not yet clear. Tetrameric DNA-binding proteins are known. The transcription factor p53 forms a tetramer. One dimer part of the tetramer binds to half site of consensus DNA sequence & the binding of second dimer to the other half enhances the affinity to DNA almost 50-fold [25]. Mutations in the mostly hydrophobic residues

responsible for tetramerization can inactivate the wild type protein suggesting the importance of tetramerization in p53 function [26] OB fold proteins that bind single stranded DNA are tetrameric [27,28]. Bioinformatics and EMSA analysis, however, shows that ZBRK1-DBD does not have the OB fold and does not bind ss DNA. (data not shown).

4.2. Zinc ions for DNA binding

The DNA-binding motif in ZF proteins comprises of two beta strands and an alpha helix, and the amino acid residues from the helix bind a triplet of bases in the major groove of the DNA helix [29–31]. Formation of the helix and proper positioning of the helical residues for optimal interaction is assumed to be ensured by coordination of zinc ion to conserved cysteine and histidine residues of the motif [32]. Therefore Zn ions are essential for repressor activity of C2H2 type of zinc fingers [33]. Interestingly, in the present study, ZBRK1-DBD binds to 15 bp consensus DNA sequence even in the absence of Zn ions (Fig. 2D, supplemental Fig. 1B). EMSA (data not shown) and ITC (Fig. 4A) also revealed interaction in absence of Zn ions.

4.3. DNA and protein conformation

It has been observed that the structure of DNA bound by ZF protein is altered from the standard B-geometry [34]. Molecular modeling has revealed that this distortion is necessitated by the requirement for neighboring fingers to bind optimally in the major groove. The actual distortions observed in the crystal structures of complexes are in agreement with predictions by modeling experiments. [35]. The relative intensity of CD signal at 275 nm is often taken as a signature for DNA geometry. An increase in its intensity suggests unwinding of DNA and/or displacement of base-pairs away from helix axis. However, in the present study on ZBRK1-DBD: DNA complex, the CD signals at 275 nm is significantly decreased. The geometry of the DNA in the complex is likely to be over wound compared to B-DNA. This probably may help ZBRK1

in repression activity, on interaction with DNA. Although previous reports mention no change occurs in ZF conformation on DNA binding, our results indicate subtle changes. Interestingly, binding of Zn and DNA to ZBRK1-DBD appear to have opposite effects on protein conformation. The $n-\pi^*$ transition at 220 nm in CD reflects the α -helical content. The $\pi-\pi^*$ excitation band at 208 nm is sensitive to whether the α -helix is involved in tertiary contacts [36]. The ratio $\theta_{222}/\theta_{208\text{ nm}}$ is, therefore, regarded as an indicator of quaternary structure through inter- α -helix coiling/interaction. The binding of protein to DNA (Fig. 2C) suggests decrease in $\theta_{222}/\theta_{208\text{ nm}}$ ratio indicating decrease in compactness of protein. On the other hand, ZBRK1-DBD with Zn ions indicated increase in $\theta_{222}/\theta_{208\text{ nm}}$ ratio (Fig. 2B) suggesting incorporation of quaternary structure in presence of Zn as compared to Zn deficient protein.

In conclusion, our report indicates that ZBRK1-DBD forms tetramer. ZBRK1-DBD binds exclusively to double stranded DNA with stoichiometry of 1:2, even in the absence of zinc ions. Interaction of ZBRK1-DBD with ZnSO_4 and DNA leads to higher thermal stability. In the complex with ZBRK1-DBD, the DNA is distorted from standard B-geometry. The future goal of this work is to crystallize the protein for understanding the structure, which will help in using it as a drug target and also to engineer it for therapeutic purposes.

Acknowledgments

We thank Dr. Thomas G. Boyer, University of Texas Health Science Center at San Antonio, Texas, USA for MBP-ZBRK1 Δ K construct. MVH thanks DAE for the award of Raja Ramanna Fellowship. AKV thanks TMC-ACTREC for funding.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bbrc.2014.05.104>.

References

- [1] H.S. Zhang, D. Liu, Y. Huang, S. Schmidt, R. Hickey, D. Guschin, H. Su, I.S. Jovin, M. Kunis, S. Hinkley, Y. Liang, L. Hinh, S.K. Spratt, C.C. Case, E.J. Rebar, B.E. Ehrlich, P.D. Gregory, F.J. Giordano, A designed zinc-finger transcriptional repressor of phospholamban improves function of the failing heart, *Mol. Ther.* 20 (2012) 1508–1515.
- [2] A. Klug, Towards therapeutic applications of engineered zinc finger proteins, *FEBS Lett.* 579 (2005) 892–894.
- [3] F.D. Bushman, M.D. Miller, Tethering human immunodeficiency virus type 1 preintegration complexes to target DNA promotes integration at nearby sites, *J. Virol.* 71 (1997) 458–464.
- [4] R.R. Beerli, B. Dreier, C.F. Barbas 3rd, Positive and negative regulation of endogenous genes by designed transcription factors, *Proc. Natl. Acad. Sci. U.S.A.* 97 (2000) 1495–1500.
- [5] P.Q. Liu, E.J. Rebar, L. Zhang, Q. Liu, A.C. Jamieson, Y. Liang, H. Qi, P.X. Li, B. Chen, M.C. Mendel, X. Zhong, Y.L. Lee, S.P. Eisenberg, S.K. Spratt, C.C. Case, A.P. Wolffe, Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A, *J. Biol. Chem.* 276 (2001) 11323–11334.
- [6] M.F. Rousseau-Merck, D. Koczan, I. Legrand, S. Moller, S. Autran, H.J. Thiesen, The KOX zinc finger genes: genome wide mapping of 368 ZNF PAC clones with zinc finger gene clusters predominantly in 23 chromosomal loci are confirmed by human sequences annotated in Ensembl, *Cytogenet. Genome Res.* 98 (2002) 147–153.
- [7] L. Zheng, H. Pan, S. Li, A. Flesken-Nikitin, P.L. Chen, T.G. Boyer, W.H. Lee, Sequence-specific transcriptional corepressor function for BRCA1 through a novel zinc finger protein, ZBRK1, *Mol. Cell* 6 (2000) 757–768.
- [8] L.F. Lin, C.H. Chuang, C.F. Li, C.C. Liao, C.P. Cheng, T.L. Cheng, M.R. Shen, J.T. Tseng, W.C. Chang, W.H. Lee, J.M. Wang, ZBRK1 acts as a metastatic suppressor by directly regulating MMP9 in cervical cancer, *Cancer Res.* 70 (2010) 192–201.
- [9] T. Ahmed, S.M. Kelly, N.C. Price, A.J. Lawrence, Activation of phospholipase A(2) by long chain fatty acyl groups involves a novel unstable linkage, *J. Biochem.* 127 (2000) 871–875.
- [10] S. Furuta, J.M. Wang, S. Wei, Y.M. Jeng, X. Jiang, B. Gu, P.L. Chen, E.Y. Lee, W.H. Lee, Removal of BRCA1/CtIP/ZBRK1 repressor complex on ANG1 promoter leads to accelerated mammary tumor growth contributed by prominent vasculature, *Cancer Cell* 10 (2006) 13–24.
- [11] W. Tan, L. Zheng, W.H. Lee, T.G. Boyer, Functional dissection of transcription factor ZBRK1 reveals zinc fingers with dual roles in DNA-binding and BRCA1-dependent transcriptional repression, *J. Biol. Chem.* 279 (2004) 6576–6587.
- [12] V. Garcia, J.M. Garcia, C. Pena, J. Silva, G. Dominguez, R. Rodriguez, C. Maximiano, R. Espinosa, P. Espana, F. Bonilla, The GADD45, ZBRK1 and BRCA1 pathway: quantitative analysis of mRNA expression in colon carcinomas, *J. Pathol.* 206 (2005) 92–99.
- [13] J. Yun, W.H. Lee, Degradation of transcription repressor ZBRK1 through the ubiquitin–proteasome pathway relieves repression of Gadd45a upon DNA damage, *Mol. Cell. Biol.* 23 (2003) 7305–7314.
- [14] H. Nishitsuji, M. Abe, R. Sawada, H. Takaku, ZBRK1 represses HIV-1 LTR-mediated transcription, *FEBS Lett.* 586 (2012) 3562–3568.
- [15] W. Tan, S. Kim, T.G. Boyer, Tetrameric oligomerization mediates transcriptional repression by the BRCA1-dependent Kruppel-associated box-zinc finger protein ZBRK1, *J. Biol. Chem.* 279 (2004) 55153–55160.
- [16] S.M. Kelly, N.C. Price, The use of circular dichroism in the investigation of protein structure and function, *Curr. Protein Pept. Sci.* 1 (2000) 349–384.
- [17] F. Wolfram, E. Morris, C.W. Taylor, Three-dimensional structure of recombinant type 1 inositol 1,4,5-trisphosphate receptor, *Biochem. J.* 428 (2010) 483–489.
- [18] C. Dulac, A.A. Michels, A. Fraldi, F. Bonnet, V.T. Nguyen, G. Napolitano, L. Lania, O. Bensaude, Transcription-dependent association of multiple positive transcription elongation factor units to a HEXIM multimer, *J. Biol. Chem.* 280 (2005) 30619–30629.
- [19] R.W. Woody, Circular dichroism, *Methods Enzymol.* 246 (1995) 34–71.
- [20] T. Rao, G. Ruiz-Gomez, T.A. Hill, H.N. Hoang, D.P. Fairlie, J.M. Mason, Truncated and helix-constrained peptides with high affinity and specificity for the cFos coiled-coil of AP-1, *PLoS One* 8 (2013) e59415.
- [21] M. Vorlickova, Conformational transitions of alternating purine–pyrimidine DNAs in perchlorate ethanol solutions, *Biophys. J.* 69 (1995) 2033–2043.
- [22] J. Kypr, M. Vorlickova, Circular dichroism spectroscopy reveals invariant conformation of guanine runs in DNA, *Biopolymers* 67 (2002) 275–277.
- [23] C.A. Sprecher, W.A. Baase, W.C. Johnson Jr., Conformation and circular dichroism of DNA, *Biopolymers* 18 (1979) 1009–1019.
- [24] R.W. Alston, L. Urbanikova, J. Sevcik, M. Lasagna, G.D. Reinhart, J.M. Scholtz, C.N. Pace, Contribution of single tryptophan residues to the fluorescence and stability of ribonuclease Sa, *Biophys. J.* 87 (2004) 4036–4047.
- [25] K.G. McLure, P.W. Lee, How p53 binds DNA as a tetramer, *EMBO J.* 17 (1998) 3342–3350.
- [26] P. Chène, The role of tetramerization in p53 function, *Oncogene* 20 (2001).
- [27] S. Raghunathan, C.S. Ricard, T.M. Lohman, G. Waksman, Crystal structure of the homo-tetrameric DNA binding domain of *Escherichia coli* single-stranded DNA-binding protein determined by multiwavelength X-ray diffraction on the selenomethionyl protein at 2.9-Å resolution, *Proc. Natl. Acad. Sci. U.S.A.* 94 (1997) 6652–6657.
- [28] D.L. Theobald, R.M. Mitton-Fry, D.S. Wuttke, Nucleic acid recognition by OB-fold proteins, *Annu. Rev. Biophys. Biomol. Struct.* 32 (2003) 115–133.
- [29] S.A. Wolfe, L. Nekludova, C.O. Pabo, DNA recognition by Cys2His2 zinc finger proteins, *Annu. Rev. Biophys. Biomol. Struct.* 29 (2000) 183–212.
- [30] D. Lu, C2H2 zinc-finger recognition of biomolecules, *Yao Xue Xue Bao* 48 (2013) 834–841.
- [31] P.V. Pedone, R. Ghirlando, G.M. Clore, A.M. Gronenborn, G. Felsenfeld, J.G. Omichinski, The single Cys2-His2 zinc finger domain of the GAGA protein flanked by basic residues is sufficient for high-affinity specific DNA binding, *Proc. Natl. Acad. Sci. U.S.A.* 93 (1996) 2822–2826.
- [32] R.N. De Guzman, H.Y. Liu, M. Martinez-Yamout, H.J. Dyson, P.E. Wright, Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP, *J. Mol. Biol.* 303 (2000) 243–253.
- [33] E.E. Zhelezanova, J.H. Crosa, R.G. Brennan, Characterization of the DNA- and metal-binding properties of *Vibrio anguillarum* fur reveals conservation of a structural Zn(2+) ion, *J. Bacteriol.* 182 (2000) 6264–6267.
- [34] L. Nekludova, C.O. Pabo, Distinctive DNA conformation with enlarged major groove is found in Zn-finger-DNA and other protein–DNA complexes, *Proc. Natl. Acad. Sci. U.S.A.* 91 (1994) 6948–6952.
- [35] M. Elrod-Erickson, M.A. Rould, L. Nekludova, C.O. Pabo, Zif268 protein–DNA complex refined at 1.6 Å: a model system for understanding zinc finger–DNA interactions, *Structure* 4 (1996) 1171–1180.
- [36] N.E. Zhou, C.M. Kay, R.S. Hodges, Synthetic model proteins. Positional effects of interchain hydrophobic interactions on stability of two-stranded alpha-helical coiled-coils, *J. Biol. Chem.* 267 (1992) 2664–2670.