Molecular Mechanism of BRCA1-BARD1-CstF50 complex and Breast Cancer Risk

By

Rajan Kumar Choudhary

[LIFE09201004019]

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

November 2016

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 Rajan Kumar Choudhary entitled **"Molecular Mechanism of BRCA1-BARD1-CstF50 complex and Breast cancer Risk"** and recommend that it may be accepted as fulfilling the dissertation requirements for the Degree of Doctor of Philosophy.

Chairperson: - Dr. Vinay Kumar Date: 07/11/2016

Guide/Convenor - Dr. Ashok K Varma Date: 7/11/2016

Co-Guide: - Dr. Rajiv Sarin Date: 7.11.16

Member 1: - Dr. Sanjay Gupta Date: 7/11/16

Invitee : - Dr. Smita Mahale Date: 7.11.2016

Technical Advisor :- Dr. C.P Rao Date: 7-11-2016

External Examiner: - Dr. Pinakpani Chakrabarti Date: 7-11-16

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 :- 07/11/2016

Place: Navi-Mumbai

Dr. Rajiv Sarin
(Co-Guide)

Dr. Ashok K Varma
(Guide)

STATEMENT BY AUTHOR

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

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

Navi Mumbai

Date: 09/11/2016 (Rajan Kumar Choudhary)

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: 09/11/2016 (Rajan Kumar Choudhary)

LIST OF PUBLICATIONS ARISING FROM THE THESIS

List of Publications

(A) Journal

1) Multimodal approach to explore the pathogenicity of BARD1, ARG 658 CYS, and ILE 738 VAL mutants. *Choudhary RK, Vikrant1, Siddiqui QM, Thapa PS, Raikundalia S, Gadewal N, Kumar NS, Hosur MV, Varma AK.J Biomol Struct Dyn. 2016 Jul;34(7):1533-44. doi: 10.1080/07391102.2015.1082149. Epub 2015 Sep 23.*

2) Structural and biophysical properties of h-FANCI ARM repeat protein. *Siddiqui MQ, Choudhary RK, Thapa P, Kulkarni N, Rajpurohit YS, Misra HS, Gadewal N, Kumar S, Hasan SK, Varma AK. J Biomol Struct Dyn. 2016 Sep 30:1-26. [Epub ahead of print]*

3) Functional basis and biophysical approaches to characterize the C-terminal domain of human-Ribosomal s6 Kinase. *Jagilinki BP, Choudhary RK, Thapa PS, Gadewal N, Hosur MV, Kumar S, Varma AK. Cell Biochem Biophys. 2016 Sep;74(3):317-25. doi: 10.1007/s12013-016-0745-6. Epub 2016 Jun 21.*

4) Mislocalization of BRCA1-complex due to ABRAXAS Arg361Gln mutation; *Vikrant1, Kumar R, Siddiqui Q, Singh N, Waghmare SK, Varma AK; J Biomol Struct Dyn. 2015;33(6):1291-301. doi: 10.1080/07391102.2014.945484. Epub 2014 Aug 8.*

5) Structural and Functional Implication of RAP80 Δ81Glu Mutation; *Vikrant1, Kumar R, Yadav LR, Nakhwa P, Waghmare SK, Goyal P, Varma AK., PLoS One. 2013 Sep 9;8(9):e72707. doi: 10.1371/journal.pone.0072707. eCollection 2013.*

6) Structural and functional characterization of the MERIT40 to understand its role in DNA repair; *Vikrant1, Nakhwa P, Badgujar DC, Kumar R, Rathore KK, Varma AK. J Biomol Struct Dyn. 2014 Dec;32(12):2017-32. doi: 10.1080/07391102.2013.843473. Epub 2013 Oct 15.*

(B) Chapters in books and lectures notes: N/A

(C) Conferences:-

- **Indian Biophysical Society (IBS) 2015**, (Jamia Millia Islamia, India, Jan 2015).
  Nature of Participation: Poster presentation (Title: Biophysical characterization of BARD1BRCT mutants). **Awarded 1st prize for best poster presentation**.

- **Tata memorial center platinum jubilee meeting** "**A Conference of New Ideas in Cancer-Challenging Dogmas" (**Mumbai Feb 2016).

Nature of Participation: Poster presentation (Title: Multidisciplinary approach to understand BARD1-CstF50 complex interface and evaluate functional implication of BARD1 Gln564His mutation).

- **10th European Biophysical society meeting** (Dresden, Germany, Jul 2015).
  Nature of Participation: Poster presentation (Title: Biophysical and functional characterization of BARD1BRCT mutants.)

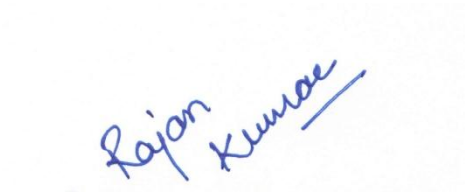
(D) Others:-

(1) Manuscript communicated:-

- RK Choudhary & Vikrant et al. Structural basis to stabilize the domain motion of BARD1-ARD BRCT by CstF50. Nature Scientific Reports; 2016. Communicated. (SREP-16-37456).

(2) Manuscripts in preparation:-

- Structural basis of BARD1BRCT and cancer predisposing mutation: A biophysical and molecular dynamics simulation approach.



Date: - 09/11/2016

Rajan Kumar Choudhary

DEDICATIONS

……dedicated to my loving grandfather and Father.

ACKNOWLEDGEMENTS

I take this opportunity to offer my sincere thanks to my mentor Dr. Ashok K. Varma and *Dr. Rajiv Sarin for his guidance, incessant encouragement and prudent suggestions. 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. Vinay Kumar, Dr. C.P Rao, Dr. Smita Mahale and Dr. Sanjay Gupta as my doctoral committee members. Their critical analysis and helpful suggestions have contributed significantly to the work. I am extremely thankful to the Dr. M.V. Hosur, 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 Dr. Vikrant, Dr. Dilip Badgujar, Dr. Bhanu Prakash, Quadir, Mrs Ulka Sawant, all my present lab members, Dr. Ram Kumar Singh, Satish Singh, Akash Dubey, Manish Pandey 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 grandfather, mother, brother and other family members for all their sacrifices and affectionate blessings. I would reserve my heartiest thanks for my brothers Gunjan Choudhary, Dr. Atul Pranay, Dr. Ravi shanker, Nitin, Abhinav, Abhishek and sister soumya for being my pillar of strength through thick and thin and loving me for who I am.

Last but not least, I would like to thank my father who has provided me love, strength, motivation and guiding lessons of life. I don't know where life will take me as I grow older, but one thing I know is that I will never be able to be a person better than you. I cherish precious memories of you and find strength in knowing that even though you are gone you still live on in my heart forever. Though, all cannot be mentioned, but none is forgotten.

CONTENTS

Homi Bhabha National Institute

Ph. D. PROGRAMME

1. Name of the Student:	Rajan Kumar Choudhary
2. Name of the Constituent Institution:	Tata Memorial Centre**,** Advanced Centre for Treatment Research and Education in Cancer
3. Enrolment No. :	LIFE09201004019
4. Title of the Thesis:	**"Molecular Mechanism of BRCA1- BARD1-CstF50 Complex and Breast Cancer risk**"
5. Board of study:	Life Sciences

SYNOPSIS

Cancer is a multifaceted phenomenon in which a cell generates an innate ability to sustain proliferation, becomes unresponsive to growth suppression, attains ability to invade, metastasis, and acquires replicative perpetuity by inducing angiogenesis [1]. Breast cancer arises due to inherited germ-line and somatic mutations in Breast cancer susceptibility gene 1 (BRCA1) and BRCA2 genes [2, 3]. BRCA1 is a tumor suppressor protein involved in different key pathways like DNA damage repair, cell-cycle control, transcription and apoptosis [4, 5]. BRCA1 associated Ring domain protein 1 (BARD1) is N–terminal (RING-RING) interacting partner of BRCA1, and also reported as binary interacting partner of CstF50 (Cleavage Stimulating factor 1) [6]. BRCA1-BARD1 complex posses E3-ubiquitin ligase activity, and ubiquitinates different molecules like γ-tubulin, H2AX, Estrogen receptor-α, RNA-polymerase CTD and SWI-SNF, thus controlling transcription coupled repair, ploidy regulation and chromatin remodeling [7-11]. BRCA1-BARD1and CStF50 complex controls transcription coupled DNA damage repair (TCR) [6, 12, 13]. BRCTs (C-terminal region of BRCA1) domain containing proteins play a pivotal role in the DNA repair by recruiting DNA-damage repair protein in phosphorylation dependent manner [14-19]. However, BRCT domain is highly conserved in BRCA1 and BARD1.

In the present study, we have investigated the binding mechanism of BRCA1- BARD1-CstF50 complex, and how germ-line mutation Gln564His in BARD1 leads to abrogation of the complex formation. Furthermore, significance of BRCA1-BARD1-CstF50 complex has been explored to understand its interplay in transcription coupled DNA damage repair mechanism. Our study also investigates the effect of pathogenic mutations reported in the BARD1BRCT domain using biophysical and computational biology based approach.

This thesis comprises of six chapters and a summary chapter.

Chapter 1 Provides an overview of binding mechanism of BRCA1-BARD1-CStF50 complex and its role in DNA damage repair.

DNA-damage-response is a result of complex cross-talk between different pathways for cell-cycle arrest, the transcriptional and post-transcriptional activation of genes including those associated with DNA damage repair, and under some circumstances, triggering of the programmed cell death [20]. The N-termini RING-RING domain complex of BRCA1-BARD1 is an active E3 ubiquitin ligase complex, which is known to be inhibited by platinum drugs [21-24]. Tumor suppressor BARD1and the mRNA 3'end processing factors and cleavage-stimulating factor 1(CstF50), 2 and 3 are required for the ultraviolet (UV)-induced inhibition of the mRNA 3' cleavage step of the polyadenylation reaction [25]. Interactive relationship between BRCA1-BARD1 and CstF50 is important in Transcriptional Coupled Repair [6, 13, 26]. Cancer predisposing mutation Gln564His in BARD1 BRCT has been found in the germ-line of a patient with adenocarcinoma [27]. BARD1 Gln564His impairs the complex formation between BARD1 and CstF50, which further leads to formation of premature transcript and loss of mRNA regulation.

BRCT domain is evolutionarily conserved domain involved in DNA damage repair [28]. BRCA1, BARD1 and MDC1 contain the conserved phosphopeptide binding motif, and act as docking sites for DNA damage repair proteins in a phosphorylation dependent manner [29-31]. Recently, it has been established that Poly A-Ribose (ADP-ribose monomer) interacts with BARD1BRCT, and is required for the early recruitment of BRCA1 to the DNA damage site [18]. Cancer predisposing mutations Cys645Arg, Val695Leu, Ser761Asn, Arg658Cys and Ile738Val are reported in the BARD1BRCT domain [18, 32, 33]. Considering the importance of BRCA1-BARD1-CstF50 complex, we have undertaken an interdisciplinary in-silico, in-vitro and

biophysical approach to understand the complex integrity, and alterations in the molecular structure due to cancer predisposing mutations.

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

Different genes for expression and purification were first sub-cloned from the provided mammalian c-DNA (generous gift from Prof Richard Bayer, Institute of Cancer Genetics, Columbia University) construct in to bacterial expression vectors. Further, mutants were generated by site-directed mutagenesis using Mis-match primer method on *wild- type* template. The cloned constructs were validated by DNA sequencing. The *wild- type* and mutant proteins were purified first by affinity chromatography, and then by molecular exclusion chromatography to get highly purified protein for biophysical characterization and crystallography. Purified *wild-type* and mutant proteins were functionally characterized using biophysical tools such as Circular Dichroism, Fluorescence, Differential Scanning Calorimetry, Dynamic Light Scattering, and Mass Spectrometry. Protein-Protein and Protein-Platinum drug interactions were analyzed using HADDOCK server and Discovery studio tool Glide [34, 35].

Chapter 3 describes biophysical, biochemical and structural characterization of ARD-BARD1BRCT *wild- type* and BARD1 Gln564His mutant protein to unravel the residues involved in the binding interface of BARD1-CstF50 complex**.**

ARD-BARD1BRCT (425-777aa) domain has been reported to bind with CstF-50, p53, and ATM [25, 36]. ARD domain of BARD1 is interconnected with BRCT domain (568-777aa) through a flexible 18 amino acid linker. Germ-line mutation BARD1 Gln564His was generated using site-directed-mutagenesis, sub-cloned into bacterial expression pGEX-kT vector, expressed and purified for structural and biochemical studies. It has been observed that *wild-type* protein is

monomer and mutation does not affect the monomeric nature of the protein having similar hydrodynamic diameter of 6.4 nm. Crystallization attempts have not yet yielded good quality single crystals. Therefore, homology model was built using the Robetta server [37]. The molecular model suggests disorderness in the linker region, which is consistent with the earlier reported structure of ankyrin repeat and BRCT domain [38, 39]. Thermal denaturation of BARD1 ARD-BRCT *wild-type* and mutant proteins indicates that both unfold through a two state pathway. However, chemical denaturation (using GuHcl) profiles suggest that BARD1 ARD-BRCT *wild-type* and mutant unfold through formation of an intermediate species. Further, molecular dynamics simulation was performed to understand the structural alterations due to mutations. Molecular dynamics simulation performed for BARD1 ARD-BRCT *wild-type* and Gln564His suggest that ARD domain is independent of BRCT domain and does not have fixed orientation in the space. BARD1 Gln564His mutation which is located in the loop region does not cause any significant effect on the secondary structure of the protein. CstF50 as a binary interacting partner of BRCA1-BARD1 regulates the transcription couple repair. The CstF50 full length (1-431aa) and WD-40 (Tryptophan Aspartate) repeat domain (92-431aa) was sub-cloned into bacterial expression vector pGEX-kT and pET28a^{+}, and the protein was expressed in *E.coli Rosetta 2(DE3)* and *BL21 (DE3)* strain. The full-length and WD-40 repeat domain were not soluble in both the expression system. Therefore, the six His tagged fusion full-length CstF50 protein was purified from inclusion bodies by using 8M urea and further refolded. The refolded protein was subjected to fluorescence spectroscopy, and it was concluded that protein is properly refolded. Crystallization was not attempted as refolding process yielded very low amount of protein and precipitated during concentration. Noting the problem in crystallization, the in-silico model of full- length CstF50 was built and docked onto the BARD1 ARD-BRCT structure. The

molecular dynamics simulation was performed for the complex structure, and observed that binding to CstF50 reduces the domain flexibility and forms a stable complex with BARD1 ARD-BRCT domain.

Chapter 4 describes the comparative biophysical characterization of BARD1BRCT *wild-type* and reported pathogenic mutants

To understand the effect of mutations on the structure, BARD1BRCT *wild-type* and mutants were sub-cloned into pET28a^{+} vector. *Wild-type* and mutant proteins were expressed and purified to carry out biophysical and crystallographic studies. BARD1BRCT is reported to be monomeric in nature [38], and to investigate whether mutation has affected the monomeric property of the protein, mutant proteins were subjected to size exclusion chromatography followed by glutaraldehyde crosslinking assay. It has been observed that mutants are monomeric in nature. CD and fluorescence spectroscopy suggests that BARD1BRCT Cys645Arg, Val695Leu, Ser761Asn, Arg658Cys and Ile738Val mutations are not driving any global secondary and tertiary structural changes in the *wild-type* protein. Thermal denaturation of *wild-type* and mutant protein indicates that except Cys645Arg, other mutant proteins Val695Leu, Ser761Asn, Arg658Cys and Ile738Val have lost thermal stability drastically, and show two state unfolding pattern. Chemical denaturation also show mutants have lost thermodynamic stability. Hence, BARD1BRCT Arg645Cys, Val695Leu, Ser761Asn, Arg658Cys and Ile738Val mutant proteins unfold via three state pathways. It has been concluded that BARD1BRCT Cys645Arg, Val695Leu, Ser761Asn, Arg658Cys and Ile738Val mutations does not affect monomeric property. However, BARD1BRCT Cys645Arg mutant protein do not show change in the thermal stability compared to *wild- type*. The BARD1BRCT (PDB ID: 2NTE) structure was used to incorporate the pathogenic mutations using SPDB viewer [40]. *Wild- type* and mutant protein

were simulated for 50ns in periodic boundary condition using GROMACS 4.5.5 [41-43]. Further trajectory was analyzed for RMSD, RMSF, R_g, number of hydrogen bond, volume, and SASA using Gromacs in built commands g_rms, g_rmsf, g_gyrate, g_bond, and g_sas respectively. Principal component analysis (PCA) was performed to understand the difference in the concerted motion of the proteins. RMSD calculation for *wild-type* and mutant protein indicates alterations in conformation of mutant structures compared to *wild- type*. BARD1BRCT Ile738Val mutant structure shows highest RMSD, volume and SASA value for entire MD production run. Further, Ile738Val also shows increased RMSF values for during MD production run, and the mutant shows loss of average intra-molecular hydrogen bond, which also validates loss in thermodynamic stability of BARD1BRCT Val695Leu, Ser761Asn, Arg658Cys and Ile738Val mutant proteins. PCA analysis for *wild -type* and mutant protein indicates that the significant changes in the atomic motion of the mutant proteins than the *wild-type*. Projection of eigenvectors on residue for BARD1BRCT Arg645Cys, Val695Leu, Ser761Asn, Arg658Cys and Ile738Val protein shows mutation has local effect on the structure of the mutant protein.

Chapter 5 describes the association between BRCA1 and platinum drugs Cis-platin and Trans-platin.

The *wild-type* gene BRCA1 (1-303aa) and BARD1 (16-119aa) were sub-cloned into the bacterial expression system, pET-41A and pGEX-kT respectively. The BRCA1 Ring domain (1-303aa) protein was not expressing in the bacterial system, but GST fused BARD1 (16-119aa) was expressing and found soluble. However, it precipitates at higher concentration 0.7 mg/ml. It is quite possible that BARD1 RING domain is not stable alone and requires BRCA1 for stabilization. We further generated molecular model for BRCA1 (1-125aa) amino acids using Robetta server [35] and performed docking study of small molecules, Trans and Cis-platin using

discovery studio [34, 36]. The modeled structure suggests that BRCA1 (1-125aa) amino acids is a predominately helical protein of RING finger like structure with two Zn^{+2} binding sites.

In conclusion, BRCA1 (1-125aa) molecular model is predominantly composed of α- helices having conserved C_3HC_4 zinc finger which binds to two Zn^{2+} ions. Cis and Trans Platinum drug binding site were observed close to the Zn^{2+} binding site in ring finger domain, and that may cause the structural perturbation in Zn^{2+} finger domain. The Pt coordination to Zn^{2+}-binding sites in Ring domain may lead to Zn^{2+} ejection from the ring domain and consequently loss of protein tertiary structure, and inhibition of biological functions regulated by BRCA1 BARD1 heterodimerization.

Chapter 6 describes important research findings and conclusions of the work

BRCA1-BARD1-CstF50 complex plays an important role in the regulation of the transcription coupled DNA damage repair. Loss of functional BRCA1-BARD1-CstF50 complex due to the germ-line mutation Gln564His in the linker region of the ARD-BARD1BRCT domain leads to loss in CstF50 binding to the ARD-BARD1BRCT. Biophysical analysis of BARD1BRCT and mutant protein indicated that the mutations Arg645Cys, Val695Leu, Ser761Asn, Arg658Cys and Ile738Val in different position of BARD1 are not abrogating the monomeric, secondary structural and overall packing of *wild-type* protein. Further, pathogenic mutations reported in the BARD1BRCT domain displays similar resistivity towards the proteases. Alterations in weak intramolecular interactions for BARD1 Cys645Arg, Val695Leu, Ser761Asn, Arg658Cys and Ile738Val mutations have been reflected in loss in thermal stability of the mutant proteins. BARD1 Ile738Val and Val695Leu, mutation resulted in loss of structural compactness. The study presented here unravels the interactions involved in BRCA1-BARD1-CstF50-complex

formation and understanding the transcriptional coupled DNA damage repair process. Comparative biophysical and *in-silico* studies of BARD1BRCT and mutants will further explore the opportunities of structure based small molecule inhibitor design for therapeutic applications.

References:

1. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-674.
2. Alberg, A.J., A.P. Lam, and K.J. Helzlsouer, *Epidemiology, prevention, and early detection of breast cancer.* Current opinion in oncology, 1999. **11**(6): p. 435.
3. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.* Science, 1994. **266**(5182): p. 66-71.
4. Venkitaraman, A.R., *Cancer susceptibility and the functions of BRCA1 and BRCA2.* Cell, 2002. **108**(2): p. 171-82.
5. Lane, T.F., *BRCA1 and transcription.* Cancer Biol Ther, 2004. **3**(6): p. 528-33.
6. Kleiman, F.E. and J.L. Manley, *The BARD1-CstF-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression.* Cell, 2001. **104**(5): p. 743-53.
7. Sankaran, S., et al., *Centrosomal microtubule nucleation activity is inhibited by BRCA1-dependent ubiquitination.* Mol Cell Biol, 2005. **25**(19): p. 8656-68.
8. Abbott, D.W., et al., *BRCA1 expression restores radiation resistance in BRCA1-defective cancer cells through enhancement of transcription-coupled DNA repair.* J Biol Chem, 1999. **274**(26): p. 18808-12.
9. Hu, Y.F., Z.L. Hao, and R. Li, *Chromatin remodeling and activation of chromosomal DNA replication by an acidic transcriptional activation domain from BRCA1.* Genes Dev, 1999. **13**(6): p. 637-42.
10. Xia, Y., et al., *Enhancement of BRCA1 E3 ubiquitin ligase activity through direct interaction with the BARD1 protein.* Journal of Biological Chemistry, 2003. **278**(7): p. 5255-5263.
11. Eakin, C.M., et al., *Estrogen receptor α is a putative substrate for the BRCA1 ubiquitin ligase.* Proceedings of the National Academy of Sciences, 2007. **104**(14): p. 5794-5799.
12. Kleiman, F.E. and J.L. Manley, *Functional interaction of BRCA1-associated BARD1 with polyadenylation factor CstF-50.* Science, 1999. **285**(5433): p. 1576-9.
13. 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.
14. Cantor, S., et al., *The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations.* Proc Natl Acad Sci U S A, 2004. **101**(8): p. 2357-62.
15. Cantor, S.B., et al., *BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function.* Cell, 2001. **105**(1): p. 149-60.
16. 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.
17. 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.

18. Li, M. and X. Yu, *Function of BRCA1 in the DNA damage response is mediated by ADP-ribosylation.* Cancer Cell, 2013. **23**(5): p. 693-704.
19. 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.
20. Zhou, B.-B.S. and S.J. Elledge, *The DNA damage response: putting checkpoints in perspective.* Nature, 2000. **408**(6811): p. 433-439.
21. Wu-Baer, F., et al., *The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin.* J Biol Chem, 2003. **278**(37): p. 34743-6.
22. Atipairin, A., B. Canyuk, and A. Ratanaphan, *Cisplatin affects the conformation of apo form, not holo form, of BRCA1 RING finger domain and confers thermal stability.* Chemistry & biodiversity, 2010. **7**(8): p. 1949-1967.
23. Atipairin, A., B. Canyuk, and A. Ratanaphan, *The RING heterodimer BRCA1–BARD1 is a ubiquitin ligase inactivated by the platinum-based anticancer drugs.* Breast cancer research and treatment, 2011. **126**(1): p. 203-209.
24. Atipairin, A. and A. Ratanaphan, *In Vitro enhanced sensitivity to cisplatin in D67Y BRcA1 RInG Domain protein.* Breast cancer: basic and clinical research, 2011. **5**: p. 201.
25. Nazeer, F.I., et al., *p53 inhibits mRNA 3' processing through its interaction with the CstF/BARD1 complex.* Oncogene, 2011. **30**(27): p. 3073-83.
26. Starita, L.M., et al., *BRCA1/BARD1 ubiquitinate phosphorylated RNA polymerase II.* J Biol Chem, 2005. **280**(26): p. 24498-505.
27. 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.
28. Yu, X., et al., *The BRCT domain is a phospho-protein binding domain.* Science, 2003. **302**(5645): p. 639-642.
29. Glover, J.N., R.S. Williams, and M.S. Lee, *Interactions between BRCT repeats and phosphoproteins: tangled up in two.* Trends Biochem Sci, 2004. **29**(11): p. 579-85.
30. Manke, I.A., et al., *BRCT repeats as phosphopeptide-binding modules involved in protein targeting.* Science, 2003. **302**(5645): p. 636-9.
31. Yu, X., et al., *The BRCT domain is a phospho-protein binding domain.* Science, 2003. **302**(5645): p. 639-42.
32. De Brakeleer, S., et al., *Cancer predisposing missense and protein truncating BARD1 mutations in non-BRCA1 or BRCA2 breast cancer families.* Hum Mutat, 2010. **31**(3): p. E1175-85.
33. Ghimenti, C., et al., *Germline mutations of the BRCA1-associated ring domain (BARD1) gene in breast and breast/ovarian families negative for BRCA1 and BRCA2 alterations.* Genes Chromosomes Cancer, 2002. **33**(3): p. 235-42.
34. De Vries, S.J., M. van Dijk, and A.M. Bonvin, *The HADDOCK web server for data-driven biomolecular docking.* Nature protocols, 2010. **5**(5): p. 883-897.
35. Studio, D., *version 3.5.* Accelrys Inc.: San Diego, CA, USA, 2012.
36. Colgan, D.F. and J.L. Manley, *Mechanism and regulation of mRNA polyadenylation.* Genes Dev, 1997. **11**(21): p. 2755-66.
37. 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.

38. Birrane, G., et al., *Crystal structure of the BARD1 BRCT domains.* Biochemistry, 2007. **46**(26): p. 7706-12.
39. Fox, D., et al., *Crystal structure of the BARD1 ankyrin repeat domain and its functional consequences.* Journal of Biological Chemistry, 2008. **283**(30): p. 21179-21186.
40. Guex, N. and M.C. Peitsch, *SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling.* electrophoresis, 1997. **18**(15): p. 2714-2723.
41. Berendsen, H.J., D. van der Spoel, and R. van Drunen, *GROMACS: A message-passing parallel molecular dynamics implementation.* Computer Physics Communications, 1995. **91**(1): p. 43-56.
42. Hess, B., et al., *GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation.* Journal of chemical theory and computation, 2008. **4**(3): p. 435-447.
43. Van Der Spoel, D., et al., *GROMACS: fast, flexible, and free.* J Comput Chem, 2005. **26**(16): p. 1701-18.

List of publications:

a) Published

1. Multimodal approach to explore the pathogenicity of BARD1, ARG 658 CYS, and ILE 738 VAL mutants. *Choudhary RK, Vikrant1, Siddiqui QM, Thapa PS, Raikundalia S, Gadewal N, Kumar NS, Hosur MV, Varma AK.J Biomol Struct Dyn. 2016 Jul;34(7):1533-44. doi: 10.1080/07391102.2015.1082149. Epub 2015 Sep 23.*

2. Structural and biophysical properties of h-FANCI ARM repeat protein. *Siddiqui MQ, Choudhary RK, Thapa P, Kulkarni N, Rajpurohit YS, Misra HS, Gadewal N, Kumar S, Hasan SK, Varma AK. J Biomol Struct Dyn. 2016 Sep 30:1-26. [Epub ahead of print].*

3. Functional basis and biophysical approaches to characterize the C-terminal domain of human-Ribosomal s6 Kinase. *Jagilinki BP, Choudhary RK, Thapa PS, Gadewal N, Hosur MV, Kumar S, Varma AK. Cell Biochem Biophys. 2016 Sep;74(3):317-25. doi: 10.1007/s12013-016-0745-6. Epub 2016 Jun 21.*

4. Mislocalization of BRCA1-complex due to ABRAXAS Arg361Gln mutation; *Vikrant1, Kumar R, Siddiqui Q, Singh N, Waghmare SK, Varma AK; J Biomol Struct Dyn. 2015;33(6):1291-301. doi: 10.1080/07391102.2014.945484. Epub 2014 Aug 8.*

5. Structural and Functional Implication of RAP80 Δ81Glu Mutation; *Vikrant1, Kumar R, Yadav LR, Nakhwa P, Waghmare SK, Goyal P, Varma AK., PLoS One. 2013 Sep 9;8(9):e72707. doi: 10.1371/journal.pone.0072707. eCollection 2013.*

6. Structural and functional characterization of the MERIT40 to understand its role in DNA repair; *Vikrant, Nakhwa P, Badgujar DC, Kumar R, Rathore KK, Varma AK. J Biomol Struct Dyn. 2014 Dec;32(12):2017-32. doi: 10.1080/07391102.2013.843473. Epub 2013 Oct 15.*

b) Manuscript communicated:-

RK Choudhary & Vikrant et al. Structural basis to stabilize the domain motion of BARD1-ARD BRCT by CstF50. Nature Scientific Reports; 2016. *Communicated*. (SREP-16-37456)

c) Manuscripts in preparation:-

1. Structural basis of BARD1BRCT and cancer predisposing mutation: A biophysical and molecular dynamics simulation approach

d) Others (Award for the poster / oral presentation if any)

- **Indian Biophysical Society (IBS) 2015, (Jamia Millia Islamia, India, Jan 2015).**
  Nature of Participation: Poster presentation (Title: Biophysical characterization of BARD1BRCT mutants). Awarded 1st prize for best poster presentation.

Biophysical characterization of BARD1BRCT mutants poster presentation delivered at the 10th Indian Biophysical society meeting New-Delhi, Jan 2015 and awarded 1st prize for best poster presentation.

Date:-1/2/2016

Signature of Student

Doctoral Committee:

S. No.	Name	Designation	Signature	Date
1	Dr. Vinay Kumar	Chairperson		3/02/16
2	Dr. Ashok Varma	Guide & Convener		03/02/16
3	Dr. Rajiv sarin	Co-Guide		03/02/16
4	Dr. Sanjay gupta	Member		03/02/16
5	Dr. Smita Mahale	Member		3/2/16
6	Dr. C.P Rao	Member		3/4/16

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

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

Prof. K. Sharma
Director, Academics
Tata Memorial Centre

PROF. K. S. SHARMA
DIRECTOR (ACADEMICS)
TATA MEMORIAL CENTRE,
PAREL, MUMBAI

ABBREVIATIONS

ACC1	: Acetyl Coenzyme A Carboxylase
ATM	: Ataxia Telangiectasia Mutated
ATR	: ATM and RAD3 related
ARD	: Ankyrin Repeat Domain
ATP	: Adenosine Triphosphate
BARD1	: BRCA1 Associated Ring Domain-1
BRCA1	: Breast Cancer susceptibility gene1
BACH1	: BTB and CNC homology 1
BLAST	: Basic Local Alignment Search Tool
BME	: β-mercaptoethanol
Bp	: base pairs
BSA	: Bovine serum albumin
CtIP	: CtBP Interacting Protein
Chk1	: Checkpoint kinase 1
CD	: Circular Dichroism
DLS	: Dynamic Light Scattering
DNA	: Deoxyribonucleic Acid
dNTPs	: deoxy Nucleoside TriPhosphate
EB	: Elution Buffer
EDTA	: Ethylenediaminetetraacetic acid
FFT	: Fast Fourier Transform
HR	: Homologous Recombination
FPLC	: Fast protein liquid chromatography
IMAC	: Immobilized Metal ion Affinity Chromatography

IPTG : Isopropyl-β-D-thiogalactoside

Kb : kilo base pairs

K_d : Dissociation constant

KDa : kilo Dalton

LB : Luria-Bertani

MALDI-TOF/TOF : Matrix Assisted Laser Desorption Ionization Time of Flight

MBP : Maltose binding protein

MS : Mass Spectrometry

MDC1 : Mediator of DNA Damage Checkpoint -1

NHEJ : Non-Homologous End Joining

NES : Nuclear Export Signal

NLS : Nuclear Localization Signal

NMR : Nuclear Magnetic Resonance

NPM : Nucleophosmin

MRE11 : Meiotic Recombination 11 homolog

MRN : Mre11-Rad50-Nbs1 complex

MD : Molecular Dynamics simulation

NTA : NitriloTriacetic Acid

OD_{600} : Optical Density at 600 nm

OmpT : Outer membrane protease

ORF : Open Reading Frame

PAGE : Poly-Acrylamide Gel Electrophoresis

PCR : Polymerase Chain Reaction

PDB : Protein Data Bank

pI : Isoelectric point

PPIs	: Protein-Protein Interactions
PARP1	: Poly (ADP-ribose) Polymerase 1
PCA	: Principal component Analysis
RNF	: RING finger containing protein
RAP80	: Receptor Associated Protein-80
RMSD	: Root Mean Square Deviation
TD	: Thermal denaturation
TEV	: Tobacco Etch Virus
TFA	: Trifluoro Acetic acid
T_m	: Melting Temperature
Top BP1	: Topoisomerase (DNA) II Binding Protein 1
UV	: Ultra Violet
UBC	: Ubiquitin Conjugating Enzyme
53BP1	: p53 Binding Protein

List of Tables

List of Figures

Figure 3.3.7	Cloning of CstF50 in pGEX-KT and pET-28a vectors
Figure 3.3.8	Expression and purification of CstF50
Figure 3.3.9	Refolding profile of CstF50
Figure 3.3.10	Ramachandran plot, and structural details of validated model for BARD1 ARD-BRCT
Figure 3.3.11	Ramachandran plot, structural details of validated model for CstF50 and domain organization of CstF50
Figure 3.3.12	Comparative NMA first mode profile of BARD1 ARD-BRCT *wild-type*, Gln 564 His and Comparative Deformation analysis of wild-type and mutant showing hinge bending.
Figure 3.3.13	Comparative RMSD and Rg profile of BARD1 ARD-BRCT (Black) and Gln564His (green) mutant and structure sampled at every 10ns for *wild-type*
Figure 3.3.14	Comparative aligned structure sampled at every 10ns for mutant Gln 564His, highest RMSD and input structure aligned for *wild-type* and mutant (Gln564His).
Figure 3.3.15	Comparative RMSD representation of RMSD of ARD and BRCT domain. ARD *wild-type*, mutant and BRCT domain *wild-type*, mutant.
Figure 3.3.16	RMSD, Rg, and RMSF attained by CstF50 during 100ns MD production
Figure 3.3.17	Structure representing the RMSF for BARD1 ARD-BRCT *wild-type*, mutant and CstF50
Figure 3.3.18	Projection of first three eigenvector for the structure BARD1 ARD-BRCT *wild-type*, mutant and CstF50
Figure 3.3.19	Cross-correlation for the structure BARD1 ARD-BRCT *wild-type*, mutant and cstF50
Figure 3.3.20	Residual fluctuation against first two eigenvector for the structure BARD1 ARD-BRCT *wild-type*, mutant and cstF50
Figure 3.3.22	Free-energy calculation for the structure BARD1 ARD-BRCT *wild-type* and CstF50 and Interacting interface for the BARD1 ARD-BRCT *wild-type* and cstF50 complex

Figure 3.3.23	Interacting interface for the structure (A) BARD1 ARD-BRCT *wild-type*, (B) Mutant and cstF50
Figure 3.3.24	Comparative RMSD and R_g of **(A and B)** (BARD1 ARD-BRCT *wild-type* (Black), Gln564His (green), CstF50 (blue) and complex (yellow) and **(C)** Hydrogen bonding profile for complex
Figure 3.3.25	Comparative cross-correlation for PCA of complex, BARD1 ARD-BRCT *wild-type* and CstF50
Figure 3.3.26	Free-energy calculation for the structure BARD1 ARD-BRCT *wild-type* CstF50 complex and representative structure of the BARD1 ARD-BRCTcstF50 complex
Figure 3.3.27	Distance calculation between ARD domain and BRCT domain for the structure BARD1 ARD-BRCT and BARD1 ARD-BRCT *wild-type*- CstF50 complex
Figure 3.3.28	Model for BARD1 ARD-BRCT and CstF50 complex formation
Figure 4.3.1	Expression and purification of BARD1BRCT and mutant proteins, FPLC purification profile of *wild-type* and mutants, Elution profile of standard proteins and stokes radii calculation of BARD1BRCT *wild-type* protein.
Figure 4.3.2	Comparative chemical crosslinking profiles of BARD1BRCT and mutant proteins
Figure 4.3.3	Secondary and tertiary structure characterization of BARD1BRCT domain and mutant protein using CD spectroscopy. Chemical denaturation profile of *wild-type*, Arg658Cys and Ile738Val and *wild-type*, Cys645Arg, Val695Leu and Ser761Asn using fluorescence spectroscopy.
Figure 4.3.4	Thermal denaturation of BARD1BRCT domain and mutant proteins using CD spectroscopy. Chemical denaturation profile of *wild-type* and mutants using fluorescence spectroscopy
Figure 4.3.5	Thermal denaturation of BARD1BRCT *wild-type* and mutant protein using CD spectroscopy in near-UV range

Figure 4.3.6	Thermal denaturation of BARD1BRCT *wild-type* and mutant protein using fluorescence spectroscopy
Figure 4.3.7	Comparative limited proteolytic profile of BARD1BRCT domain and mutant protein using trypsin and chymotrypsin w*ild-type*, Val695Leu, Ser761Asn and Cys645Arg respectively
Figure 4.3.8	Limited proteolytic profile of Arg658Cys and Ile738Val respectively using trypsin and chymotrypsin
Figure 4.3.9	Comparative representation of RMSD, RMSF, H-bond, Rg, Volume, and SASA. The *wild-type* is represented as black, Cys645Arg as green, Val695Leu as Blue, Ser761Asn as cyan, Arg 658 Cys as purple and Ile738Val yellow respectively.
Figure 4.3.10	Comparative average intra-molecular hydrogen bond formed by the *wild-type* and mutant structures.
Figure 4.3.11	Scree plot of mutant and *wild-type* structure
Figure 4.3.12	Projection of first three eigenvectors for the structure BARD1BRCT *wild-type*, Cys645Arg, Val695Leu, and Ser761Asn
Figure 4.3.13	Projection of first three eigenvectors for the mutant structure Arg 658 Cys and Ile738Val.
Figure 4.3.14	Residual fluctuation against first two eigenvectors for the structure BARD1BRCT *wild-type*, Cys645Arg.
Figure 4.3.15	Residual fluctuation against first two eigenvectors for the mutant structures Val695Leu and Ser761Asn
Figure 4.3.16	Residual fluctuation against first two eigenvectors for the mutant structures BARD1 Arg658Cys and Ile738Val.
Figure 4.3.17	Cancer predisposing mutations in different regions of BARD1BRCT domain.
Figure 4.3.18	Comparative representation of RMSF structure of *wild-type* and different mutants of BARD1. Red show regions of highest fluctuation and blue regions indicate areas of minimum fluctuations
Figure 4.3.19	Comparative cross-correlation for PCA of *wild-type*, Cys645Arg, Arg658Cys, Val695Leu, Ser761Asn and Ile738Val mutant

Figure 4.3.20	Free-energy calculation for the structure BARD1BRCT *wild-type* and mutant structures
Figure 4.3.21	Intra-molecular interaction analysis for the mutant structures (A, B) Cys645Arg, Val695Leu and Ser761Asn.
Figure 4.3.22	Comparative intra-molecular interaction analysis for the mutant structures BARD1BRCT Arg658Cys and Ile738Val
Figure 5.3.1	Cloning of BRCA1 and BARD1 Ring domain region. PCR profile of BARD1 Ring domain and clones showing insert release. PCR profile of BRCA1 Ring domain and clones showing insert release
Figure 5.3.2	Expression of BARD1 Ring domain region. Expression profile of BARD1 Ring domain and FPLC purification profile
Figure 5.3.3	Model structure BRCA1 Ring domain (1-125aa)
Figure 5.3.4	Hydrogen bonding and hydrophobic interactions between BRCA1 and cisplatin complex model structure of cisplatin and BRCA1. The dotted lines indicated H-bond and brush lines indicates hydrophobic interactions
Figure 5.3.5	Intermolecular analysis and complex structure of transplatin and BRCA1. The dotted lines indicated H-bond and brush lines indicates hydrophobic interactions

Chapter 1

Introduction & Review Of Literature

1.1 DNA damage and Cancer

In nature, every form of life preserves its genomic stock to ensure the high-fidelity communication of genetic information by repairing its damaged genetic material. To achieve very high accuracy, a plethora of evolved molecular systems including sensors, mediators, and effectors molecules continuously detect DNA damage, signal its presence, and mediate the damage repair. Living cells in an organism have advanced mechanisms for monitoring genome integrity and respond to DNA damage by triggering multifaceted DNA-damage-response pathway. DNA-damage-response is a result of cross-talk between different pathways like cell-cycle arrest, transcriptional and post-transcriptional activation of genes including those associated with DNA damage repair. DNA damage may arise during normal cellular metabolic processes like DNA replication, free radical generation and exposure to UV or ionizing radiations. A sensor molecule detects the DNA damages, and transduces the damage signals through mediator molecules to terminal effector molecules, which activate the cell-cycle checkpoints and recruits DNA repair proteins to the damage site. Inability to respond properly to damage and repair, lead to genetic instability, and an ablation in the genetic content which in turn may enhance the chances of cancer development [44, 45]. Cancer is intimately related to the accumulation of DNA damage, and repair fiascos. However, due to advancement in functional genomics, it is now becoming progressively clear that deregulated DNA-damage signaling and repair pathways is essential to the etiology of the most of the human cancers. However, it is understood that highly conserved DNA-repair and cell-cycle checkpoint pathways allow cells to handle both endogenous and exogenous causes of DNA damage. Furthermore, it is now accepted that inherited or acquired deficiencies in genome conservation system contribute ominously to the commencement of cancer [1]. Mammalian

genome contain proto-oncogenes that function to regulate normal cell proliferation and differentiation [1, 46]. Oncogenes activation play central role in the predisposition to cancer risk. In cancer cells, oncogenic protein level is frequently up or down-regulated due to the acquired mutations [1]. The after effects of these mutations, act through different mechanisms, including gene amplification that leads to increased copy number of oncogenes and chromosomal translocation which replace the original promoter with a constitutively active one [47, 48]. The cellular response to DNA damage involves change in the localization of a number of nuclear proteins [13]. It is now clear that Double strand DNA damage breaks (DSBs) lead to the triggering of an intricate regulatory reaction that impinges on a variety of cellular measures, including cell-cycle control, apoptosis, transcription, DNA replication, and telomere maintenance. Several proteins like RAD family of proteins have been anticipated as DNA damage sensors [20]. However, the identity of sensor molecules and their intermingled molecular interplay, which leads to DNA damage repair response and role in cancer, is yet to be fully understood.

It is well documented that mutations in the tumor suppressor proteins such as p53, ATM, BRCA1 (breast cancer and ovarian cancer susceptibility gene 1) and BRCA2 have been linked to breast and ovarian cancer**.** Breast cancer develops from a heterozygous population of mutated cells in the breast. Breast cancers and neoplasms are closely associated with DNA damage repair defects in cell-cycle checkpoints, which allow damaged DNA to go unrepaired. Different proteins of the homologous recombination repair pathway are binding partner of BRCA1 and BRCA2. However, it is still a challenge to understand, how mutations in BRCA1 and BRCA2 lead to higher chances of predisposition to breast and ovarian cancers.

1.2 Breast Cancer and BRCA1

Breast cancer is one of the most commonly diagnosed cancer after melanoma (skin cancer), and it is the second leading cause of cancer demises after lung cancer [49]. Breast cancer has been categorized into four clusters based on their tissue origin. These are ductal carcinoma, which is also known as intraductal carcinoma, lobular carcinoma, invasive (infiltrating) ductal carcinoma, and invasive (infiltrating) lobular carcinoma [50]. Most breast cancer cases are sporadic and only 5 to 10% are genetically linked [51]. In response to endogenous and exogenous sources of DNA damage, elementary signal transduction pathways involve majorly three proteins, BRCA1, BARD1 and PARP-1. Hereditary breast cancer arises due to inherited genetic alterations in the tumor suppressor genes BRCA1 and BRCA2 [52]. BRCA1 was mapped by genetic linkage on to the p-arm of chromosome 17 [53]. During early 1990's it was first reported that germline mutations in BRCA1 are associated with early-onset of breast and ovarian cancer [53].

BRCA1 is implicated in various cellular processes like transcriptional regulation, cell-cycle control, and cellular response to DNA damage [54-57]. BRCA1 is a large protein comprising of 1863 amino acids, and well characterized into different domains such as N-terminal RING domain, DNA binding domain, transactivation domain and C-terminal BRCT domain (**Figure:- 1.1**). Co-Immunoprecipitation studies have revealed that BARD1 is an obligatory binding partner of BRCA1 [58]. 90% of the cellular BRCA1 exist as a BRCA1-BARD1 heterodimer. BRCA1 associated ring domain (BARD1) comprises of 777 amino acids **(Figure: - 1.2).** BARD1 is characterized by N-terminal RING domain and conserved BRCT domain which shares a great deal of structural similarities to RING domain and BRCT repeats of BRCA1 [59].

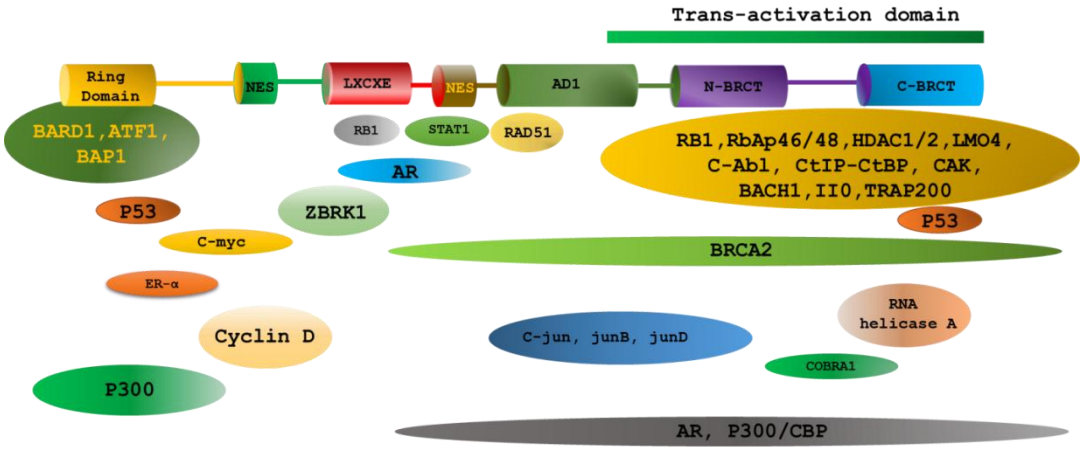


Figure: - 1.1: - Domain organization and interacting partners of BRCA1

In addition, BARD1 also possesses four ankyrin (ANK) repeats from 425-550 amino acids that are highly conserved in other cellular proteins [60-62]. Ankyrin repeats are reported to act as protein-protein interactions motif. In view of the much lower levels of sequence and structural homology for the rest of the protein (BARD1), it is reported that ARD and BRCT domains facilitate indispensable functions for BARD1.

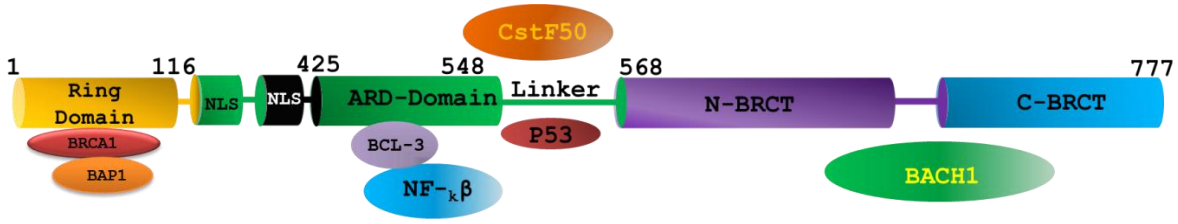


Figure 1.2: - Domain organization and interacting partners of BARD1

BRCA1 and BARD1 heterodimerizes *in-vivo* and *in-vitro* by interacting through its RING-RING domain and forms active E3 ubiquitin ligase complex which ubiquitinates several cellular proteins and has an essential role in tumor suppression **(Figure: - 1.3)** [63-65].

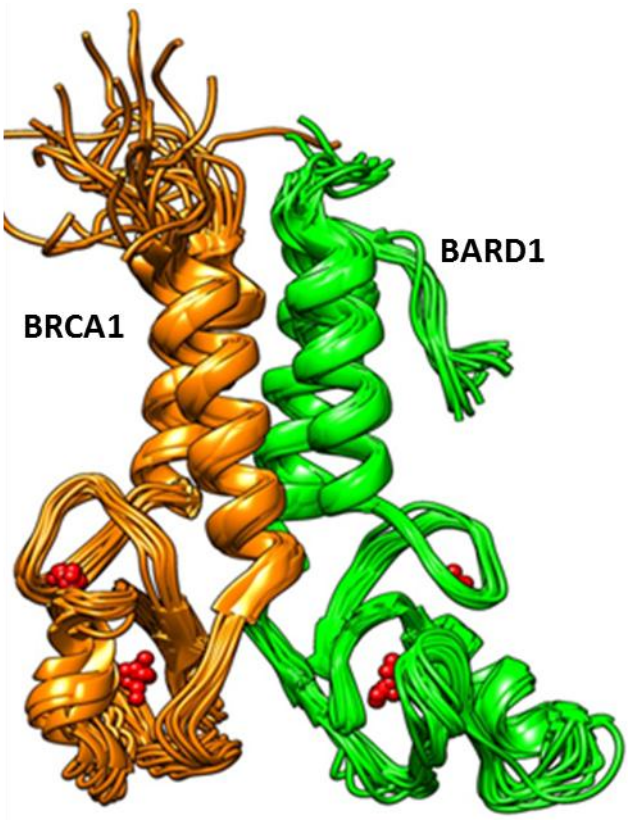


Figure 1.3: - NMR solution structures of BRCA1-BARD1 RING-RING domain E3 ubiquitin ligase complex (PDBID:-1JM7)

Ubiquitination is a traditionally conserved pathway known for degradation of proteins in cells. A wider view suggests that ubiquitination is an important posttranslational modification which includes proteasomal degradation, targeting proteins to their working niche in the cell, and regulating their activity [66].

Ubiquitins are small poly peptides of 76 amino acids attached to target molecule with the help of E3-Ubiquitin ligase **(Figure:- 1.4) [67, 68]**.

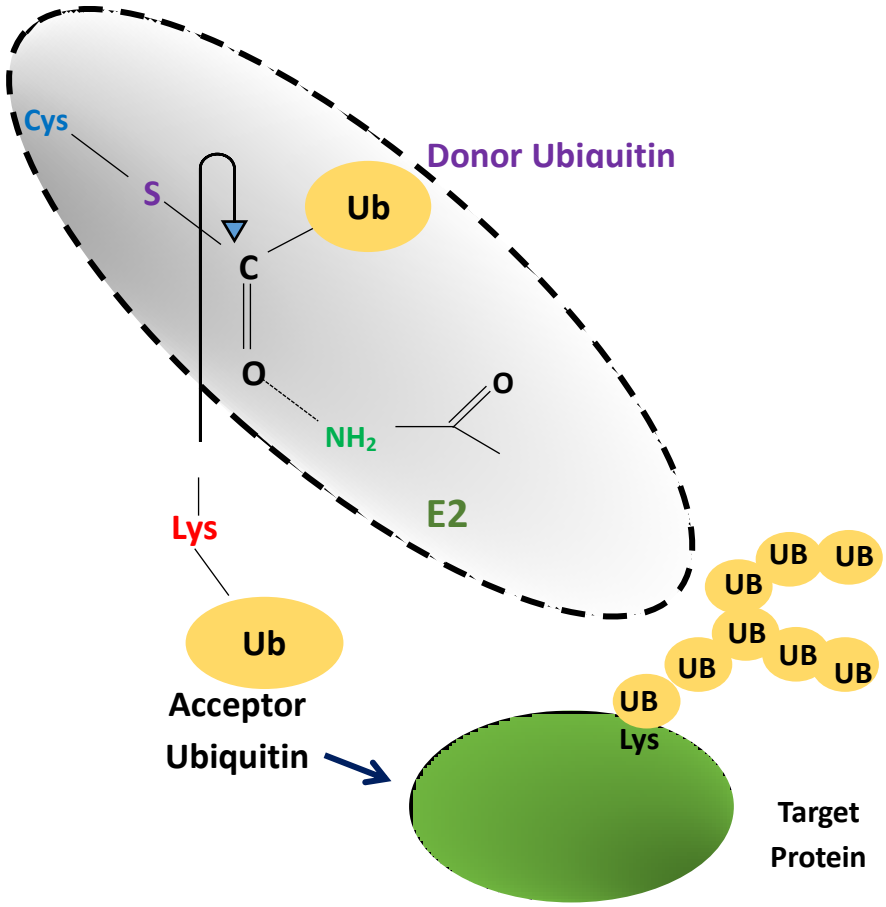


Figure 1.4: - Mechanism of ubiquitination in protein

Earlier studies have established that BRCA1-BARD1 acts as E3 ubiquitin ligase *in-vitro* [21, 69, 70]. The BRCA1-BARD1 heterodimer specifically mediates the formation of 'Lys-6'-linked polyubiquitin chains and coordinates cellular pathways such as DNA damage repair, ubiquitination and transcriptional regulation to maintain genomic stability [71]. BRCA1-BARD1 E3 ubiquitin ligase complex plays a central role in the cell-cycle check point control

in response to DNA damage. The BARD1-BRCA1 interaction is ablated by amino acid substitutions in BRCA1 ring domain, implying that the formation of a stable complex between these proteins may be a critical aspect of BRCA1 facilitated tumor suppression [63, 64]. A tumor associated mutation of BRCA1 Cys61Gly within the RING domain leads to loss in binding to BARD1, thereby suggesting a role of BARD1 in mediating tumor suppression [72, 73]. The BRCA1 Cys61Gly is devoid of ubiquitination activity, supports that a complex of BRCA1 and BARD1 mediates ubiquitination and tumor suppression [63].

In BRCA1-BARD1 E3 ubiquitin ligase pathway the activity of the BRCA1–BARD1 ubiquitin ligase leads to RNA Pol II degradation [12, 13, 26], cell-cycle arrest, γ-tubulin degradation [7, 74, 75], control of centrosome duplication, H2A/H2AX ubiquitination [69, 76], epigenetic control and NPM ubiquitination and stabilization [77]. BRCA1 exists as a part of BASC (BRCA1-associated genome surveillance complex) and because of ATM/ATR-dependent phosphorylation of BRCA1, it generates binding sites for its associated factors [78]. Furthermore, BRCA1 phosphorylation by CHK2 at Ser988 is essential for HR activity [79], G2/M cell-cycle checkpoint activation and suppression of genotoxicity induced mammary and uterine cancer [80]. In response to DNA damage, BRCA1 interacts with MRN complex in checkpoint kinases ATM or Chk2 dependent manner [81]. MRN (Mre11-Rad50-Nbs1) complex intervenes the G2/M cell-cycle checkpoint signal and plays roles in DSB repair by holding two DNA ends close to each other, which is required for non-homologous end-joining function and BRCA2-Rad51 mediated HR [82]. BRCA1 acts in a genome surveillance complex by specifically interacting with hyperphosphorylated processive RNA polymerase II (IIO), rather than hypophosphorylated RNA- polymerase II (IIA) found at promoters [15, 83].

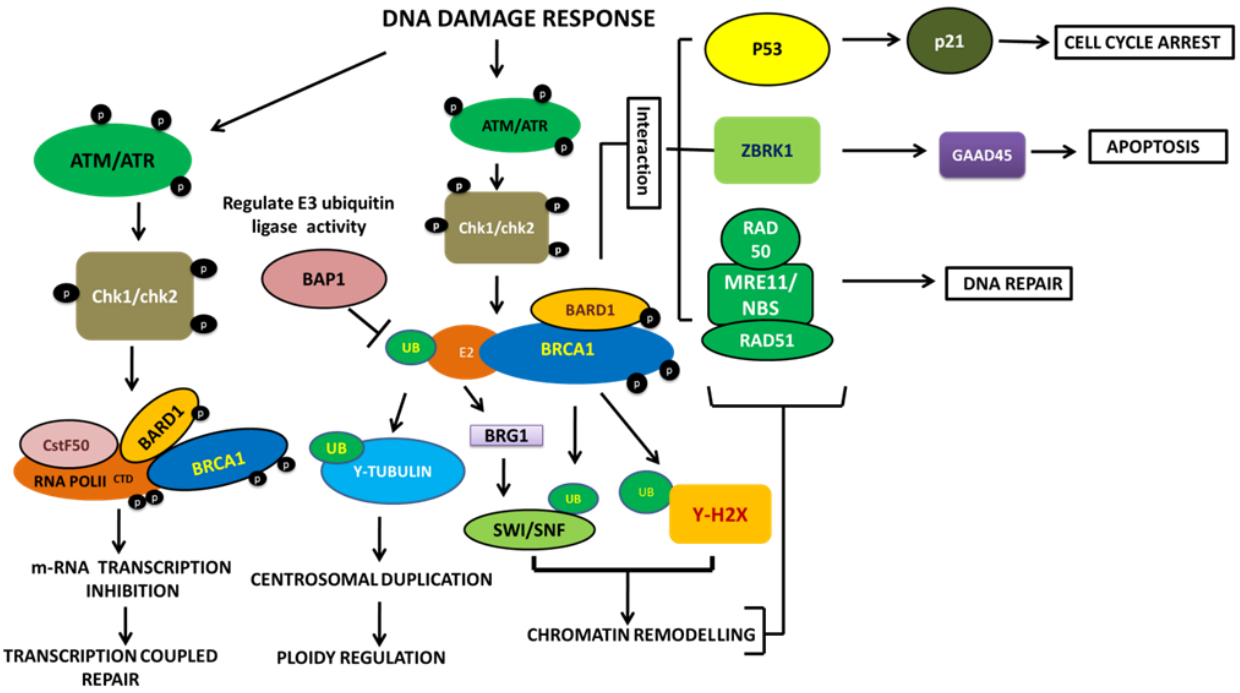


Figure 1.5:- Role of BRCA1-BARD1 E3-ubiquitin ligase complex in cellular pathways.

BRCA1 is known to associate with RNA polymerase II (Pol II) holoenzyme, which supports recent finding of cessation of transcription at the site of DNA damage leading to the BRCA1–BARD1 mediated ubiquitination of the Pol II holoenzyme [84, 85]. BRCA1 is a nuclear protein and its transport to the nucleus is facilitated by BARD1 whose degradation depends on the presence of RING structure. The full-length BARD1 might be exported and degraded while its excess is translocated to the nucleus. The BARD1 binding masks the nuclear export signal of BRCA1, thereby retaining it in the nucleus. BARD1 and BRCA1 are also required for localization of each other to the DNA damage induced foci [86]. Ubiquitination is a regulatory

mechanism that controls shuttling from nuclear to cytoplasm and vice versa, e.g. the nuclear export of p53 [45]. BRCA1 regulates a large number of chromatin remodeling activities through its activation domain 1 (AD1) and the C-terminal tandem repeats (BRCTs).

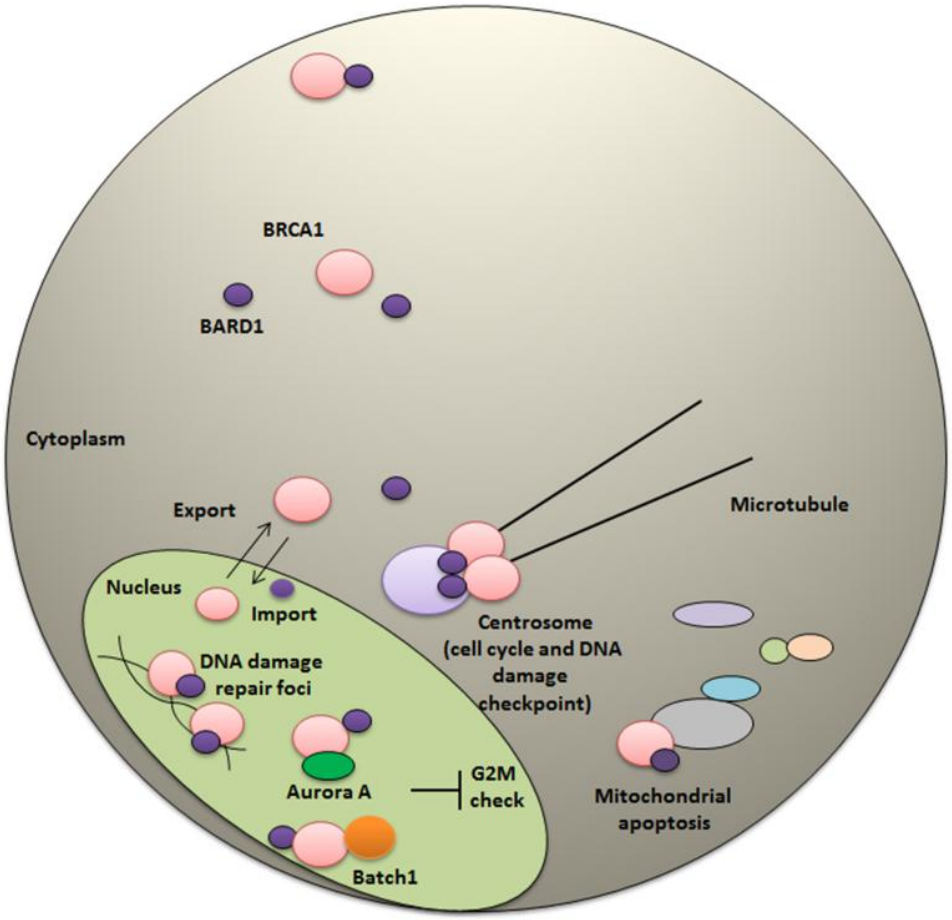


Figure 1.6:- Mechanism of nuclear import and export of BRCA1-BARD1 E3-ubiquitin ligase complex.

BRCA1 remodeling activities do not require histone acetylation, and are mediated through the BRCA1-dependent recruitment of a cofactor of BRCA1 (COBRA1), also called NELF-B [87]. NELF-BRCA1 complex binds to the estrogen receptor-alpha (ER-α) and down regulates its

transcription activity, thus contributing towards BRCA1-mediated repression of ER-α activity [88]. Hence, it acts as a sensor of microtubule disorganization elicited by various microtubule dis-assembly agents such as drugs, and promotes apoptosis through JNK activation [89].

The BRCA1-BARD1 complex plays a central role in the homologous recombination repair process. In the early Homologous recombination repair (HR), ATM/ATR phosphorylated γ-H2AX recruits MDC1 and RNF8/UBC13 ubiquitin ligase, that ubiquitinates the γ-H2AX. Ubiquitinated γ-H2AX interacts with UIM repeat of RAP80 and binds to ABRAXAS [90, 91]. The pS-X-X-F ($_{p}$S-T-X-F) binding motif at the C-terminal of ABRAXAS interacts with BRCA1-BRCT phospho-peptide binding domain, thus achieving foci formation [92].

The residues Ser404 and Ser406 positions of ABRAXAS are phosphorylated by ATM/ATR after exposure to Ionizing radiation (IR) and are required for the phosphodependent interaction between ABRAXAS and BRCA1 BRCT domains **(Figure :- 1.7)** [49, 50]. Further, BRCA1 surveillance complex in concert with RAD51, BRCA2, and XRCC3 initiates the homologous recombination repair [78, 93, 94].

1.3 BRCA1 independent function of BARD1

BARD1 is an interacting partner of p53, and facilitates p53 dependent apoptosis and tumor suppression by stabilizing p53 and activating caspase-3 [25, 95]. Hence, BARD1 has BRCA1 independent but p53 dependent tumor suppressor function [95]. Functional aspects of BARD1 interacting partners like CstF50, p53 and ATM are not well studied [25, 36]. In BRCA1 independent pathway, expression of cellular BARD1 has been shown to be up regulated by DNA damage mediated by hormone signaling, exposure to ultraviolet light, and hypoxia [95].

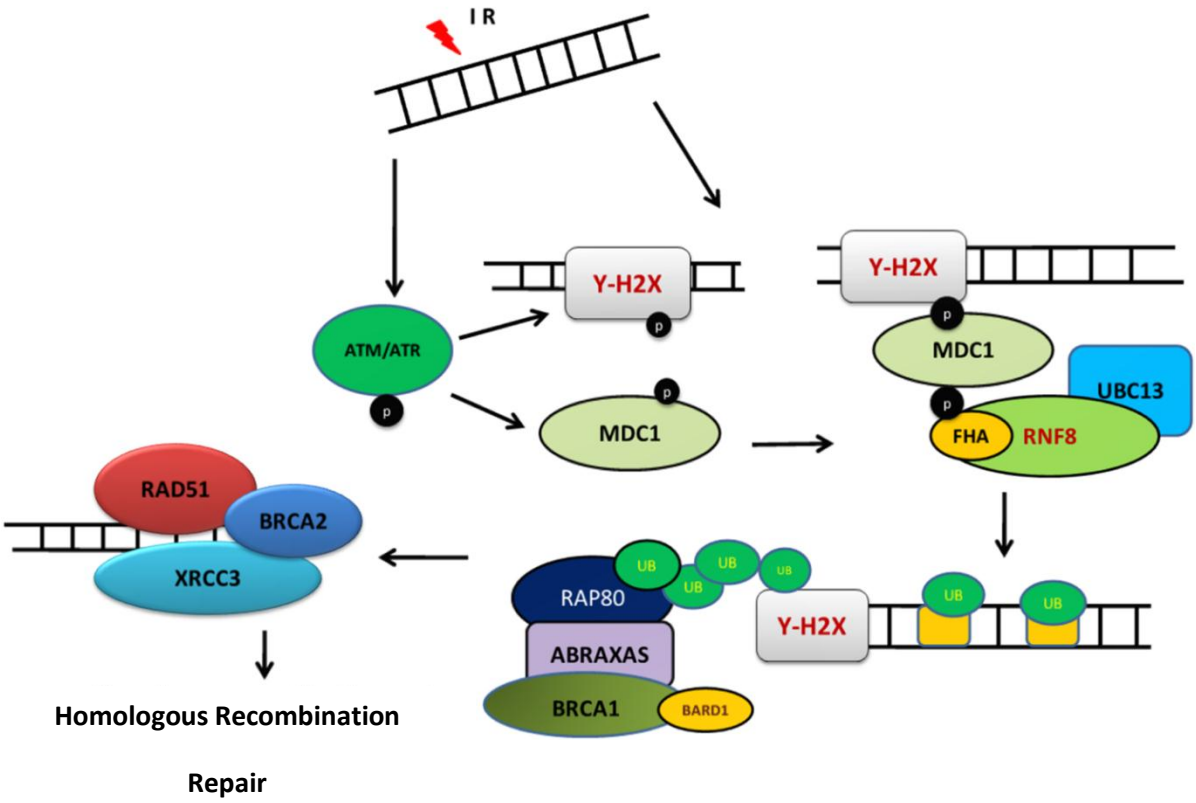


Figure 1.7:- Role of BRCA1-BARD1 E3-ubiquitin ligase complex in homologous recombination repair.

The molecular mechanism of BARD1 mediated phosphorylation of p53 by ataxia telangiectasia mutated (ATM) is still unknown. BARD1 up-regulation under genotoxic stresses stabilizes p53 by facilitating its phosphorylation at Ser15 position by ATM kinase which leads to the interaction of p53 with Ku-70, a subunit of DNA-dependent protein kinase (DNAPK) [96]. DNAPK plays an important role in DNA damage repair (DDR) by non-homologous end-joining

(NHEJ) to maintain genomic stability [96]. The inter-domain region of ankyrin repeat and BRCT domain of BARD1 interacts with p53.

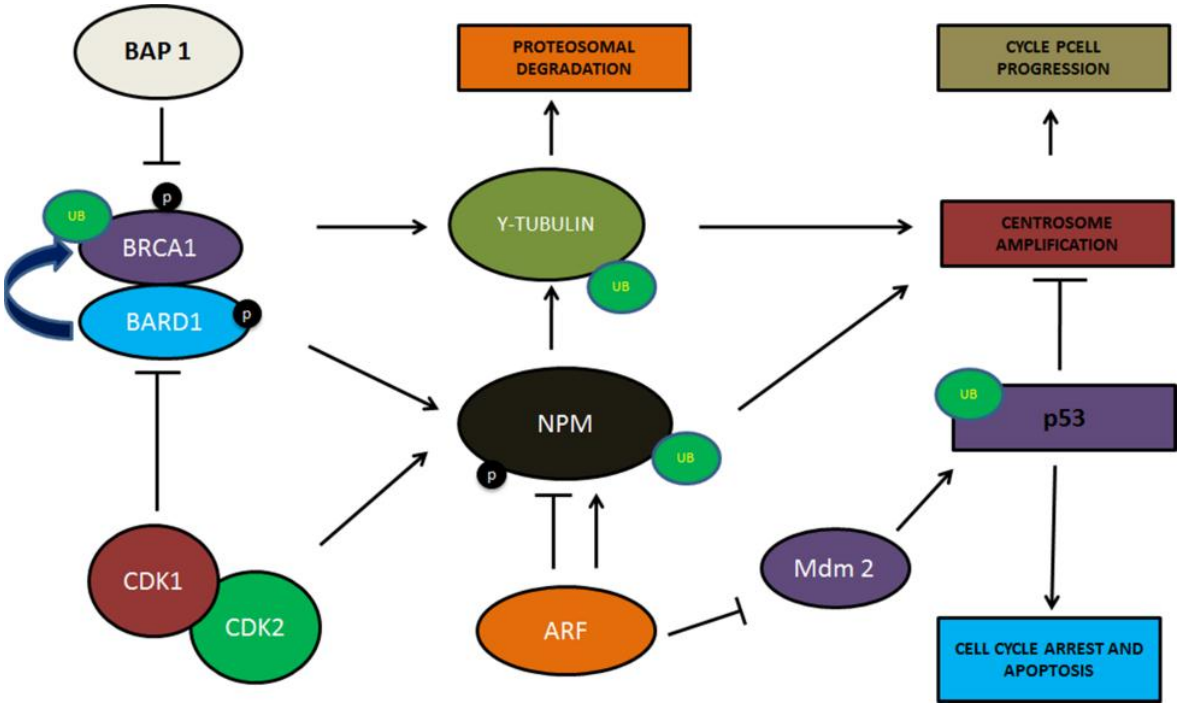


Figure 1.8:- Regulation of BRCA1-BARD1 E3 ubiquitin ligase complex.

However, this process is independent of BARD1 RING-finger domain because even deletion mutants lacking this domain co-immunoprecipitated with p53 [96]. Hence, BARD1 alone has apoptosis inducing capability in a p53 dependent manner [95]. It is believed that CDK2–cyclin complexes like CDK2–cyclin A1/E1 and CDK2–cyclin B1 regulate the interactions between BARD1 and BRCA1 and trigger its mitotic activity in phosphorylation dependent manner [97] (**Figure: - 1.2)**. Interestingly, ubiquitination of NPM *in-vivo* and BRCA1 auto-ubiquitination are disrupted by co-expression of CDK2–cyclin A1/E1, but not by CDK2–Cyclin B1 (**Figure: - 1.8)**.

BARD1 also showed transcriptional activation of NF-κB through binding to BCL3, a NF-κB co-factor. However, BCL3 does interact with BRCA1 and whether its activation mechanism is BARD1 dependent or not, has not been elucidated clearly [98]. BARD1 is

susceptible to proteolytic cleavage by caspase dependent protease caplin, and proteolytic products are found to be immunogenic and also suspected to have anti-tumorigenic properties like BCL3 [61, 98]. Another BRCA1-independent function of BARD1 is interaction with NF-κB and modulation of its transcriptional activity via Bcl-3 [99].

1.4. Regulation of BRCA1/BARD1 E3 ligase function

Sumoylation is one of the mechanisms of fine-tuning in regulatory pathways. Sumoylation enhances the E3 ubiquitin ligase activity of BRCA1-BARD1 complex. BRCA1 sumoylation mutants Lys199Arg and Asp121Arg have shown drastically reduced ubiquitination affinity and foci formation [100]. Furthermore, BRCA1-BARD1 E3 ubiquitin ligase activity is found to be inhibited by phosphorylation of DNA damage response molecule CDK1/CDK2, ATM, ATR, CHK1/2 in vivo **(Figure:- 1.8)** [79, 97, 101-105]. UBXN1 binds to BRCA1 through its N-terminal UBA domain, whereas the C-terminal domain binds to BRCA1-BARD1 heterodimer in an ubiquitin independent manner [106]. BRCA1 Associated Protein 1 (BAP1) is a novel nuclear UCH–type (ubiquitin carboxyl-terminal hydrolase) deubiquitinating enzyme, which interacts with the RING domain of BRCA1 (**Figure: - 1.9**) [107].

BAP1, like other deubiquitinating proteases, catalyses the hydrolysis of the isopeptide linkage joined by the COOH-terminal glycine of ubiquitin and lysine side chain of substrate. This leads to removal of single ubiquitin moieties from ubiquitin chains or cleavage of the isopeptide bond between ubiquitin and the substrate protein [108]. BAP1 has ubiquitin COOH-terminal hydrolase (UCH) domain at N-terminus, two nuclear localization signals (NLS) and a host cell binding motif (HBM) [94]. BAP1 was first identified as a BRCA1 associated protein, which binds to the RING finger domain of BRCA1 and BARD1 (**Figure: - 1.9**) [107, 109]. BAP1 binds to *wild–type* BRCA1 but not to mutated BRCA1 RING finger domain and

enhances the growth suppression properties of BRCA1. Binding with zinc finger RING domain of BRCA1 suggests that BAP1 may serve as regulator/effector of BRCA1 mediated growth control and differentiation pathways. BAP1 region from 182-365 amino acids binds to BARD1 RING finger domain [110].

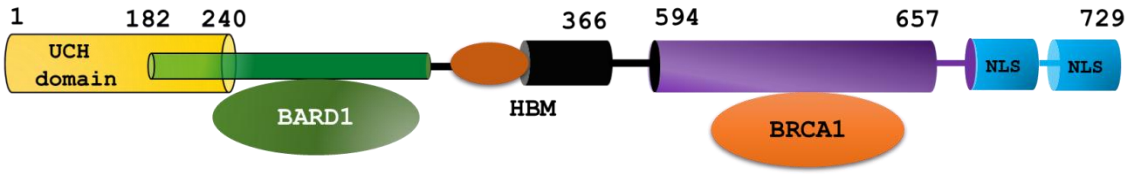


Figure 1.9:- Domain organization and interacting partners of BAP1

It has been reported that BAP1 interferes with BRCA1-BARD1, RING-RING domain E3 ligase activity by binding to BARD1 [84]. This suggests that BRCA1-mediated ubiquitination and BAP1-mediated deubiquitination may coordinately function in these cellular processes. The results show that a BAP1 directly inhibits E3 ligase activity of BRCA1-BARD1 complex [84]. BAP1 mutations, having correlation with metastatic uveal melanomas of eye, has also shown to reduce BARD1 binding [111]. Hence, BAP1 is likely the key regulatory factor behind modulation of BRCA1-BARD1 E3 ubiquitin ligase mediated tumor suppressor activity, and thus reflects its potential to control ubiquitin-dependent biological processes [112]. Furthermore, it has been recently observed that HECT type E3 ubiquitin ligase HERC2 interacts with BRCA1 and ubiquitinates BRCA1 unbound to BARD1 [113]. BRCA1-BARD1 interaction serves to protect BRCA1 from HERC2 [113]. Hence, it can be concluded that HERC2 have some role in the cell-cycle checkpoint control by BRCA1 degradation via ubiquitination [114].

1.5 BRCA1-BARD1 E3 ubiquitin ligase complex inhibited by platinum drugs

Deregulation of fundamental cell survival pathways is the prime cause of genetic instability and cell death. Recent advancements in cancer therapy involve altering pathways that are vital for cell survival, and particularly to target pathways involved in maintenance of genomic integrity [115]. As discussed earlier BRCA1 has critical role in the DNA damage repair. Several clinical and preclinical studies have demonstrated that modulation of BRCA1 function has proven to be a promising therapeutic solution for breast and ovarian cancer [116-118]. Dysfunctioning the E3 ubiquitin ligase activity of BRCA1 can be considered as a potential therapeutic target [119, 120].

Recent report suggests that tumor suppressor activity of BRCA1-BARD1 E3-ubiquitin ligase complex is inhibited by platinum compounds like transplatin, cisplatin, oxaloplatin and carboplatin [23]. This inhibition occurs in decreasing order of their efficiency *in-vitro* as well *in-vivo* respectively and making tumor cells more sensitive to chemotherapeutic drugs. Binding of anticancer drug cisplatin has been shown to affect the Apo-conformation of the BRCA1 RING finger domain. Mass spectrometric analysis shows that cisplatin binds to BRCA1 RING finger domain forming inter-molecular as well as intra-molecular adduct at His117 position of the RING domain [121]. Hence, platinum compounds or metal-based drugs, which principally target BRCA1 E3 ubiquitin ligase activity, can be significantly helpful in improvising the combinatorial therapy and efficacy of anti-breast cancer drugs **(Figure: - 1.10**).

Hence, to understand the mechanism of E3 ubiquitin ligase activity inhibition of BRCA1-BARD1 in *vitro* or in-vivo, their protein-inhibitor complex structure should be determined to elucidate the mechanism of inhibition by platinum compounds.

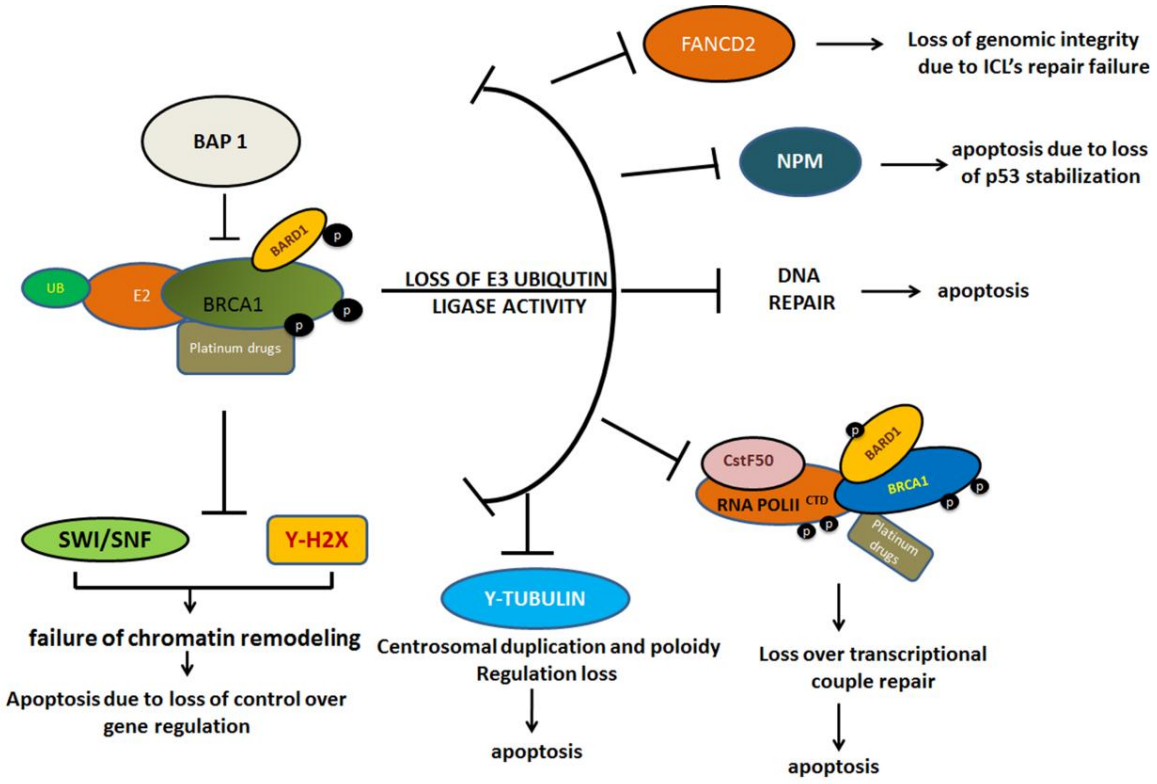


Figure 1.10:- BRCA1-BARD1 E3 ubiquitin ligase complex inhibition by platinum and its effects

1.6 3'-cleavage and polyadenylation are closely coupled to the transcription

The poly (A) tail is characteristic of eukaryotic mRNAs and plays an imperative role in regulation of mRNA stability, transportation, and translation [122-124]. In humans, ~75% of the core poly (A) sequences contain a conserved AAUAAA or AUUAAA upstream

hexameric motif located upstream of the poly(A) site and U or GU-rich sequence element (DSE) [125, 126].

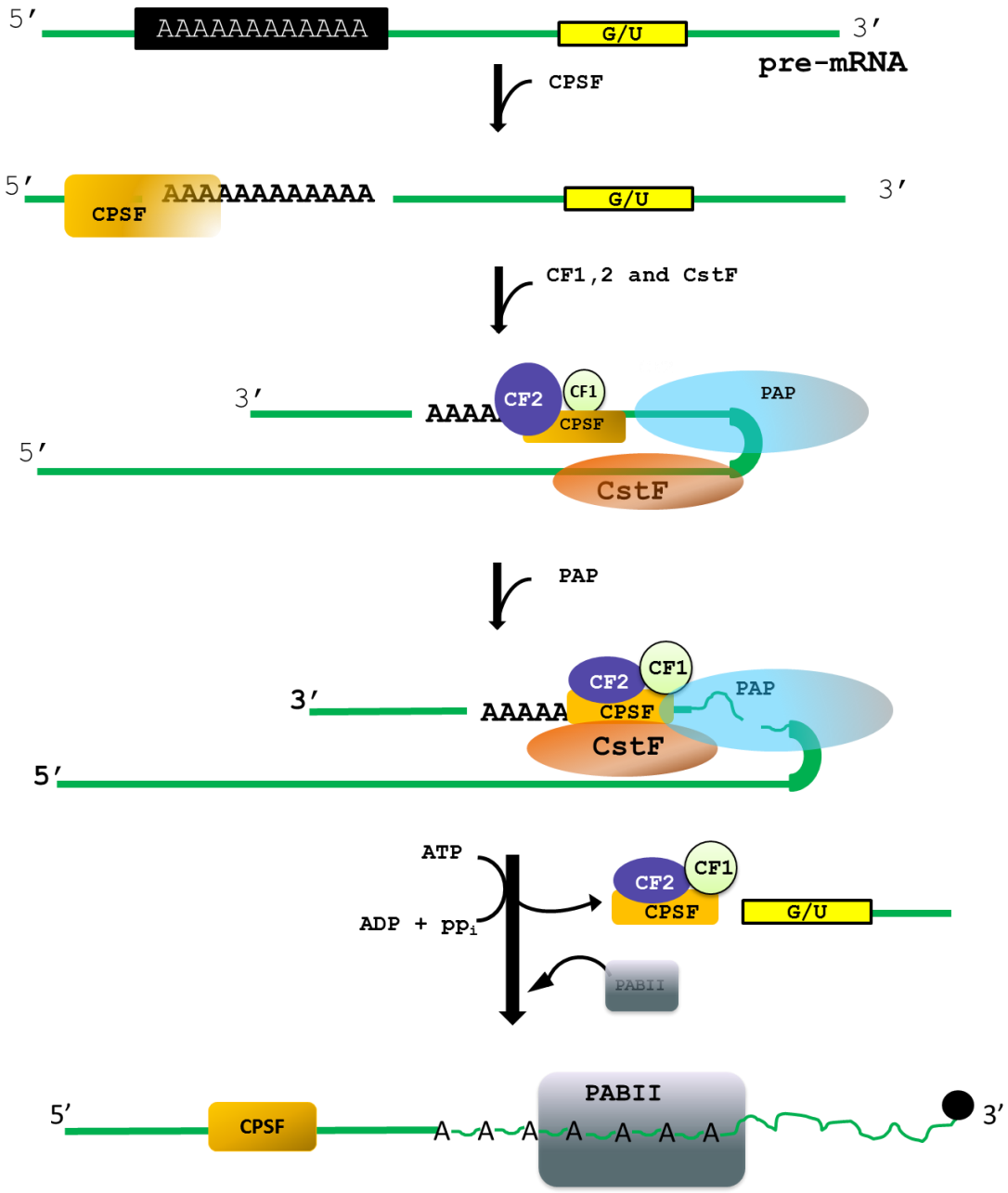


Figure 1.11:- Molecules involved in mRNA processing and polyadenylation *in-vitro*

This hexameric sequence is recognized by the Cleavage stimulating Factor (CstF) [127-129]. Cleavage stimulation factors (CstF), heterotrimeric in nature CstF50, CstF64, CstF77) and are responsible for processing of 3'end of premature mRNA transcript generated by RNA polymerase II.

CstF50 (Cleavage stimulating factor 1) is involved in the processing of premature transcript and has also been implicated in direct interaction with the RNA polymerase II CTD through its N-terminal region [130, 131]. CstF-50 is a WD (Tryptophan Aspartate) 40-repeat containing protein, interacts with the carboxyl-terminal domain (CTD) of RNA polymerase II (Pol II). N-terminal of CstF50 has predominant helical regions, followed by a connecting linker having high flexibility loop region.

1.7 CstF50 structure and domain organization

CstF50 belongs to the highly conserved family of WD40 repeat proteins which contains 7 WD40 repeat domains spanning C-terminal region (92-431aa) comprising of beta-sheets and folding in to a propeller like structures **(Figure:- 1.13)**.

The WD repeat in CstF50 comprises of ~44-60 residue sequence that typically contains the GH dipeptides at the N-terminus, and the WD di-peptides at the C-terminus. In a homodimer of CstF50, N-terminal domain is a repertoire of three α-helical segregation and WD40 repeat, which collectively bury >1500 $Å^2$ of surface area per monomer. Gβ subunit of GPCR is also a WD repeat protein which shows very high structural homology with CstF50 model structure **(Figure: - 1.12)**.

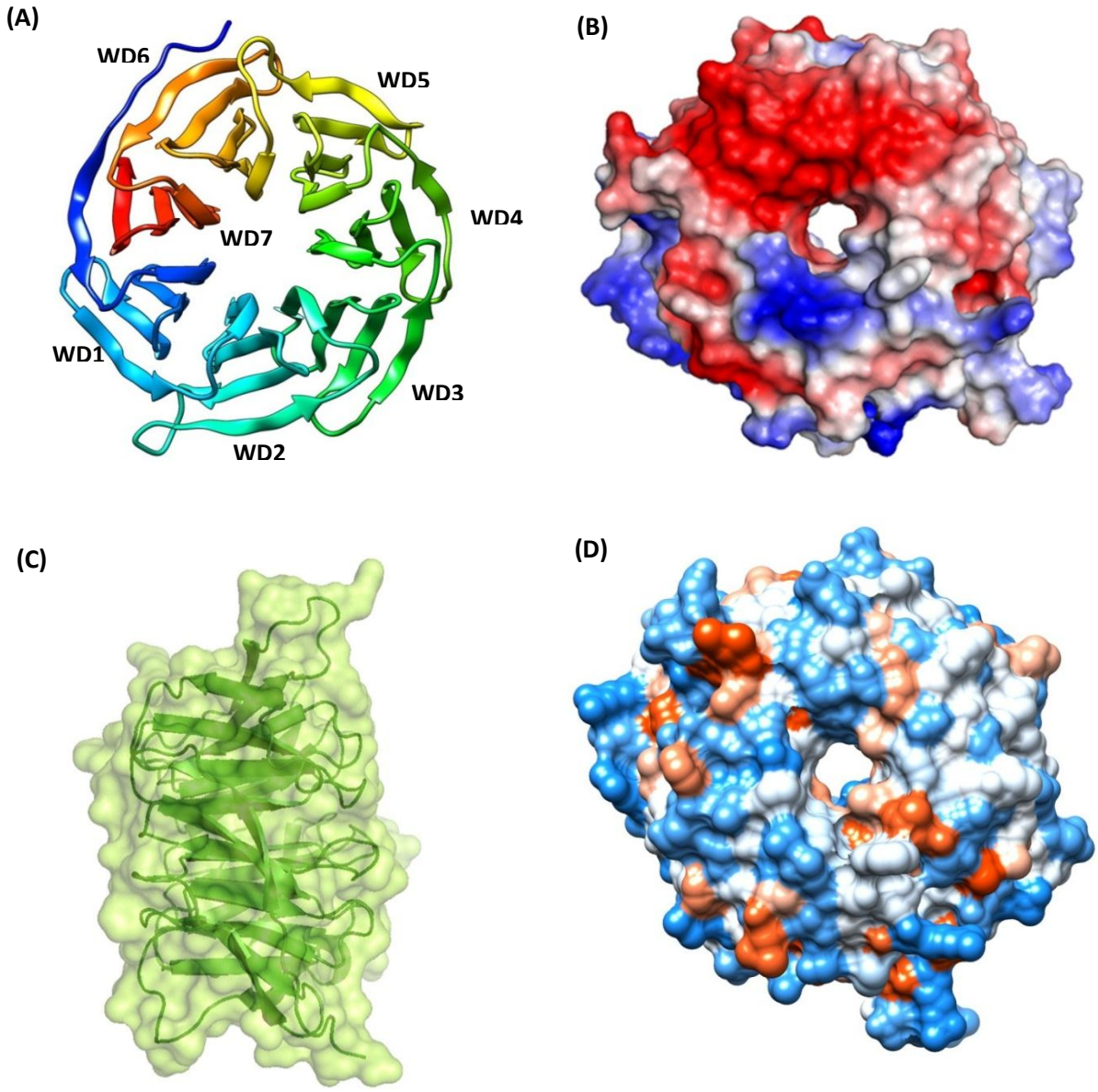


Figure 1.12:- 7 WD40 repeat protein β-subunit of G-protein. **(A)** Secondary structure **(B)** Electrostatic surface **(C)** Secondary merged with surface and **(D)** Solvent accessible surface view (PDBID: 1XHM).

CstF50 protein has an important role in the processing of premature mRNA transcripts and polyadenylation by interacting with BARD1 ankyrin repeat domain through its WD40 repeats region [12, 36, 131]. BARD1 interacts to CstF50 through its flexible linker region consisting of

554-568 amino acids (**Figure:-1.2**). BARD1 and the cleavage-stimulating factor 1 (CstF 1, CstF50) are required for the (UV)-induced inhibition of the mRNA 3' cleavage polyadenylation reaction [25]. Interactive relationship between BRCA1-BARD1 and CstF50 has been found to be important in Transcription Coupled Repair (TCR) [6, 13, 26]. Ionizing radiation activates TCR response through ATM/ATR kinases which phosphorylate their substrate ChK1/ChK2 and regulate cell-cycle and DNA damage repair. These results indicate that BRCA1-BARD1-CstF50 mediated inhibition of polyadenylation may prevent post-transcription defective pre-mRNA processing [12]. Treatment of cells with hydroxy-Urea leads to inhibition of transcription, suggesting that BARD1 is an important factor with CstF50 in inhibition of polyadenylation of nascent mRNA [72, 93]. BARD1 forms heterodimer with CstF50 to modulate mRNA processing and RNAP II stability, by inhibiting pre-mRNA 3'end cleavage [6, 7, 12, 13]. Depletion of CstF50 enhances sensitivity to UV treatment and reduces the ability of BRCA1-BARD1 to ubiquitinate RNAP II and repair DNA damage, which leads to cell-cycle arrest and apoptosis [132].

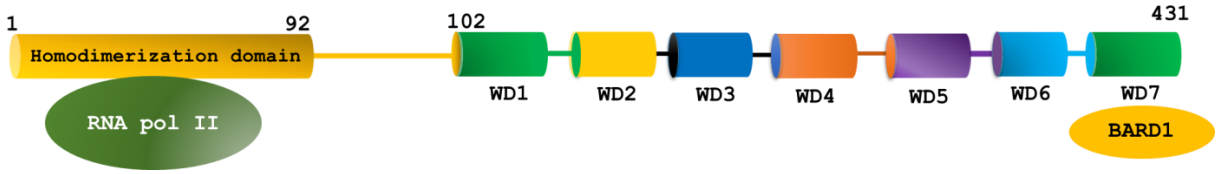


Figure 1.13:- Domain organization of CstF50.

CstF50 also interacts with the DNA replication and repair factor PCNA [12]. PCNA co-localizes with BRCA1-BARD1 at sites of DNA repair [72, 78], and associates with DNA repair proteins as part of the TCR response [133]. Hence, PCNA is the repair factor that associates with the unprocessive RNAP II complex to the DNA damage repair machinery during TCR. Supporting these facts, studies have shown that cells lacking BRCA1 have

defective transcription coupled repair (TCR) [26]. Since the results of two independent studies link BRCA1-BARD1 to the stimulation of ubiquitination of RNAPII in non-breast cells, it is improbable that the BRCA1-BARD1 mediated RNAPII degradation by ubiquitination is a breast and ovarian cell-specific event [134]. Linker region between ARD and BRCT domain of BARD1 binds to p53, CstF50 and ATM [25]. However, ARD-BRCT binding to DNA damage repair protein in a phosphorylation dependent manner does not indicate its direct involvement in DNA damage repair mechanism.

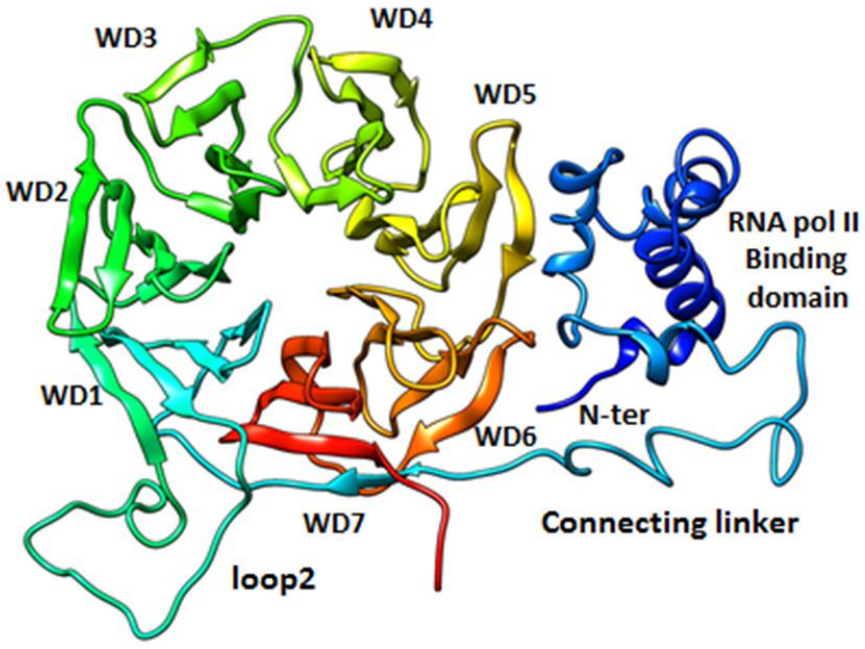


Figure 1.14:- Model structure of CstF50.

The CstF50 N-terminal domain (1-91aa) is the homodimerization domain predominantly comprising of α- helices and is required for interaction with RNA Pol-II CTD. The C-terminal WD40 repeat domain has barrel shaped structure providing an interaction interface at top region, the bottom region, and the cylindrical surface. Therefore, CstF50 WD40 repeat

domains can thus act as a platform for interacting proteins, making them best suited to be hubs in cellular interactions networks.

Structural studies of WD40 repeat domain containing protein complexes indicate that proteins of this family interact with protein partners at the groove formed at the centre of the WD40 β-propeller structure [135]. Such interactions can form an anchor for association of larger protein complexes. Similar, interactions were also characterized for heterotrimeric G-proteins where the β-subunit is a WD40 family protein [136, 137].

1.8 BRCA1-BARD1, E3 ubiquitin ligase mediates CstF50 dependent transcriptional repair control

It is known that regulation of mRNA 3' end formation has significant role in cancer, as unprocessed pre-mRNA are susceptible to nuclease digestion [6], and pre-mRNA cleavage regulates gene expression of diverse mRNAs in tumor cells during cell differentiation [138].

After DNA damage, mRNA 3' processing is inhibited by CstF50-BARD1-BRCA1 complex [12] and subsequently proteasomal degradation of RNA polymerase II CTD [13]. Further, prematurely terminated transcripts are known to activate a nuclear surveillance pathway which allows elimination of defective mRNAs by exosome-mediated degradation [139]. Cancer predisposing BARD1BRCT mutation, Gln564His has been found in the germ-line of a patient with adenocarcinoma [27]. BARD1 Gln564His mutant impairs the binding between BARD1 and CstF50, which leads to the formation of premature transcript and uncontrolled mRNA regulation **(Figure: - 1.15)**.

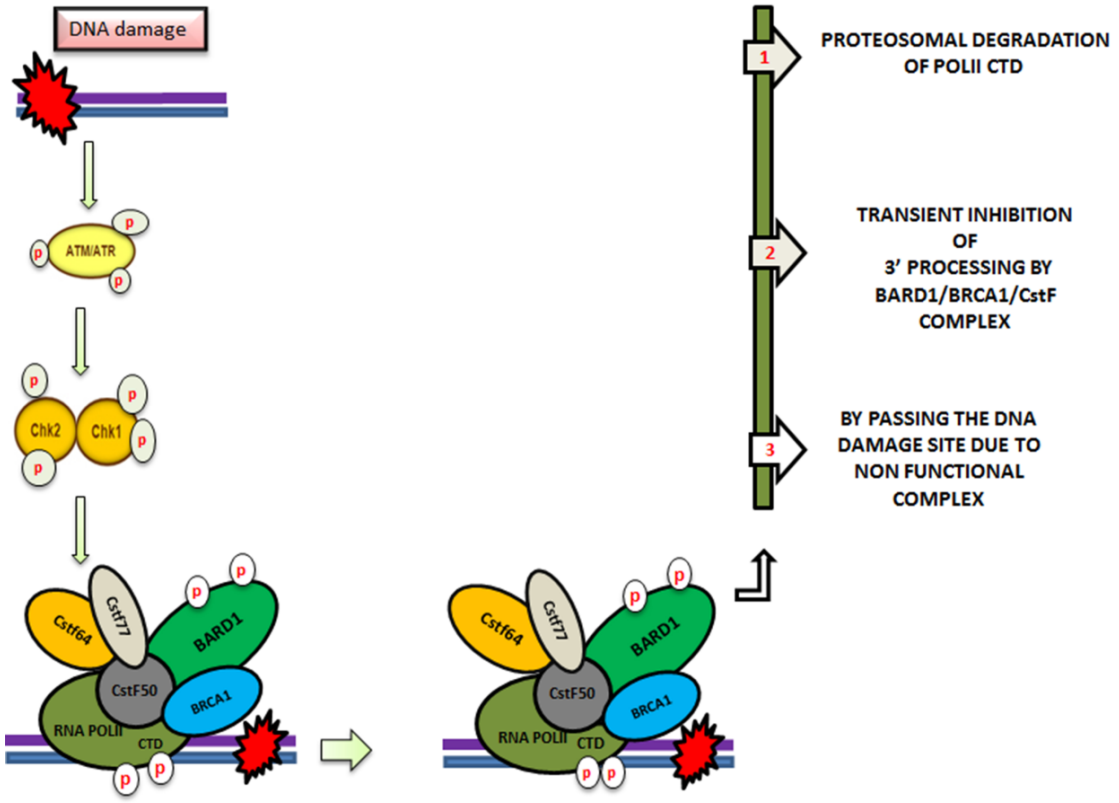


Figure 1.15:- Role of BRCA1-BARD1-CstF50 complex in transcriptional coupled repair.

The BARD1 Gln564His mutation induces apoptosis less efficiently when transfected in mammalian cells, demonstrating that the BARD1 loop region connecting the ARD and BRCT domain is essential for its tumour-suppressor and pro-apoptotic functions **(Figure:- 1.16)** [95]. An in depth understanding of the structure, and functions of CstF50-BARD1-BRCA1 complex, will be helpful in deciphering of cellular processes at atomic level. Furthermore, it will ultimately provide new means to tinker with biological functions via synthetic and systems biology approaches. Reasons for WD40 domains to act as scaffolds evidently epitomise one of the important domain families for indispensable cellular processes.

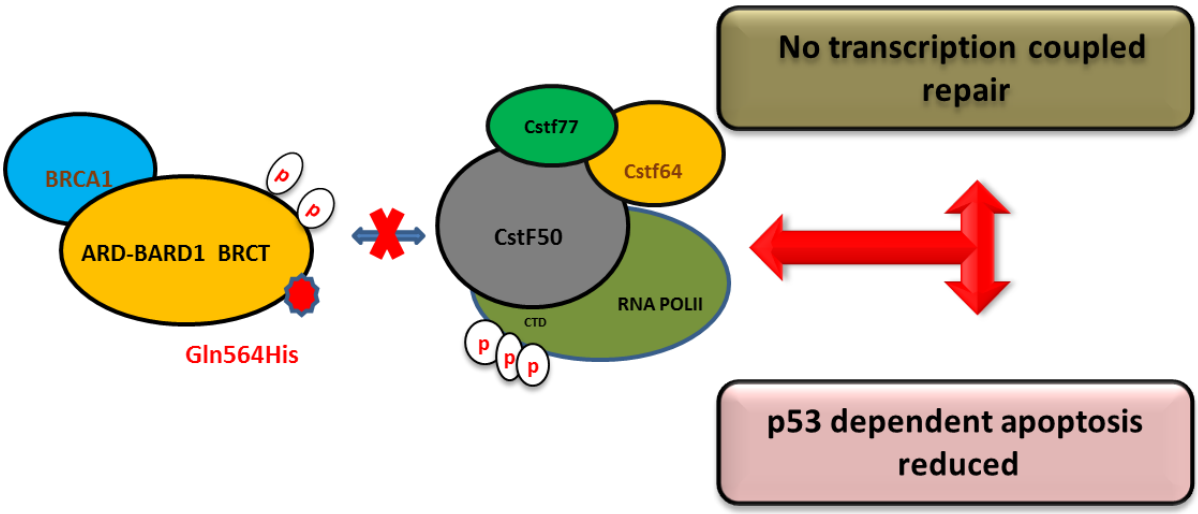


Figure 1.16:- Consequences of BARD1 Gln564His mutation on the p53 and CstF50 interactions.

1.9 BRCT domain as tumor suppressor

The BRCT domain was first characterized and identified in BRCA1. The tumor suppression function of BRCA1 is largely contributed by its BRCT domain as, truncation and missense mutations in BRCA1 BRCT region leads to increased risk of breast and ovarian cancers [140-142]. BRCT domains are present in different proteins, predominantly involved in DNA metabolism and repair mechanism [142, 143].

BRCT domain of BRCA1 recognizes specifically the phospho peptide containing signature sequences pS-X-X-F (pS-phospho Ser, X-any amino acids, F-Phenylalanine) motif present in proteins such as (BACH1/FANCJ), [30, 31] CtIP11 and

CCDC98 [17, 144, 145]. BRCT-BRCT domain packing in different BRCT repeat structures such as Mediator of DNA damage checkpoint protein 1 (MDC1), BRCA1, BARD1 and their functional role in phospho-peptide binding is well recognized [38, 146-148]. First crystal structure of BRCT was inferred for the XRCC1 protein [149].

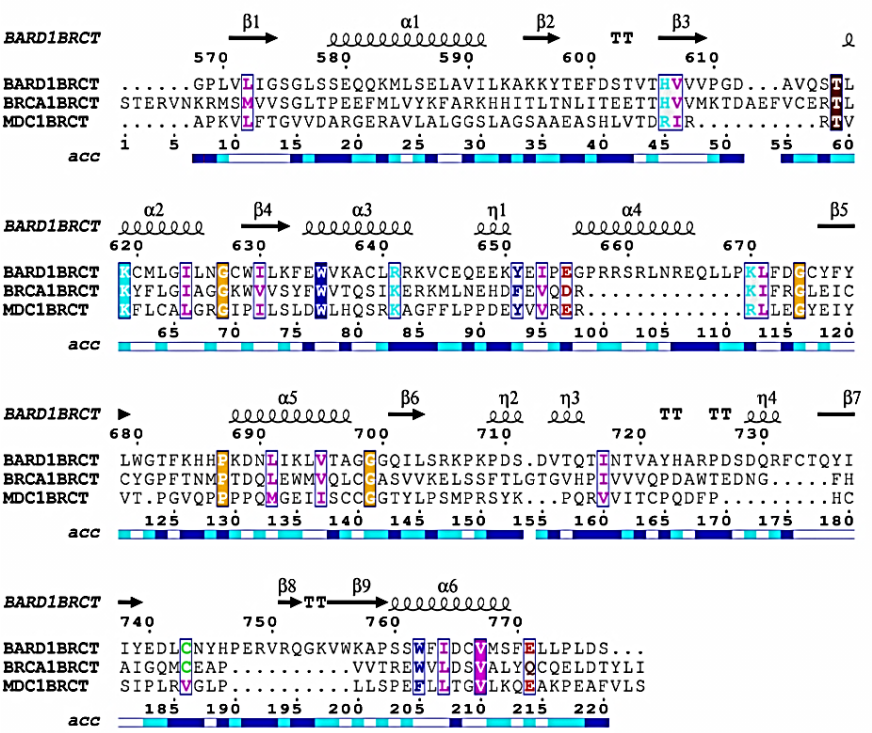


Figure 1.17: - Sequence alignment of BRCA1, BARD1, and MDC1BRCT**.**

The BRCA1BRCT structure shows that BRCT domain is comprised of a central three β-strands separated by a single α-helix on one side and two α-helices on the opposite side (**Figure:- 1.19)** [150]. Structural and sequence evaluation of different BRCT domain indicates that variation in structure and sequences are restricted mainly to the connecting loop regions (**Figure: - 1.17**). The crystal structures of proteins with multiple BRCT domains (tandem repeats) are separated by flexible linker regions. The BRCT tandem repeats in BRCA1

consists of the two sets of three α-helices and three β-strands at the N-terminal and C-terminal domain respectively (**Figure:- 1.19)** [115].

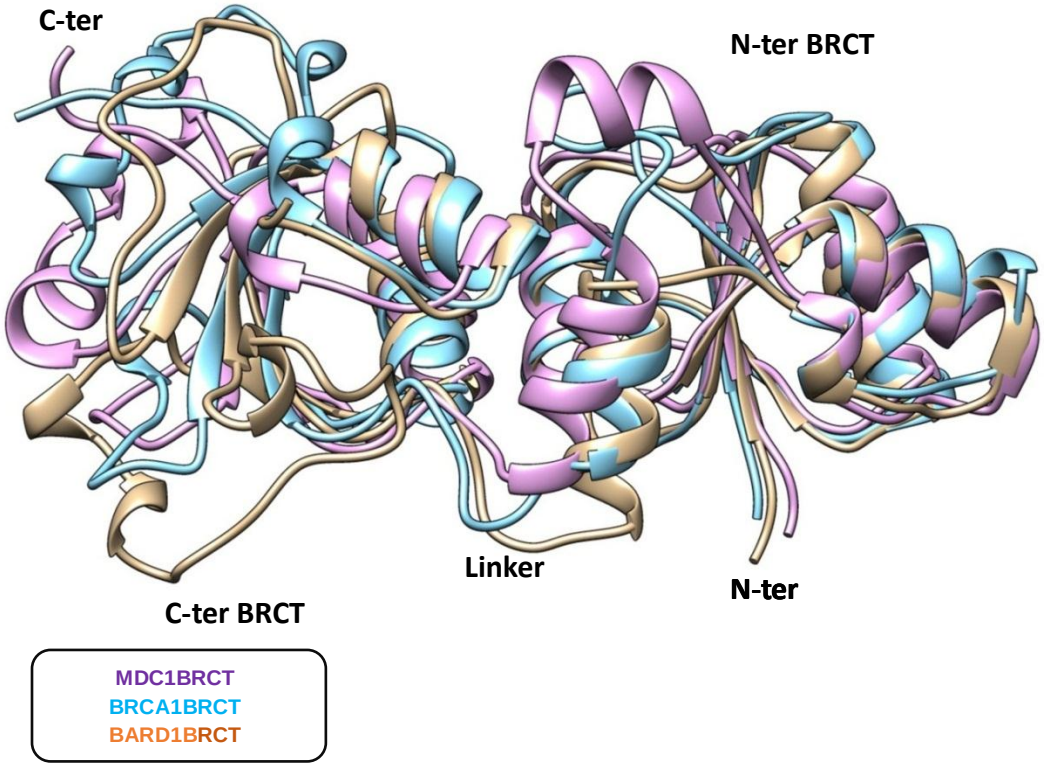


Figure 1.18: - Structure alignment of BRCA1, BARD1, and MDC1BRCT.

So far, two diverse classes of phosphate binding pockets have been recognized in the tandem BRCT repeat domain. Amino acids located at phospho-peptide binding pocket and hydrophobic core are highly conserved [29, 151].

The phospho-peptide binding compartments of BRCA1 and MDC1 serve as models for the first class; as the phosphate moiety of phosphopeptide interacts with Ser1655/Thr898, Lys1702/Lys1936, and Gly1656/Gly1899 in BRCA1/MDC1 respectively. The chemical natures of these residues are also conserved in BARD1BRCT **(Figure 1.17)**.

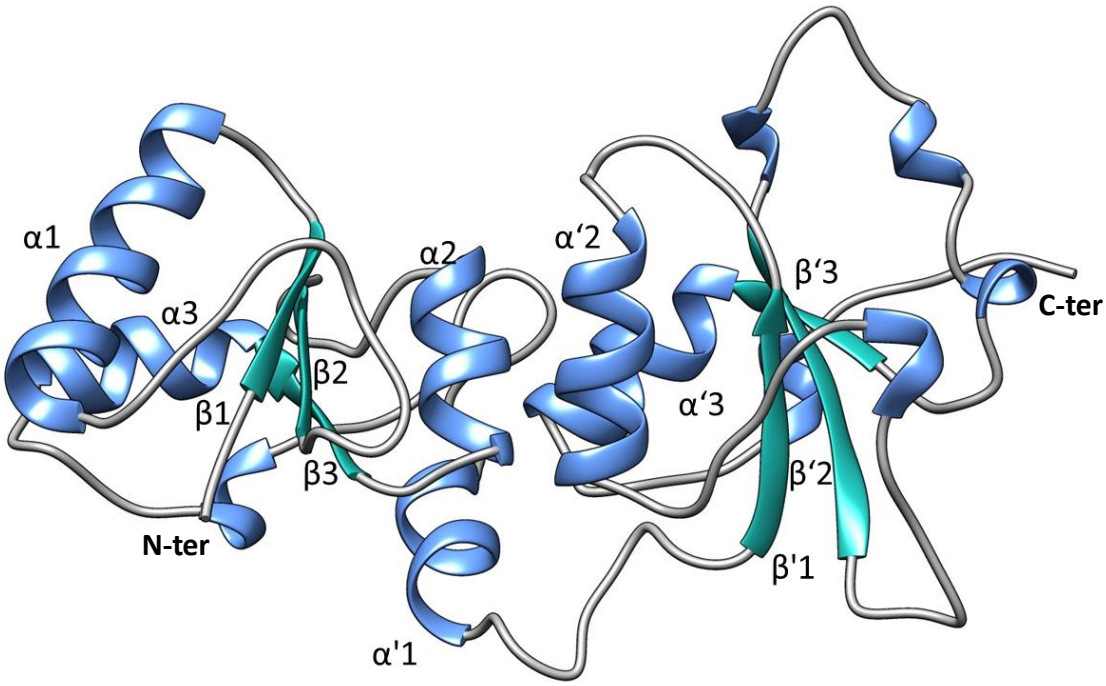


Figure 1.19: - Structure of BRCA1BRCT

In phosphopeptide chain, the +3 residue sits in the hydrophobic interface between the two tandem BRCT repeat and determines the binding specificity [152]. Conserved arginine residues Arg1699/Arg1933 in BRCA1/MDC1, respectively has a pivotal role in distinguishing the peptide main chain at the +3 position.

In the highly homologous structure of BARD1BRCT, the conserved arginine is substituted by a serine residue (Ser616). Such replacement indicates that the phosphopeptide binding properties for BARD1BRCT might differ from MDC1 and BRCA1 BRCT [29, 38, 153, 154]. Hydrophobic interactions between α2-α1'-α3' helices bring the BRCT domain pair in close vicinity and only show small non-conformities in domain movement upon peptide binding **(Figure: - 1.20)**.

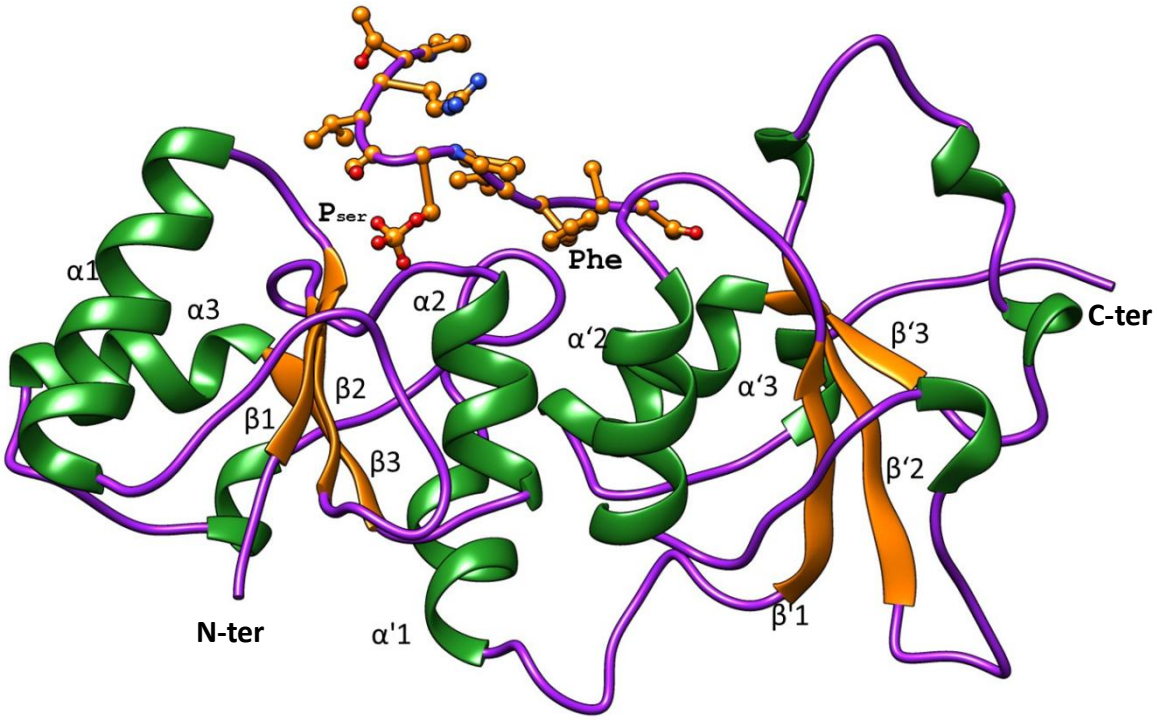


Figure 1.20:- Complex Structure of BRCA1BRCT and CtIP phosphopeptide (PDBID: 1Y98).

Alterations in the hydrophobic packing between BRCT repeats may also alter phosphopeptide specificity. This possibly imitates in the absolute conservation of the cognate phospho-peptide motif sequence for BRCA1, where all confirmed interacting partners contain the consensus motif, pS-x-x-F [17, 30, 31, 91, 144, 145, 155, 156].

1.10 Domain Organization of BARD1BRCT

Like other BRCTs (BRCA1, XRCC, 53BP1) BARD1BRCT also has conserved phosphopeptide interacting residues and have tumor suppressor activity by interacting with other DNA damage repair protein **(Figure:-1.21)** [105]. Phosphorylated BARD1 at BRCT Thr734 in BRCA1-BARD1 complex degrades the actively elongating RNA polymerase, thus preventing cell cycle proliferation after DNA damage, induced by IR [105].

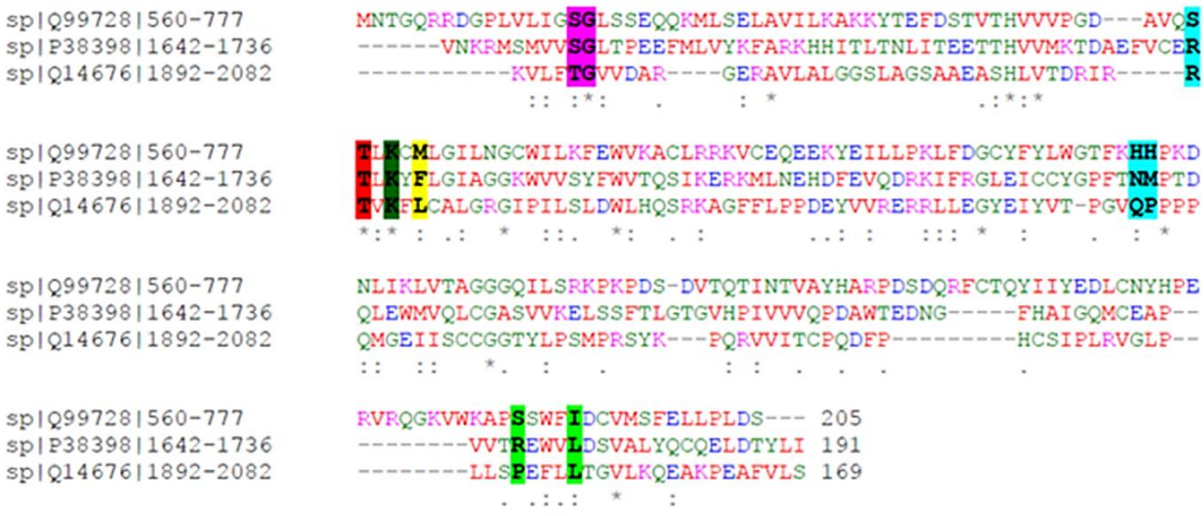


Figure 1.21:- Sequence alignment of BRCA1, BARD1 and MDC1 BRCT showing conserved binding pocket residues.

Site-directed mutagenesis (Thr734Ala) study has revealed that phosphorylation of BARD1 at Thr734 position is required for the BRCA1-BARD1 facilitated ubiquitination and proteasomal degradation of RNA pol II CTD [105]. IR irradiated cells shows activation of ATM/ATR kinases and that further activates CHK1/CHK2 which in turn phosphorylates BARD1BRCT at Thr734 and Thr714 position [105, 157]. Depletion of BRCA1-BARD1 complex using siRNA mediated knockdown leads to significant reduction of polyubiquitination of RNA Polymerase II following IR induced DNA damage [13, 105].

CHK1/CHK2 cell lines having BARD1 Thr734Ala mutation failed to show RNA Pol II CTD degradation, and uncontrolled mRNA regulation. This phenomenon was reconstituted when complemented with *wild-type* BARD1 [61]. However BARD1 Thr 714 Ala $^{(--)}$, Thr734 $^{(++)}$ containing cells did not show significant change in RNA Pol II CTD and mRNA level. Hence, it can be concluded that Thr 734 phosphorylation of BARD1BRCT

is important for RNA Pol II degradation by ATM/ATR, mediated through Chk1/Chk2 under IR induced DNA damage response [105].

Recently, it has been identified that the BARD1BRCT domain interacts with ADP-ribosylated proteins, which are generated after DNA damage. Moreover, it was shown that BARD1 Leu44Pro at RING domain mutant abolishes the BRCA1-BARD1 heterodimer formation and failed early recruitment of BRCA1 to DNA damage foci. Similarly, Cys61Gly mutation in BRCA1 ring domain abrogates the heterodimer formation and lead to similar results amalgamating that BARD1BRCT is required for the early recruitment of BRCA1 at the DNA damage foci. However, ITC experiment doesn't show binding between ADP-ribose and PAR with BRCA1BRCT [18]. This braces the fact that protein with less sequence similarity can have conserved folding pattern and different binding affinity.

Cancer predisposing mutations in BARD1 Cys645Arg, Val695Leu, Ser761Asn, Arg658Cys and Ile738Val are present in the BRCT region [27]. BARD1BRCT Ser761Asn missense mutation is suspected in uterine and breast cancer, Cys645Arg was identified in breast and ovarian cancer but Val695Leu was associated with breast cancer only **(Figure:- 1.24)** [27]. BARD1BRCT Cys645Arg and Val695Leu mutants have not shown interactions with ADP-Ribose, and its polymeric form PAR [18]. Cancer predisposing BARD1BRCT Gln564His mutant was found in the germ-line of a patient with adenocarcinoma [27]. BARD1 mutation Gln564His impairs the binding between BARD1 and CstF50, which leads to the formation of premature transcript and uncontrolled mRNA regulation [6, 13].

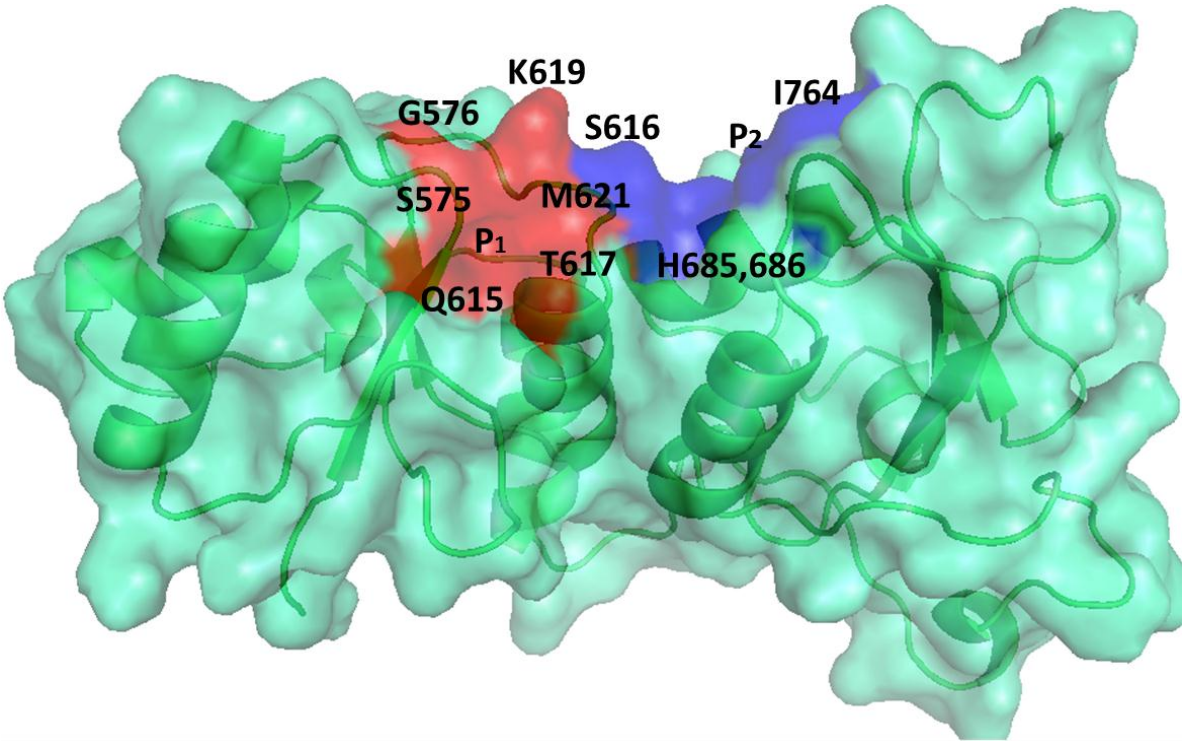


Figure 1.22: - Structure of BARD1BRCT showing conserved phosphopeptide binding pocket.

The Gln564His mutation has shown to induce apoptosis less efficiently when transfected in mammalian cells, indicating that the BARD1 region around Gln564His is necessary for its tumour-suppressor and pro-apoptotic functions [95]. BARD1BRCT Gln564His mutation abrogates the p53 and CstF50 binding to BARD1 [6, 95].

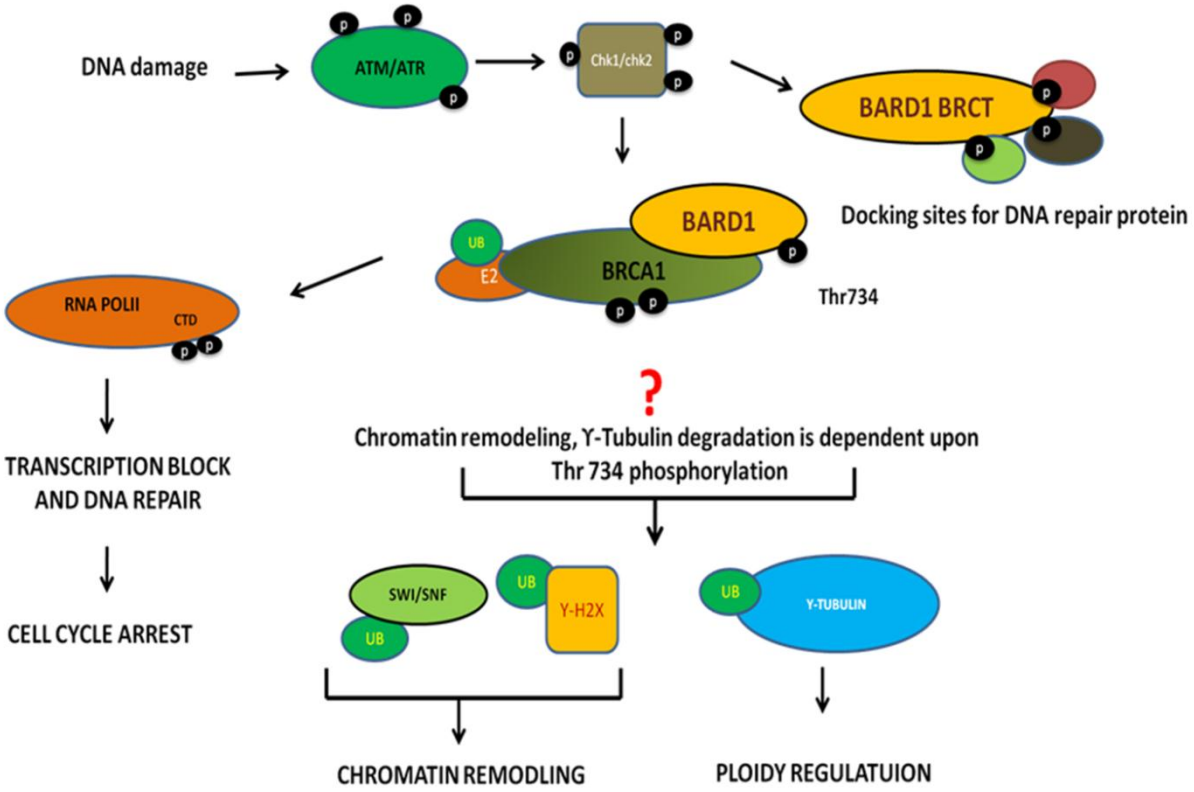


Figure 1.23: - Role of BARD1BRCT Thr734 phosphorylation and consequences upon the DNA repair, chromatin remodeling complex and ploidy regulation**.**

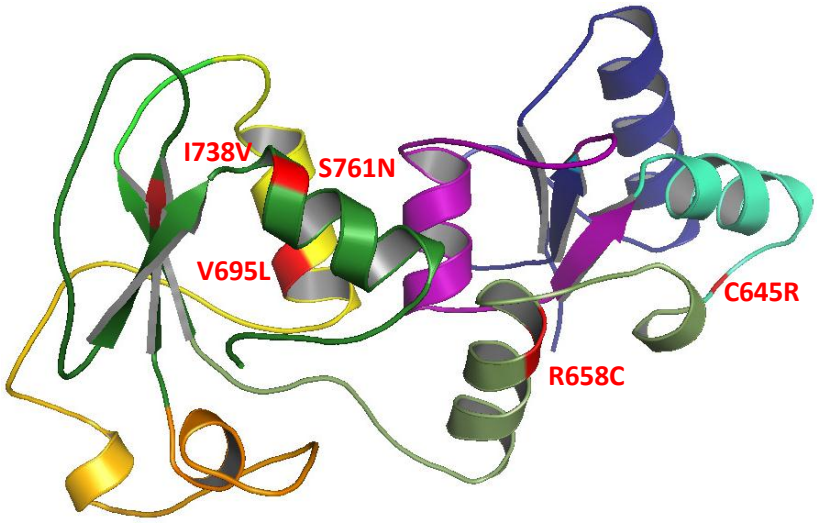


Figure 1.24: - Structure of BARD1BRCT showing cancer pre-disposing mutations.

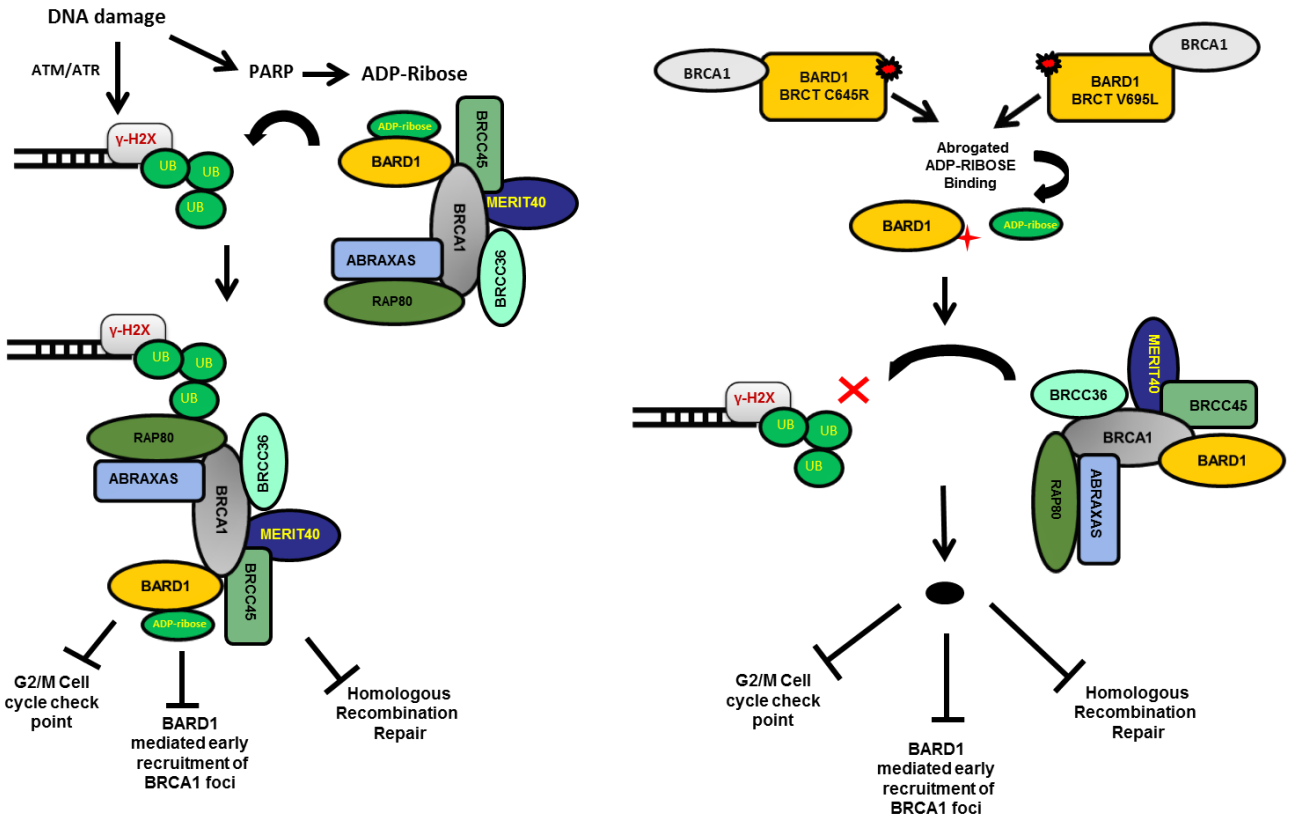


Figure 1.25:- Mechanism of BARD1BRCT in BRCA1 recruitment and consequences of Cys645Arg and Val695Leu mutation upon the DNA damage repair.

1.11 Conclusions

BRCA1 interacts with BARD1 to form a potent E3 ligase [63, 64] and regulate different cellular pathway by ubiquitinating multiple protein substrates [13, 26, 75, 84, 97, 134]. BRCA1-BARD1 E3 ligase activity has been implicated in the BRCA1 mediated control of mitotic spindle assembly [75], checkpoint activation [158], centrosome duplication [159] and tumor cell motility [160]. E3 ligase activity and tumor suppression activity of BRCA1 is impaired due to tumor associated missense mutations Cys61Gly and Cys64Gly at the RING

domain, which abrogates the RING-RING interactions of BRCA1-BARD1 complex [63, 140]. Tumor suppression activity of BRCA1 is mediated by the BRCA1-BARD1 heterodimer and mammary specific inactivation of either *brca1* or *bard1* elicits breast tumors [143]. Unlike the tumorigenic Cys61Gly and Cys64Gly mutations, the synthetic Ile26Ala mutation specifically ablates the enzymatic activity of BRCA1 but allows appropriate assembly of the BRCA1-BARD1 heterodimer. The RING domain comprising zinc finger motif mediates protein-protein interactions while the BRCT domains are evolutionarily conserved and recurrently found in proteins that are involved in cell cycle checkpoint or DNA repair functions in response to DNA damage [142, 143, 161]. However, enzymatically inert BRCA1 is sufficient to suppress tumor formation in several settings, including a model of human basal-like breast cancer. The E3 ligase activity of BRCA1 appears to be dispensable for tumor suppression. Thus, several of the recognized and unidentified functions of BRCA1 may be facilitated independent of its capability to ubiquitination; E3 ligase activity. The BRCA1 protein displays E3 ubiquitin ligase activity, and this enzymatic function is thought to be required for tumor suppression. However, it was found that BRCT mutations that ablate phosphoprotein recognition in BRCA1 elicit tumor formation thus signifying that phosphoprotein recognition by BRCT domain is required for tumor suppression not the E3 ligase activity [162].

BRCTs are evolutionarily conserved tumour suppressor proteins act as docking site for phospho proteins involved in DNA damage repair and cell cycle control [29, 30, 146, 150, 155, 162]. Binding of BRCT domain to phospho-ligands is critical for BRCA1 tumor suppression and 95% of breast and ovarian cancer patients carry the mutations in BRCT region of BRCA1. The crystal structure of BRCT domain from different molecules

from the PDB shows different molecular orientations. Furthermore, truncation and missense mutations in BRCA1BRCT and BARD1BRCT region leads to increased risk of breast and ovarian cancers [27, 140-142]. Cancer predisposing mutations reported in BARD1 Cys645Arg, Val695Leu, which shows failed recruitment of BRCA1 at DNA damage site, suggests that BARD1BRCT is required for BRCA1-dependent tumor suppression. Mutations responsible for breast and ovarian cancer in BRCA1BRCT and BARD1BRCT indicate that BRCT has significant role in tumor suppressor function.

Chapter 2

Materials And Methods

1. Materials:-

Restriction enzymes for gene cloning were procured from Ferments (https://www.thermofisher.com). All other chemicals were purchased from Sigma-Aldrich, unless otherwise specified. Chemicals of analytical grade or molecular biology grade were used to conduct the experiments. Double distilled, 0.44 µM filtered Milli-Q water (Millipore, USA) was used to prepare the buffers for protein purification and other purposes. Purified protein and buffer solutions were either centrifuged or filtered through 0.44-micron filter (Millipore, USA) and degassed prior to use for any experiments.

2. Methods:-

2.1. Gene Cloning:-

2.1.1 Cloning of target gene: Full length cDNA for gene of interest cloned in mammalian or prokaryotic vectors were used as template to clone full-length or desired gene regions of interest. PCR (polymerase chain reaction) was performed to amplify the region of interest. Complementary forward primers were designed in such a way that they have TEV protease nucleotide sequence was incorporated between the gene and fusion TAG region. Following are the steps performed for gene cloning:

2.1.2 Primer Designing: Primers were designed manually; complementarity and thermodynamics parameters were cross-validated using Primer-3 software (www.bioinformatics.org/primerx/). The following parameters were taken into consideration [163].

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

Tm: 50-65°C, GC clamp: Yes (3')

2.1.3 PCR amplification: Different combination of oligonucleotides were pre-tested and evaluated with the Primer-3 suite. The following parameters were used for the PCR 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

To check the desired PCR amplification the amplified PCR product was loaded onto 1% agarose gel and with the help of DNA markers and the region of interest was excised. The gel block-containing product of interest was purified to remove gel contents using commercially available kit (Qiagen). The purified PCR product was subjected to restriction enzyme digestion. The digested gene of interest was purified to remove restriction enzyme, and then ligated in to suitable bacterial expression vector.

2.1.4 Restriction digestion: The PCR amplified gene of interest and expression vector were digested with restriction enzymes (RE) as per the choice of selection of restriction site in the PCR primers. The digestion protocol is as follows.

1. 1-2 µg (20µl) of DNA (gene interest) was taken in the microfuge tube
2. 10X digestion buffer was added as per total volume of reaction mixture.

3. 1µl each (10-unit) restriction enzymes were added to the reaction mixture.
4. Reaction volume was then supplemented with DNAase and protease free distilled water.
5. The reaction cocktail was kept at 37°C in the water bath for 120-180 minutes.
6. Digestion reaction was terminated by heat inactivation of the enzymes at 85°C for 5 minutes.
7. Digested product was purified using enzyme clean-up kit (Qiagen).

The digested product was then ligated using following protocol:-

2.1.5 Ligation: Double digested gene of interest (sticky or blunt end insert) and bacterial expression vector were ligated using T4 DNA ligase.

1. Digested vector backbone and insert DNA were taken in such a way that they are in molar 1:3 ratio (1 vector and 3 insert).
2. Reaction mixture was then supplied with 1 µl of T4 DNA ligase.
3. Ligation buffer (2x) was added to final reaction volume.
4. Ligation cocktail was adjusted to the final volume using nuclease free distilled autoclaved water and incubated at 16°C for 3 hours.

The ligation reaction product was then used for transformation in bacterial host such as *E.coli,* DH5α strain.

2.1.6 Transformation:-

Insertion or uptake of extracellular genetic material in prokaryotes from external agents or by means of extracellular stimulus such as temperature is called transformation. In prokaryotic transformation, fragments, or intact vector with or without foreign gene are taken up by living

cells under *in-vitro* condition. However, cells have to be made competent enough to uptake the naked DNA into the cells.

Competency of bacterial cells is defined as biological nature of bacterial cell(s) to uptake DNA more voluntarily than in the normal growth conditions. Competency in a bacterial system can be externally induced by chemical or physical methods. The treatment of bacterial cells with buffers enriched with divalent cations such as Mg^{2+} or Ca^{2+} most regular practice to make bacterial cells of desired competency. Under suitable condition, the transformation efficiency of most genes would be about 1 in every 10^3 cells [164].

2.1.7 Competent expression host cells:-

The BL21 (DE3) is commonly used bacterial strain, which expediently expresses foreign proteins using T7 promoter. The (DE3) stands for the host is a lysogen of lambda-prophase (DE3), and carries T7 RNA polymerase gene under the control of lacUV5 promoter, which can be induced by non-hydrolysable analogue of lactose IPTG (Isopropyl β-D-1-thiogalactopyranoside). This strain also comes with an advantageous property that they are devoid of expressing Lon and OmpT proteases making them an excellent host for foreign protein expression host.

Expression of eukaryotic proteins in bacterial system often gets compromised due to the different codon usage. To overcome such problems Rosetta 2(DE3) which is a derivative of BL21 is designed to enhance the expression of foreign proteins having rare codons. The Rosetta strain has provided with supply of extra set of 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 the control of their own promoters. In the pLacI and pLysS derivatives of these bacterial strains, the rare tRNA genes are present on the plasmids that carry the T7 lysozyme and *lac* repressor genes, respectively.

Protocol of preparation of competent cells:-

1. Streaked out the desired competent cells from glycerol stock onto a LB agar plate and, grown over night at 37˚C.
2. Prepared an inoculum by inoculating a single colony from LB streaked plate into 2ml LB media and, let it grown at 37˚C for 10 hours.
3. Added the starting inoculum into 500 ml LB media without antibiotic or with antibiotics selective for the host competent cell, and incubated it at 37˚C with constant shaking condition till the O.D reached between 0.5-0.6 at λ=600 nm.
4. Stopped the growth by replacing culture to low temperature (preferably ice) condition for 5-10 minutes.
5. Harvested the bacterial cells in collection tube (sorvall tubes) by centrifuging it at 5000 rpm for 15 minutes at 4˚C.
6. Discarded the broth supernatant and kept the centrifuge tube containing pellet on ice immediately (sterile condition).
7. Added 100 ml of pre-chilled TB1 buffer[*1] and resuspend the bacterial cell pellet such that no clumps should left over (maintain on ice while re–suspending the competent cell pellet)

8. Centrifuged the re-suspended competent cells at 5000 rpm for 8 minutes at 4°C in pre-chilled rotor (sorvall tubes).
9. Discarded the supernatant and kept the bacterial pellet on ice immediately (sterile condition).
10. Added 10 ml of pre-chilled TB2[*2] buffer and resuspend the pellet (maintain on ice while re-suspending).
11. Made the even aliquots of 100 µl in 0.5 ml microfuge tubes.
12. Snap frozen with liquid nitrogen and stored at -80°C for further use.

[*1]**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.

***[2]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 were tested by streaking on LB agar plates with and without ampicillin.

Protocol of bacterial DNA transformation:-

1. Allowed the competent cells to thaw on ice.
2. Added 1µl (50-100ng) of gene of interest cloned to 100 µl of competent cells.
3. Incubated the cells containing foreign DNA on ice for 45 minutes.

4. Provided a brief heat shock for 90 seconds at 42°C and quickly put back the tube in ice and kept it for 2 minutes.

5. Added 800 µl of sterile LB medium in sterilized conditions.

6. Grown cells at 37°C for ~45 minutes on shaker incubator.

7. Centrifuged at 13000 rpm for 1 minute, and re-suspend the pellet in 100 µl of LB media. Spreaded the resuspended cells on the surface of LB agar plate containing the selective antibiotic with sterile L-shaped glass road.

8. Placed the plate in inverted condition at 37°C maintained incubator for overnight to obtain well-isolated transformed colonies.

2.2 Protein Expression and Purifications:-

Recent advancements in the gene manipulation and recombinant DNA technology have enabled us to over-express protein of interest *in-vitro* using a bacterial host. A controlled but strong protein overexpressing promoter can efficiently convert bacterial cells into a micro-protein production factory with a yields as high as 10-20% of total native cellular proteins. Since, we are overexpressing non-bacterial protein, which possess a threat to host system by intoxicating the host cells. Thus, desired gene of interest must be put under very stringent promoter control system. The expression vector system available generally contains modified inducible promoters and can sense the presence of inducer in the system e.g. an inducing agent induces Lac operon induced by addition of IPTG. In the absence of inducer in the growth medium, the lac promoter is repressed by Lac repressor protein (LacI). Induction or turning on of the Lac promoter is mediated by the addition of either Lactose or its non-hydrolysable analogue, IPTG (isopropyl-β-D-thioglactopyranoside) to the medium. Either of

these substances prevents the Lac repressor from binding to the Lac operator, thereby enabling over-expression of the genes.

In BL21 (DE3), the lacUV5 promoter controls the expression of T7 polymerase. BL21 (DE3) carries the LacI gene in its chromosome and exerts a negative control over the expression of T7 polymerase. Consequently, the addition of inducer releases LacI inhibitory effect and expression of T7 polymerase is turned on. Furthermore, addition of IPTG to the culture of BL21 (DE3) strain transformed with expression vector containing gene of interest induces the expression of T7 polymerase which in-turn over-express the protein of interest.

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:** A single colony was selected from freshly transformed agar plate and inoculated it in 100 ml LB broth containing 100 µM of antibiotic for which cells are carrying resistant gene. Incubated the pre-inoculum on shaker incubator at 37°C overnight (12-16 hrs).
2. **Dilution:** Diluted the 1000 ml culture with 10 ml of pre-inoculum (1:100 ratios) of autoclaved LB broth containing 100 µM of antibiotic. Grown the diluted culture on shaker incubator at 37°C until all the cells reached mid-log phase i.e. A_{600} between 0.6-0.8.
3. **Induction:** The cell growth was arrested by cooling the grown culture, added optimized concentration of IPTG, and further incubated the culture on shaker incubator at optimal temperature and time.

4. **Harvesting:** The culture was transferred to centrifuge bottles and centrifuged for 10 minutes at 6000 rpm at 4°C. The pellet was re-suspended 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 until further used.

2.2.1 Protein Purification:-

Protein purification is a sequential way to fish-out your protein of interest from the cellular proteomic pool. Homogenous and highly pure protein is prerequisite for structure determination, functional and interactions analysis (protein with ligand, protein-DNA and RNA). Separation of desired protein from bacterial extract is usually the most difficult part of purification. Purification steps also depend on protein size, binding affinity to its respective tag and folding pattern. Chromatography is a general technique used for separation of desirable protein molecule from the mixture of complex pool of cellular proteins. Selection of stationary phase and mobile phase depends on the biophysical property of the proteins. Proteins are generally purified using size exclusion chromatography, hydrophobicity chromatography, ion exchange chromatography, and affinity chromatography. Regularly, purified proteins are detected as they elute from the column by their absorbance at λ=280 nm. The natural affinity between ligand and acceptor is the basis for affinity chromatography. Well-folded protein molecule recognizes its ligand through one of its structural motifs.

General Protocol for purification of Proteins:-

1. **Re-suspension**: The IPTG induced bacterial pellet was re-suspended in desired volume of purification buffer; supplemented with protease inhibitor (200 µl of 200 mM PMSF and 20 µl of protease inhibitor cocktail).

2. **Cell lysis**: Transferred the suspension devoid of any clump into centrifuge tubes and sonicated the cell suspension at 70-pulse rate and 70 power with 1 minutes of duty cycle in ice bucket. This process disrupts the cell wall using high frequency sound waves. Sonication process involves application of sound energy using an ultrasonic bath or probe to disrupt the cell wall and cellular contents oozes out.

3. **Centrifugation:** The cell lysate was centrifuged at 18000 rpm for 45 minutes at 4°C to remove cell debris. Collected the supernatant and discarded the pellet.

4. **Equilibration of GST/Ni-NTA beads**: Passed five column volume of water to remove traces of ethanol, as the beads are stored in 20% ethanol. Further, provided 5 column volume washes to calibrate the resin.

5. **Binding**: The soluble protein was carefully transferred into the column containing the affinity resin, and allowed the binding by passing the cell lysate by gravity flow, collected the flow through of the column to recheck the fraction unbound to resin.

6. **Washing**: To remove non-specifically bound proteins from the affinity matrix passed ten column volume of wash buffer. Taken ~10 µl of protein bound resin to load on SDS-PAGE gel to detect the binding.

7. **TEV Protease Cleavage**: To remove the affinity tag like GST/his tag, 200 µl (20 units) of 6His-TEV protease with wash buffer was incubated with eluted fusion protein for 3 hours. ~10 µl of resin was taken out to run on the SDS-PAGE gel. The TEV protease is highly site-specific cysteine protease found in Tobacco Etch Virus (TEV). The optimum recognition sequence for the enzyme is Glu-Asn-Leu-Tyr-Phe-Gln (Gly/Ser), and cleavage occurs between the Gln and Gly/Ser residues.

8. **Equilibration of centricon:** Rinsed the centricon (10-50 KDa molecular weight cut off) with water to eliminate traces of alcohol, calibrated it with wash buffer by centrifuging it at 4000 rpm for 15 minutes at 4°C.

9. **Gel filtration:** 2 ml (or as per availability of FPLC loop size) of concentrated protein was injected on AKTA- FPLC pre-equilibrated with FPLC buffer and selected the size exclusion column (superdex-75 120ml) as per size of protein. Eluted protein fractions were collected by reading the absorbance at wavelength λ=280 nm.

10. **Loading on SDS-PAGE**: 10 µl of each FPLC fractions were loaded on desired percentage of SDS-PAGE gel, and stained with coomassie dye. Furthermore, destained to visualize the protein bands under trans-illuminator.

11. **Concentrate the protein**: The purified fraction of proteins, which appears to be pure at the expected size, was concentrated as per the requirement.

2.2.2 Regeneration of GST resin:-

Composition of Wash buffer: 300 mM NaCl, 50 mM Tris, 5% glycerol, pH-7.5

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

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

Protocol:-

1. 20 mM reduced glutathione solution prepared in 50 mM Tris pH 8.0 (final pH) was passed twice, through the column containing GST resin, to remove traces of bound GST and GST-fusion protein.

2. Five column volume distilled water was passed.

3. Provided 1 column washes with washing buffer followed by 2 column washes with 1% Triton-X 100.
4. Passed one column volume of 70% ethanol followed by extensive washes with buffer-A and buffer-B alternatively.
5. Finally, passed single column volume wash buffer and a wash with distilled water subsequently.
6. Stored the resin in 20% ethanol at $4^{o}C$ for further use.

2.2.3 Regeneration of Ni-NTA resin:-

Regeneration buffer: 6 M GuHcl, 0.2 M acetic acid

Protocol:-

1. Two column volumes of 750 mM imidazole solution prepared in lysis buffer (pH 7.5) was passed, followed by two column washes with regeneration buffer.
2. Provided excess of column washes with distilled water to completely remove traces of GuHcl and three subsequent washes with 2% SDS.
3. Passed single column volume of 25%, 50%, 75%, and 100% ethanol. Repeated the washing steps in reverse order.
4. One column volume of distilled water was allowed to pass through the column, followed by another five column washes with 100 mM EDTA (pH-8.0) to strip of all the chelated Ni^{2+}.
5. Again, re-charged the NI-NTA resins by applying two column volumes of 100 mM $NiSO_4$, and passed 3 column volume of water to remove unbound nickel.

6. Stored the resin in 20% ethanol at 4^{o}C.

2.3 Protein characterization (BARD1 ARD-BRCT *wild-type* and Gln564His mutant, BARD1BRCT *wild-type*, mutants and CstF50)

2.3.1 Protein Estimation:-

Bradford (Expedon) assay and Nanodrop were used to measure the concentration of the purified proteins.

1. Different concentration of BSA 0.1, 0.5, 1.0, 2.5, 5, 10 mg/ml were prepared as a standard reference and absorbance was recorded at λ=595 nm using split beam UV spectrophotometer (Shimadzu).

2. For most of the biophysical study, it is very important to estimate accurate protein concentration. Using protein sequence, the concentration of the protein can be obtained by determining the absorbance at λ=280nm and extinction coefficient using Nanodrop spectrophotometer.

2.3.2 Mass Spectrometry:-

The mass spectrometry is very powerful tool for identification of protein or DNA sequence, mass determination, post-translational modification, and other biophysical behaviour of proteins. This technique needs only miniscule amount (~ 1-10 ng) of protein. When the charged molecules or analytes are introduced in an electric and/or magnetic field, their paths through the field are function of their mass-to-charge (*m/z)* ratio. Mass/Charge ratio of the analyte can be used to determine the mass (M) of the analyte with very high accuracy.

Mass spectrometry can be applied to sequence short stretches of a polypeptide using a technique called tandem *MS* or *MS/MS.* A solution containing the protein to be identified is first digested with protease or chemical reagent to hydrolyse it to a mixture of oligopeptides. The mixture is then subjected into device that is mass spectrometers in tandem.

Types of Mass spectrometry: Mass spectrometry of biomolecules can be characterized based on sample introduction/ionization approaches:-

1. Electron Impact (EI): This is a harsh but versatile method, which provides fundamental information for the molecule up to 1000 Daltons.
2. Chemical Ionization (CI): It is a lenient method for generation of molecular ion peak $[M+H]^+$ and can be engaged for molecules up to 1000 Daltons.
3. Electrospray Ionization (ESI): It is a soft method for production of manifold charged species, and can be employed for molecules up to 200,000 Daltons.
4. Matrix Assisted Laser Desorption (MALDI): It is a soft method for generation of multiple charged species and can be used for molecules up to 500,000 Daltons.

Protocol for mass spectrometry sample preparation:-

1. Expurgated the gel into small pieces and, incubate into fresh destainer by keeping on shaking conditions overnight for complete removal of staining from the gel pieces.
2. Provided repeated washing (3 changes for 10 minutes each) to the gel pieces with distilled water to eliminate any staining content from the gel.
3. The excised gel pieces was treated with 50 mM NH_4HCO_3 + 50% acetonitrile (1:1) cocktail to offer buffering and destaining, kept it for 15 minutes at 25˚C.

4. Excluded excess of liquid and submerged it with sufficient amount of 100% acetonitrile.
5. Removed the excess of acetonitrile, and add 50 mM NH_4HCO_3 to rehydrate the gel plugs.
6. Added equal volume of 100% acetonitrile subsequently, kept it for 5 minutes to allow the gel to shrink.
7. Quick dried the gel particles in vacuum spin-dry (for 10-25 minutes) to vaporize the solvent.

Reduction and Alkylation-

1. Swelled the gel pieces in 10 mM DTT, 50 mM NH_4HCO_3 (300 µl/tube) (freshly prepared) for reduction reaction.
2. Incubated the gel particles at 56°C for 45-55 minutes. Brought the tubes at room temperature when incubation period is over.
3. Eliminated additional fluid and substituted it rapidly by approximately the identical volume of recently prepared 55 mM iodoacetamide (IAA) in 50 mM NH_4HCO_3 for alkylation.
4. Kept it for 30 minutes at room temperature in dark.
5. Removed the IAA solution and washed the gel particles with the mixture of 50mM NH_4HCO_3 and 50% acetonitrile (1:1).
6. Added plenty of 100% acetonitrile (about 200 µl) to cover the gel pieces.
7. Once the gel pieces shrinks, acetonitrile was removed.
8. Dried down the gel particles in speed vacuum (5-6 minutes).

In Gel Digestion:-

1. Provided superfluous amount of freshly prepared enzyme solution (Trypsin + NH_4HCO_3) to the gel particles. The working concentration of trypsin can be in the ratio of 1:100 or 1:10.
2. Discarded the additional solution and incubated the tubes containing gel pieces at 37°C overnight in upright position (16 hours minimum).

Extraction of peptides:-

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

1. Added ~ 30 µl of extraction buffer each time after vortexing (three times, and kept the extracted peptides in fridge for supplementary use).
2. Centrifuged the extract from the sample for 1-2 minutes at 1000 rpm.
3. Dried the sample (45-50 minutes) using vacuum drier.
4. Stored the vacuum dried samples at -20°C until next step.

2.3.3 Circular Dichroism:-

Circular Dichroism (CD) is an optical phenomenon which arises due to the differential absorption of left and right circularly polarized light by chiral molecules. Only asymmetrical molecule can display such kind of activity when exposed to circularly polarized light. Biological molecules like amino acids, proteins can display differential absorption of R-CPL (right circular polarized light) and L-CPL (left circular polarized light). Proteins having different composition of secondary structural element like α- helices, β-strand or random coil has unique signature in CD spectrum. The far-UV region, (λ=250-190nm) CD spectroscopy

of proteins reveals secondary structural contents of the protein. However, near-UV region (λ=350-250nm) CD spectroscopy predicts tertiary structure of the protein. Proteins showing absorption minima at λ=208 nm and 222 nm indicate predominant α-helical structure, whereas a minimum at λ=218 nm is a characteristic content of β-strands or sheet. CD signal is often affected due to the composition of buffer in which protein is present. Hence, it is important that spectra of buffer components may also require additional monitoring at lower wavelengths. For example, CD spectrum at λ= 208, 218, 222 and 228 nm allows acquisition of higher signal to noise ratio for the secondary structural changes at different urea concentration. While avoiding amplification of the voltage at the higher concentrations of denaturant may incorporate noise in the data acquisition.

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

Protocol:-

1. CD cuvette was washed with 0.1% t-Octylphenoxypolyethoxyethanol (Triton) and subsequently washed with 100% ethanol and air-dried the cuvette.
2. Scanned 5-10 µM of protein solution (in 10 mM HEPES, pH 7.5, 50 mM NaCl) in a wavelength range of λ 200-240 nm at 15°C.
3. Calculated the mean residual ellipticity considering the protein concentration and mass.

2.3.4 Fluorescence Spectroscopy:-

At room temperature, each molecule poses minimum energy state, and attains higher energy level when excited from the external source. Each electronic level comprises of closely

spaced energies bands, which exist at different vibrational modes. Electronic states are characteristically divided by energies roughly 10000 cm^{-1}. Each electronic state is fragmented into multiple sublevels demonstrating the vibrational modes of the molecule. The energies of the vibrational levels are also separated by about 100 cm. Fluorescence intensity measurement allows the determination of the local environment presence of intrinsic fluorophores and their absorptions. The fluorescence emission spectrum of a well-folded protein is a combination of the fluorescence emission of individual aromatic residues. Aromatic amino acids like phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp) have intrinsic fluorescence properties due to presence of conjugated ring system. Emission from Tyrosine and Tryptophan are generally used because their quantum yields are high enough to provide a decent fluorescence signal. The major involvement in intrinsic fluorescence emissions of a folded protein is due to tryptophan residue, with little complement from tyrosine and phenylalanine. Disulphide bonds also have considerable absorption in the same range. Emission spectra is collected from λ=310 to 400 nm, following excitation at λ=280 or 295 nm. Temperature and chemical denaturants are used to monitor the folding pattern using fluorescence spectroscopy. In both, we can collect emission spectrum for the native and unfolded protein using a single wavelength range scan to monitor the structural changes.

Spectrofluorimeter: JOBIN YVON Horiba Fluorolog 3

Protocol:-

1. The tryptophan and tyrosine microenvironments were monitored using Fluorolog 3 (Horiba, USA) at excitation wavelength of λ=280 nm and temperature of 10°C.
2. Fluorescence emission scans were recorded from λ=310-380 nm wavelength.

There are two widely used methods to monitor the protein folding and melting temperature.

1) Thermal Denaturation:-

A protein can acquire different conformations, and each conformation is closely related to the previous one. Week intermolecular interactions like hydrogen bonding, ionic interaction, van der Waal interactions, and hydrophobic interactions play very important role in protein folding. The kinetics and thermodynamic folding parameters of protein can be elaborated by utilizing intrinsic probes such as fluorescence and CD. Some proteins have internal features, which leads to its aggregation or precipitation, which in turn poses experimental difficulty to apprehend the folding pattern. For thermal denaturation using CD, spectra of 10 µM protein at a temperature range of 10-80°C and λ=200-250 nm wavelengths were collected respectively. Fraction unfolded were calculated at each temperature interval and plotted upon each other to determine the melting temperature and folding pattern. In fluorescence spectroscopy, the tryptophan and tyrosine loci can be monitored at excitation wavelength of λ=280 nm and tryptophan specifically at λ=295 at 10°C. Emission spectra were collected from λ=310-400 nm wavelength. 2-µM *wild-type* and mutant proteins were unfolded in a temperature range of 15-80°C and emission maxima were obtained.

Protocol for BARD1 ARD1-BRCT *wild-type* and Gln564His mutant thermal denaturation

For thermal denaturation in CD:-

1. Purified *wild-type* and mutant protein (10µM) were unfolded in a temperature range of 10-60°C and unfolding pattern was analyzed at λ= 222 nm wavelength.

2. Data fitting was done in two-state transition model and thermal parameters were calculated respectively.

Protocol for BARD1BRCT *wild-type* and mutant proteins thermal denaturation:-

1. 10 µM proteins were unfolded in a temperature range of 15-80°C and spectra were obtained.
2. Data was analyzed by fitting in to a three state transition model and thermodynamic parameters were calculated.

Chemical denaturation: Fluorescence spectroscopy is spectrophotometric technique used to assess minor conformational changes in the protein structure even at low protein concentration (2µM). To examine the alterations in proteins tertiary structures, we have taken GuHcl as chaotropic agent. Equilibrium denaturation studies allowed us to compute the free energy, protein stability and presence of intermediates.

The protein concentration required for measuring the equilibrium of unfolding experiment in the micro molar range, but it also depends on number of aromatic residues. Different suitable buffer systems are available for fluorescence and CD experiments. However, absorbance required is to be corrected to obtain native protein fluorescence contribution.

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

1. Prepared 7 M GuHcl stock.
2. Checked that the protein is completely denatured.
3. Adjusted the instrument background signal using native and denatured protein in their respective buffers.
4. Incubated the samples in increasing (0-6M) concentrations of denaturant.

5. Established the equilibration time and reversibility.

6. Performed equilibrium-unfolding experiments.

7. Repeated experiments at different protein concentrations to determine the model used for fitting the data.

Emission scans were collected between λ=310-390nm and blank correction was done to get high signal to noise ratio. Data fitting was performed in two or three state based on intermediate presence and further thermodynamic parameters were calculated.

Data analysis and curve fitting:-

Data fitting (CD, Fluorescence) was performed using Origin (Origin Lab, 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 [165]

$$N \overset{K_1}{\longleftrightarrow} I \overset{K_2}{\longleftrightarrow} U \quad (1)$$

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

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

$$\mathbf{F_F = (Y_x - Y_u) \div (Y_N - Y_u)} \quad (2)$$

Y_x and Y_U are the normalized blank corrected signals and the signal of the unfolded protein respectively, and Y_N is the signal of the folded protein. F_F represents the fraction folded at a particular temperature or denaturant concentration. A plot of the normalized signal (fraction

folded, f_F) against the denaturant variables, generate the unfolding curve for spectroscopic probes.

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

$$\mathbf{K=f_u/f_F=f_u/ (1-f_u)} \quad (3)$$

Following equations were employed to calculate the free energy (ΔG) and enthalpy (ΔH) of unfolding reaction from K.

$$\mathbf{\Delta G=-RTlnK} \quad (4)$$

$$\mathbf{d(lnK)/d(1/T)=-\Delta H/R} \quad (5)$$

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

$$\mathbf{\Delta G=\Delta G^{H2O}-m\ [urea]} \quad (6)$$

ΔG^{H2O} 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 dependence of ΔG on urea concentration [166, 167].

2.3.5 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. The target proteins would be induced to acquire the partially folded state under specific conditions, such as low temperature and pH. It can also be useful to isolate the protein fragments that can fold autonomously, and thus behave as domains.

Protocol for limited proteolysis of BARD1 ARD-BRCT *wild-type* and Gln564His mutant:-

1. Equal concentration of BARD1 ARD-BRCT *wild-type* and Gln 564 His (1.5 mg/ml) was incubated with trypsin independently in two different set of experiment, keeping the final concentration of trypsin 40ρg/µl.
2. Reaction mixture was incubated for different period of time 0, 5, 10, 15, 30, 60, 120 minutes at 37°C (trypsin).
3. Reaction was terminated individually by adding protease inhibitor 1mM PMSF.
4. Samples were heated by adding equal volume of laemmli buffer and analyzed by SDS-PAGE gel.

Protocol for limited proteolysis of BARD1BRCT *wild-type* and mutants:-

1. Trypsin and chymotrypsin were mixed to purified *wild-type* and mutant proteins (2mg/ml) separately with their final concentration of 40 Þg/µl and 10 Þg/µl respectively. 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, 120, 180 minutes. Reaction was terminated by adding 2 µl of 200 mM PMSF.
3. Samples were heated with laemmli buffer and analyzed on SDS-PAGE gel. To our conclusion, band corresponding to 24 KDa was considered as a most stable fragment.
4. Domain of interest was identified by mass spectrometry and Mascot analysis with Biotool software (Bruker Daltonics, USA) [168].

2.3.6 Molecular Modeling and Docking:-

Protein structure determination and functional characterization is an essential area of investigation in structural biology. Determination of protein tertiary structure unravels the functional domains present in the molecule at the atomic level. X-ray crystallography, electron microscopy and NMR spectroscopy are the important experimental methods for protein structure determination [169-171]. Structure determination of protein using X-ray crystallography or NMR spectroscopy requires large amount of pure protein, which sometimes becomes the limitation of these techniques. However, recent advancement in computational resources has enabled us to model the protein structure. Currently, *ab initio* prediction and homology modeling are widely used to generate the protein 3-D structure. *Ab initio* or *de novo* prediction is established on the fundamental physical and chemical properties and principles. *Ab-initio* or *de novo* method is limited by computational power and accuracy. However, homology modeling of protein depends on sequence similarity with structures available in the Protein Data Bank (PDB) and other structural databases. Homology modeling is based on the principle that proteins of similar amino acid sequences, commonly evolutionarily interrelated, fold into analogous 3D structures [170]. There are two widely used approach for protein modeling:-

1. Knowledge-based model that combines the sequence data to other information, such as homology modeling [172, 173].
2. Energy-based calculations through theoretical models and energy minimization, such as *ab initio* prediction [174].

1. Homology Modeling

When sequence homology with known protein is high and stereo-chemical structure is known, modeling of an unknown structure by comparison can be carried out with reasonable success. It is shown that protein pairs with a sequence homology more than 50%, have 90% of the residues within a structurally conserved common folds [175]. If one examines sets of related proteins, they may find different primary amino acid sequence but tertiary fold is similar which signifies that general folding pattern of the structural core is conserved. Given a protein sequence of interest, the homology modeling procedure requires investigation of homologous sequences with solved crystal structures, alignment of the query sequence to the template structure, 3D model building, and refinement of the proposed model.

The homology modeling normally consists of four steps:

1. Begin with the known sequences.

2. Assemble fragments/substructures from known homologous structures.

3. Carry out investigation about the limited structural changes from a known adjoining protein.

4. Optimize the structure by energy minimization.

Homology modeling relies on the accepted similarity in 3D structure between proteins of similar sequences. Hence precise identification of homologous sequences is critical to generate the correct structural representations. Most commonly, identification of close homolog occurs through comparison of protein sequence profiles or Hidden Markov Models (HMMs). Additional, use of more innovative methods of sequence assessment, such as

sequences to profiles alignment (PSI-BLAST), profiles to profiles (FFAS), or HMMs to HMMs, helps greatly in homolog identification.

First and foremost BLAST searches are performed against the input sequence for homology identification by pairwise sequence alignment. The reliability of BLAST and PSI-blast output in the "twilight zone" (<30% identity with probe) is often compromised, which limits the effectiveness of homologues identified by BLAST. In the profile based identification methods, that allows identification of more distantly related proteins. In a standard study of the SCOP database, PSI-BLAST validates with twice the accuracy of BLAST, while FFAS search enhances the accuracy by additional 20% upon PSI-BLAST. These increments in search specificity ultimately led to greater alignment of the query sequence with the experimental structures [169, 170]. Once a set of homologues is cautiously chosen to use as templates, alignment of the probe sequences to the template structures can often be followed by adjustments to optimize placement of insertion and deletions outside tight secondary structure elements. Building the large loops with low structural homology or no homology can be performed by searching structural libraries, loops longer than six or seven residues increases challenges because of the large number of random possible conformations [176]. Contemporary homology modeling techniques have achieved structure predictions of such accuracy that it can be successfully used in drug designing, virtual screening, and site directed mutagenesis applications [170].

2. *Ab initio* Prediction of Protein Structure

The prediction of the 3D structure from the sequence of amino acids in the absence of template structure has been a great challenge for protein modeling [174]. A 3D stereochemical structure determination of protein in the absence of homology depends on the

energy based structure prediction, which relies on energy minimization and molecular dynamics. The energy based structure prediction method is faced with the problem of a large number of possible multiple structures having very similar energy minima. A protein of 100 amino acids can have $(2x5)^{100}$ different main chain conformations, if only two torsion angles per residue are assumed to have five likely values per torsion angle. If we assume and calculate other features like optimal (constant) bond lengths, bond angles, and torsion angles for the side chains, it is computationally impossible to evaluate and analyze all the possible conformations to find the global minimum. Various constraints and approximations have to be introduced to simplify the problem of protein structure prediction. The total molecular energy E can be expressed as under the assumption of constant bond lengths, and bond angles as:

$$\mathbf{E = E_{tor} + E_{elec} + E_{vdw} + E_{pse}} \quad \textbf{(1)}$$

Where E_{tor}, $E_{elec,}$ E_{vdw}, and E_{pse} are torsion angle potential, electrostatic potential, van der Waals and pseudo entropic interaction energies respectively. The secondary structure prediction may be helpful in reducing the search space so, calculation of realistic torsion angles or restrictions for torsion angles permitting to the predicted secondary structures can be used to constrain main-chain torsion angles. It is reasonable that an incorrectly predicted structure will have fewer stabilizing hydrogen bonds, electrostatic, and van der Waals interactions, larger solvent accessible surface and a greater fraction of hydrophobic side-chain atoms exposed to the solvent [177]. Typically, force field provided to *ab initio* modeling conducts a conformational quest under the direction of a pre-designed energy function. This method commonly produces a number of possible conformations (structure decoys), and final

model is selected from them. Therefore, successful *ab-initio* modeling depends on three factors:

(1) A precise energy function with which the native structure of a protein resembles to the utmost thermodynamically stable state, associated to all likely decoy structures.

(2) An effective search method, which can rapidly and precisely identify the low-energy states over conformational search,

(3) Assortment of native-like models from a pool of decoy structures.

One of the best known methods for *ab-initio* modeling is one pioneered by Bowie and Eisenberg, who generated protein models by collecting small fragments (mainly 9-mers) taken from the PDB library [178]. Grounded on a related idea, Baker and coworkers have established ROSETTA [179, 180], which is extremely successful for the free energy modeling (FM). In the recent expansions of ROSETTA, models were generated in a reduced form with conformations specified with only heavy backbone and C_β atoms. Furthermore, the low-resolution decoys were subjected to all-atom refinement technique by an all-atom physics defined energy function, which includes van der Waals interactions, pair wise solvation free energy, and an orientation-dependent hydrogen-bonding potential [181, 182]. For the conformational search, multiple rounds of Monte Carlo minimization are performed [183]. Despite the significant success, the computational cost of the procedure is relatively exclusive for routine use. Partially because of the remarkable achievement of the ROSETTA algorithm, as well as the restricted accessibility of its energy functions to others, several groups started developing their own energy functions using the idea of ROSETTA. Derivatives of ROSETTA comprise Simfold [184] and Profesy [185].

2.3.7 Molecular Docking:-

Protein-protein interactions are the basis of cellular function. The complex molecule adopts energetically and entropically best binding interface for each other. Furthermore, determining the structures of complex protein by experimental methods (X-ray and NMR solution) is still challenging. Hence, *in-silico* protein-protein docking can help to characterize the binding interface. It is usually achieved in three steps:

(**i**) **Sampling**: A large number of randomized conformations of the protein structures to be docked are created.

(**ii**) **Optimizing**: The binding complexes are optimized with respect to different energy terms, geometrical reasons or other supplemental information by performing translational, rotational, and conformational changes.

(**iii**) **Scoring:** The output structures are ranked to find the best solutions (close to the native complex structure) in top positions. This is achieved by calculating appropriate scores that measure the quality of the complexes.

In molecular docking, one molecule (ligand) is brought into close proximity of another (receptor) while investigating the energies of interaction in different possible mutual orientations. In the complex, ligand and receptor molecules are bonded in such a way that they adopt the most favorable (energetically) docked structures. There are majorly two approaches for docking. The first class treats a whole ligand molecule as a beginning point and engages a search algorithm to investigate the energy contour of the ligand at the binding site. Furthermore, it examines optimal reasonable solutions for a specific scoring function. The search algorithms comprise geometrical complementary fit, simulated annealing, genetic

algorithms and molecular simulation dynamics. Representative examples are DOCK3.5, AutoDock, and GOLD [186-188]. The second methodology of docking starts by placing one or several substructures (geometrical conformers) of a ligand into a binding pocket, and then it builds the rest of the molecule in the site. Representative examples are DOCK4.0, FlexX, LUDI, GROWMOL, and HOOK [189-194]. In an interactive docking, an initial knowledge of the binding site is normally required which helps in placing the ligand interactively onto the binding site. Furthermore, geometric restraints such as distances, angles and dihedrals between bonded or non-bonded atoms may simplify the docking process. For interatomic distance, bond length can be restrained to remain constant during a simulation for representative complex structure. Such constraints also guarantee the confinement of the ligand molecule at juxtaposition of the binding site of the receptor during energy minimization process. In an automatic docking, ligand is permitted to fit into possible binding cleft of the receptor of known molecule [186-188]. HADDOCK (High Ambiguity Driven biomolecular DOCKing) [195-197] is developed by EMBL which takes advantage of biochemical and biophysical data documented during chemical shift perturbation data resulting from NMR titration experiments, mutagenesis data or bioinformatics predictions. HADDOCK uses Ambiguous Interaction Restraints (AIRs) to carry out the docking process. An AIR can be defined as an ambiguous distance between all residues shown to participate in the interactions. Additionally, HADDOCK also uses more quantitative restraints between pairs of nuclei from NOEs and RDCs (residual dipolar couplings). HADDOCK has been extended to deal with large variety of data and complexes in addition to protein-protein docking [196, 197]. Docking by HADDOCK is performed by transforming the data into ambiguous interaction restraints (AIRs) that describe a huge grid of ambiguous spaces

between residues expected to be involved in the binding mode without forcing it to any specific position or orientation on the components. HADDOCK also handles other NMR sources of knowledge such as residual dipolar couplings (RDCs) [195] diffusion relaxation [198], and pseudo-contact shifts [199]. Other low-resolution data such as small angle X-ray scattering (SAXS) [200] and cross-link data from mass-spectrometry can be used for scoring and/or generating models.

The HADDOCK docking protocol

Docking in HADDOCK is accomplished in three successive stages:

1. **Randomization of orientations and rigid body energy minimization.**

The receptor and ligand are placed at a geometrical separation of 150 Å from each other in space and randomly rotated around their center of mass. Then both molecules are allowed to rotate to minimize the intermolecular energy function. After the translations and rotations, ligand and receptor are docked by rigid body energy minimization. The best decoys are selected based on intermolecular energies.

2. Semi rigid simulated annealing in torsion angle space.

First, the two proteins are presumed as rigid bodies and their particular orientation is optimized. Then the side chains at the interface are allowed to move in. Further, chains and backbone at the interface are allowed to move in with conformational rearrangements. The best top scoring models 200 out of 1000 from the 1st stage are refined using a simulated-annealing (SA) method in torsion-angle space procedure during which the interface is treated as flexible by taking first side-chains only, then both side-chains and backbone.

3. Final refinement in Cartesian space with explicit solvent.

Models from simulated annealing are subjected to molecular dynamics simulation in an explicit solvent shell. In the final phase, they are solvated in a 8Å shell of TIP3P water molecules. In the TIP3P model, each atom is assigned a point charge, and the oxygen atom gets the Lennard-Jones parameters [201]. A large number of energetically improved structures generate the problem of selecting the better results. However, comparative analysis of micromolecular property, such as steric surface complementarity, electrostatic interactions, hydrogen bonding and knowledge based pair-potentials, desolvation energies and van der Waals interactions helps to find better scored model [202-211].

2.3.8 Molecular Dynamics Simulation (MD simulation):-

Computer based MD simulations act as a connection between microscopic length and time scales i.e. to develop the prospect of the complication that splits 'solvable' from 'unsolvable'. MD simulation ascertains the motion of atoms on a multi-dimensional potential energy landscape resultant to one electronic state using classical mechanics or quantum mechanical equation. MD simulation imposes Newtonian mechanics on a particle system to solve the time evolution of a set of interactions. MD simulation comprises of the mathematical, solution of the conventional calculations of motion.

$$\mathbf{F}_i = m_i \frac{\mathrm{d}^2 \mathbf{r}_i(t)}{\mathrm{d}t^2} \qquad m_i \ddot{\boldsymbol{r}}_i = \boldsymbol{f}_i \qquad \boldsymbol{f}_i = -\frac{\partial}{\partial \boldsymbol{r}_i} \mathcal{U} \qquad \textbf{(1)}$$

In order to compute the forces f_i acting on the atoms, which is originate from a potential energy U (r^i), where r_i = (r1; r2; ….. r_i) symbolizes the entire set of 3N atomic coordinates. The position and velocity vector describes the time evolution of the system in phase space, which is defined in the MD trajectory. The objective of the numerical integration of Newton's equations of motion is to find an expression that describes positions r_i ($t_1\Delta t$) at time $t_1\Delta t$ in terms of the previously known positions at time t. Verlet algorithm is commonly used in MD simulations because of its simplicity and consistency. The elementary formula for this algorithm is derivative of the Taylor series for the positions r_i (t); it. Reads as in equation

$$\mathbf{r}_i(t+\Delta t) \cong 2\mathbf{r}_i(t) - \mathbf{r}_i(t-\Delta t) + \frac{\mathbf{F}_i(t)}{m_i}\Delta t^2 \qquad \textbf{(2)}$$

Equation [2] is precise up to the fourth power in Δt. Velocities can be obtained from the propagated explicitly as in alternative velocity Verlet scheme [212]. The atomic interactions and potential energy of the particle system can be described by the potential U (r1… rN), which signifies the potential energy of N interacting atoms as a function of their locations r_i= (x_i, y_i, z_i). Given the potential, the force acting upon i_{th} atom is obtained by the gradient with respect to atomic displacements, as shown in equation

$$\mathbf{F}_i = -\nabla_{\mathbf{r}_i} U(\mathbf{r}_1, \cdots, \mathbf{r}_N) = -\left(\frac{\partial U}{\partial x_i}, \frac{\partial U}{\partial y_i}, \frac{\partial U}{\partial z_i}\right) \qquad \textbf{(3)}$$

The physical and chemical properties in a functional form of a system are defined by empirical potentials, which are defined by the atomic force field. The adaptable constraints

are chosen such that the empirical potential signifies a good fit to the appropriate sections of the *ab initio* Born–Oppenheimer surface [213]. The atomic force field model defines physical systems as collections of atoms kept together by inter atomic forces.

A characteristic force field, used in the simulations of bio-systems, where the potential energy *U* is calculated from bonded and non-bonded interactions can be written as

$$V = E_{bonded} + E_{non-bonded} = \\ = \sum_{bonds} k_b(r - r_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \sum_{torsions} \frac{V_n}{2}[1 + \cos(n\tau - \phi)] \\ + \sum_i \sum_{j>i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}^{12}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}^{12}}\right)^{6}\right] + \sum_i \sum_j \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_{rel} r_{ij}} \quad \textbf{(4)}$$

where $r_{ij} = r_i - r_j$, *kb* is the bond stretching constant, r_0 equilibrium bond distance, K_θ is the bond angle constant, θ_0 is the equilibrium bond angle, τ is the torsion angle, ϕ is the phase angle, and V_n is the torsional barrier. The last two non-bonded terms in the potential are Lennard-Jones potential and coulomb interaction, in which is the van der Waals well depth, σ is the van der Waals diameter, q is the charge of each atom, and ε is dielectric constant. Biomolecular force fields are generally accounts for long-range electrostatic and dispersive interactions and an additive of order N^2 required accounting for all non-bonded pairs. The new positions and velocities are upgraded after each stepwise numerical integration which also requires the forces acting upon the atoms have to be recomputed at each step. An aqueous medium is the illustrative environment for biological macromolecules and it has to be accounted for in realistic simulations. The treatment of long-range forces is associated to the choice of boundary conditions imposed on a particle system to deal with its limited size and surface effects. The two common techniques are generated on both periodic boundary conditions and the Ewald method for lattice summations or on spherical boundary conditions

and the reaction field method. Particle meshed Ewald method or fast multiple methods allows more effective computation of the long-range interactions and do not accounts to basic cutoff approximation, in which contributions of sites separated by distance larger than a certain cut-off are abandoned.

Molecular modeling & docking Protocol for BARD1 ARD1-BRCT and CstF50:-

1. Protein structure of BARD1 ARD-BRCT, ARD domain (425-550aa) and BARD1BRCT (568-777) amino acids and CstF50 (1-431aa) was modeled using Robetta server.
2. Good-quality model was selected based on overall stereochemistry, Ramachandran plot and "**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 100 ns using GROMACS [41-43].
4. Monomer BARD1 ARD-BRCT-CstF50 geometrical docking was carried out using HADDOCK server [211].
5. PDBsum was obtained to analyze the interactions.
6. DALI server is an important tool that predicts the best structure homologs of the novel protein [214]. It has been reported that structurally similar proteins most likely play similar role in a particular biological occurrence [215].

Protocol for BARD1BRCT domain and mutant's simulation and mutational analysis:-

1. Protein structure was downloaded from (PDB ID: 2NTE).

2. Mutants were generated by swiss PDB viewer and chemical quality, Ramachandran plot and validated by "**SAVES**" server.

3. Protein was simulated for 50 ns using GROMACS [41].

4. Monomer-monomer geometrical docking was carried out using HADDOCK server [197].

5. PCA analysis and clustering were performed on the *wild-type* and mutant trajectory using R 3.2 and Prody software [216, 217].

6. 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 [218].

2.3.9 Principal Component Analysis (PCA):-

In cellular milieu, proteins are dynamic in nature and extraction of physical principles that govern such directional dynamics may prove crucial in understanding their mode of functioning at the molecular level. MD simulations of macromolecules can mimic the cellular conditions, but extract the meaningful information from large diverse data set can be derived from PCA [219, 220]. A characteristic MD trajectory contains the data of time-evolution of the coordinates of all the component atoms of the system being investigated. Usually, MD time steps are roughly 1 fs while the simulation time may fluctuate from a few to tens of nanoseconds. So, a single resultant trajectory after MD simulation comprise huge amount of data. For an *N*-atom system, the input dataset for PCA can be constructed as a trajectory matrix and each column comprises a cartesian coordinate for a given atom at each output time step ($\boldsymbol{x}\ (t)$). Prior to performing PCA, it is mandatory to remove any net translational and

rotational motion of the system by aligning the coordinate data to a reference structure (input structure) to get the appropriate trajectory matrix (*X*). The standardized trajectory data is then used to generate a covariance matrix (*C*) and elements matrix can be defined as

$$C_{ij} = < (X_i - <X_i>) (X_j - <X_j>) \quad \textbf{(1)}$$

Where <···> signifies an average done over all the time steps of the trajectory. The following step comprises of diagonalization of the 3*N*x3*N* covariance matrix and can be accomplished via eigenvector decomposition method as where *T*

$$C = T\Lambda T^T \quad \textbf{(2)}$$

is a matrix of column eigenvectors and Λ is a diagonal matrix comprising the corresponding eigenvalues. This procedure leads to transformation of the actual trajectory matrix in a new orthonormal basis set composed of the eigenvectors. The eigenvalues are indicative of the mean squared displacements of atoms along the matching eigenvector. There will be 3*N* following eigenvalues if the number of arrangements (*M*) is more than 3*N*. If *M*<3*N* there will be the number of eigenvalues will decrease to *M*.

The simplest mode of visualizing these results involves categorization of eigenvectors in a descending order with respect to their eigenvalues. The plot of eigenvalues against the respective eigenvector index can then be plotted as 'scree plot'. Typically, a scree plot indicates that only a few first eigenvectors possess large eigenvalues and higher eigenvector shows lower eigenvalues. As a result, most of the variance or atomic motion in the data is content and illustrated by only a few first modes. It is then imperative to postulate that the motions along these 'essential eigenmodes' explain the most significant global information which shows predominantly the dynamics of the macromolecular systems. Visualization of

the components of individual eigenvector can be helpful to estimate the landscape of the eigenmodes and atomic fluctuation. After the identification of a subset of essential eigenmodes, further investigation specifying each mode can be undertaken by projecting the native trajectory along a specified eigenvector. The respective projection matrix (*P*) can be acquired as

$$P=XT \quad \textbf{(3)}$$

The time evolution given by the projection matrix yields excitation amplitude of a given eigenvector. The column vectors in *P* (**p**(t)) are termed as the 'principal components'. To investigate the motion along any specified eigenvector, the column vector from *P* multiplied by the respective eigenvector in *T* yields a compact trajectory comprising motion only along the selected mode. First 20 modes which occupy high eigenvalue position and thus contain the essential information of the macromolecular dynamics.

2.3.10 Normal Mode Analysis (NMA):-

The vital goal while examining the protein dynamics is to explain slow large-amplitude motions. These motions characteristically indicate conformational fluctuations which are crucial for the functioning of proteins. Normal mode analysis of protein can provide valuable insight regarding the mechanism of slow large-amplitude motions within multidomain protein assembly in a time independent manner [221, 222]. Normal mode analysis (NMA) is a quick and non-expensive technique to calculate vibrational modes and protein flexibility. In NMA the modeled atoms are considered as point masses associated by springs, which indicate the interatomic force fields. In NMA each interatomic bonds are considered as springs connecting each atom (which is also called as node) to all other neighboring nodes are of equal strength,

and only the atom inter-connected within a limited distance are accounted. A standard NMA requires a set of Cartesian coordinates, a molecular force field defining the interactions between constituent atoms. NMA in Cartesian coordinate space usually involves three main calculation steps:

(1) Minimization of the conformational potential energy as a function of the atomic Cartesian coordinates.

(2) The calculation of the "Hessian" matrix, which is matrix of second derivatives of the potential energy with respect to the mass-weighted atomic coordinates.

(3) The diagonalization of the Hessian matrix. This final step produces eigenvalues and eigenvectors (normal modes).

All three steps are computationally demanding and depend on the size of the macromolecular system. Usually, the first and final steps are the bottlenecks. Normally, energy minimization and diagonalization are computationally demanding (CPU time and memory) because it encompasses the diagonalization of a 3N×3N matrix, where N is the number of atoms in the particle system. It is usually considered that at a potential energy minimum, the potential energy function V can be expanded in a Taylor series in terms of the mass-weighted coordinates. Which can be written for Cartesian coordinate (q_i) as $qi=\sqrt{mi}\Delta x_i$, where Δx_i is the displacement of the $_i$th coordinate from the energy minimum and m_i is that mass of the respective atom. If the expansion is completed at the quadratic level, then since the linear term is zero at an energy minimum: the potential energy function for the particle system can be written as

$$V = \frac{1}{2} \sum_{i,j=1}^{3N} \left. \frac{\partial^2 V}{\partial q_i \partial q_j} \right|_0 q_i q_j \qquad \textbf{(1)}$$

The second derivatives in Eq. 1 can be transcribed in a matrix "Hessian," F. Determination of its eigenvalues and eigenvectors (equivalent to diagonalization) indicates:

$$Fwj = \omega^2_j w_j \qquad \textbf{(2)}$$

where wj is the jth eigenvector and $\omega^2 j$ will be jth eigenvalue. There are 3N such eigenvector equations so each eigenvector classifies a normal mode coordinate through an equation:

$$Q_j = \sum_{i=1}^{3N} w_{ij}\, q_i . \qquad \textbf{(3)}$$

Yet, assessment of low-frequency normal modes and the directions of large-amplitude fluctuations in molecular dynamics simulations designate a close and clear similarity [219, 220, 223]. Close directional coincidence of the first few normal mode and the first principal component acquired from molecular dynamic simulations has been detected [223].

2.4 Protein oligomerization and size characterization:-

Protein oligomerization can be distinguished and characterized using different approaches. Analytical ultracentrifugation, size exclusion chromatography, scattering techniques, NMR spectroscopy and mass spectrometry are day to day techniques to detect and quantify oligomerization [224-226].

2.4.1 Glutaraldehyde Cross-Linking:-

Protein-Protein Interactions can be temporary or stable depending upon their role in several biological functions [227]. However, chemical cross-linking helps to find both transient and stable interactions involved between two proteins [228]. In chemical cross-linking, covalent bonds forms between bi-functional reagents such as glutaraldehyde containing reactive end group. The reactive end group reacts with the protein functional groups, which are primary – NH_2 and sulfhydryl's of amino acid residues. Furthermore, cross-linking experiment can also reveal the areas of interacting interface between two interacting proteins.

Protocol for BARD1 ARD1-BRCT and Gln564His glutaraldehyde cross-linking:-

1. FPLC purified wild-type and mutant proteins at concentration 0.5mg/ml were prepared.
2. Reaction mixtures with 10 µg of protein in 10 mM HEPES buffer (pH 7.5) were incubated with freshly prepared solution of glutaraldehyde (final concentration 0.1%) for 2 minutes at, 37°C in a time dependent manner [229].
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 over 12% SDS-PAGE.

Protocol for BARD1BRCT *wild-type* and mutants glutaraldehyde cross-linking:-

1. FPLC purified *wild-type* and mutant proteins at concentration 1 mg/ml were prepared.
2. Reaction mixtures with 15 µg of protein in 10 mM HEPES buffer (pH 7.5) were incubated with freshly prepared solution of glutaraldehyde (final concentration 0.1%) for 2 minutes at, 37°C in a time dependent manner [229].

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 over 12% SDS-PAGE.

2.4.2 Dynamic Light Scattering:-

Dynamic Light Scattering (Photon Correlation Spectroscopy) measures time-dependent fluctuations in the scattering intensity from particles undergoing arbitrary Brownian motion. Dynamic scattering is less precise in differentiating small oligomers and dimer from monomer. However, DLS is capable in measuring several other biophysical parameters such as molecular weight, radius of gyration, translational diffusion constant etc.

Dynamic Light Scattering Protocol for BARD1 ARD-BRCT and Gln564His mutant:-

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

Dynamic Light Scattering Protocol for Protocol for BARD1BRCT and mutants:-

1. Molecular size measurement was performed using Brookhaven 90 plus particle size analyzer (Brookhaven, NY).
2. 1 mg/ml *wild-type* and mutant proteins were scanned at an interval of 5 minutes for 15 minutes and effective diameter of each measurement was considered.

2.5 Crystallization:-

A three-dimensional structural analysis of a protein is essential to understand the protein function and design small molecule effectors to modify its function. During crystallization, purified protein forms nucleation from an aqueous solution. Discreet arrangement of protein molecules in the form of lattice are held together by non-covalent interactions, and make up smallest "unit cells". Protein crystallography is used to determine protein's three-dimensional structure using X-ray diffraction method.

The diffraction pattern can then be processed to derive the three-dimensional structure of the protein. In order to produce a well-shaped crystal, appropriate uniformity with purity of the protein is required at an optimal pH conditions [230]. For crystallization, protein should be present in suitable buffer system and precipitants.

Several elements can affect crystallisation of proteins, which comprises protein purity, concentration, temperature, pH, ionic strength, and volume of crystallization solution, etc. Process of crystallization is distinguished into two steps: Nucleation process and crystal growth.

Nucleation and growth can happen in the supersaturated regions as represented in phase diagrams (**Figure: - 2.1**). The illustration mainly includes three regions; unsaturated region, saturated region, and supersaturated region. Crystal grows in saturated or supersaturated region while nucleation most habitually starts in supersaturated region. The major focus of crystallization process is to obtain diffraction quality crystals. The best method to get crystals is through organized exposure of the protein solution to a diverse range of buffers, which have different amalgamations of suitable precipitants, such as salts, polyethylene glycols. For better

crystallization, the protein-precipitant mixture must reach to the nucleation phase gradually to deliver necessary time for crystal growth. Most commonly used precipitants in crystallization trial are salts of sodium, ammonium and potassium ($(NH_4)_2SO_4$, NaCl, KH_2PO_4), organic polymers (PEG) and alcohols (methanol, ethanol, propanol, acetonitrile). Salt generally preserve the ionic environment of the solution although organic polymers diminishes protein solubility by dropping the dielectric constant of solvent. Solubility of a protein is found lowest at its isoelectric point (pI) since protein carries a net zero charge.

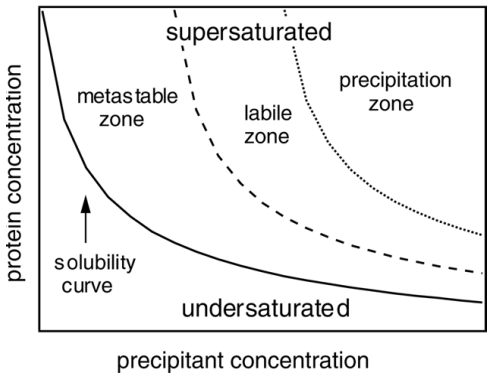


Figure 2.1: The phase diagram of crystallization.

2.5.1 Methods of crystallization: Different strategies have been established to grow crystals of functionally important proteins. Some of the more frequently used methods are vapor diffusion, micro-batch, dialysis and liquid–liquid diffusion technique.

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

In both the methods, a drop of purified protein mixed well with buffer and precipitant is present in a closed system and permitted to equilibrate with reservoir solution comprising the same precipitant. In the beginning, protein and precipitant concentration in the droplet is different, but as the system equilibrate 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 but potent technique since it consumes less amount of protein, and several crystallization conditions can be explored with a limited volume of protein. It is also appropriate for crystal growth monitoring, crystal manipulation and harvesting.

2) **Micro batch**: The micro batch method of crystallization is introduced to resolve the problem of microcrystal formation. In micro batch system a low density paraffin oil (~0.87 mg/ml) which float on the surface of protein-precipitant mixture, thus diminishes the rate of evaporation can be used [231]. In this process, a mixture of paraffin and silicon oil in combination can also be used. It is an excellent technique for filtering the known crystallization conditions to optimize the crystal growth [232].

3) **Dialysis**: The dialysis process is based upon the principal of slow diffusion of inorganic molecules through a semi permeable membrane, which leads to supersaturation of protein solution. The system tries to maintain equilibration between precipitant and solute particles by permitting selective passage of water and precipitant. The dialysis bag is kept in a chamber containing the pool solution with a precipitant [233].

4) **Liquid–liquid diffusion method**: In this method, protein and precipitant solution keep a direct contact with the dense solution at the bottom. A concentration gradient is created due to the diffusion of protein and precipitant and thus crystallization may occur at suitable protein and precipitant concentration. Free interface diffusion is a method of choice for fine-tuning the crystallization conditions.

3.0 Domain organization of expression contructs:-

BARD1 ARD-BRCT: BARD1 ARD-BRCT (425-777aa) *wild-type*, BARD1 ARD-BRCT (425-777aa) Gln564His mutant domains of BARD1 ARD-BRCT were sub-cloned in pGEX-

kT vector (Amersham Pharmacia). BARD1 ARD-BRCT domain (425-777aa) was expressed in bacterial system Rosetta 2D3 (DE3) strain. BARD1 N-terminal ring domain (16-119aa) and BRCA1 (1-303aa) was also expressed in bacterial system Rosetta 2(DE3) strain.

BRCA1: BRCA1 N-terminal ring domain (1-303aa) was sub-cloned into bacterial expression vector pET-41a using one-step cloning method as described earlier.

BARD1 BRCT and mutants: BARD1BRCT (568-777aa) region was sub cloned into pET28a vector. BARD1BRCT domain mainly comprises of C-terminal region, which includes phosphopeptide interacting region. Site-directed-mutagenesis with mismatch primers for respective mutations was performed over the *wild-type* template. All the reported mutations were confirmed by DNA sequencing. *Wild-type* and mutant construct was transformed into bacterial expression host, *E. coli* BL21 (DE3) strain, and the fusion protein was expressed.

Chapter

3

Biophysical characterization and

Molecular mechanism

Of BARD1 ARD-BRCT

CstF50 binding

Interface

3.1 Introduction:-

Germ-line mutations in *Breast cancer susceptibility gene 1 (BRCA1)* result in early predisposition of 40–50% familial breast cancer cases and >75% cases of hereditary breast and ovarian cancer [53, 141, 234, 235]. The gene product of *BRCA1* comprises of 1863 amino acid protein that interact with different cellular partners such as BARD1, BAP1, ABRAXAS, CtIP and RAP80 [65]. BARD1 in association with BRCA1, present itself as E3 ubiquitin ligase and helps in tumor suppression function of BRCA1 [21, 236-239]. CstF50 (Cleavage Stimulating Factor 1) is key reported binding partner of BARD1 and has an important role in processing and polyadenylation of premature mRNA transcripts. BARD1 is the only binary interacting partner reported for CstF50 until now [12, 36, 131].

The polyadenylation reaction is a series of steps consisting of an endonucleolytic cleavage followed by synthesis of the poly(A) tail [130]. Cleavage stimulation factors ($CstF_s$) are the essential processing factors, responsible for processing of 3'-end of premature mRNA transcript generated by RNA polymerase II [240]. CstF50 is a 50 KDa protein implicated in direct interaction with the RNA polymerase II CTD (C-terminal domain) through its N-terminal region [130, 131]. BARD1 interacts with 7^{th} WD40 repeat of CstF50 through the flexible linker region residues from 554-568 [12]. It has been reported that 3'-end processing is repressed after UV-induced DNA damage as a result of proteasome mediated degradation of RNAP II by the BRCA1–BARD1–CstF50 complex [6]. Depletion of CstF50 enhances sensitivity to UV irradiation and reduces the ability of BRCA1-BARD1 to ubiquitinate RNAP II [132] and further leads to cell-cycle arrest and apoptosis [241]. CstF50 is also reported to interact with the DNA replication and repair factor PCNA [12]. PCNA co-localizes with BRCA1/BARD1 at the sites of DNA damage [72, 78] and in association with

other DNA damage repair proteins [133]. Supporting these facts, studies have shown that cells lacking BRCA1 have defective transcription coupled repair mechanism [8, 242]. Since, the results of two studies link BRCA1-BARD1 to the stimulation of ubiquitination of RNAPII in non-breast cells, signifying that the BRCA1-BARD1 mediated RNA Polymerase-II CTD ubiquitination and degradation is not a breast and ovarian cell-specific phenomenon. Genomic analysis of BARD1 in 58 ovarian tumors, 50 breast tumors and 60 uterine tumors has been previously performed and mutational investigation leads to detection of seven different polymorphism [27]. Cancer predisposing mutation in BARD1 ARD-BRCT linker region Gln564His was found in the germ-line of a patient with adenocarcinoma along with loss of the *wild-type* BARD1 allele [27]. BARD1 Gln564His mutation abrogates the p53 and CstF50 binding to BARD1 as demonstrated by GST pull down assay [6, 95]. Loss of binding between BARD1 and CstF50 leads to the formation of premature transcript and uncontrolled mRNA regulation. The BARD1 Gln564His mutation has shown to induce apoptosis less efficiently when transfected in mammalian cells, indicating that the BARD1 linker region is necessary for the tumor-suppressor and pro-apoptotic functions [95]. These findings suggest role for BARD1 mutations in the development of sporadic and hereditary tumors.

3.2 Material and Methods:-

A comparative analysis of BARD1 ARD-BRCT *wild-type* and mutants were performed using different biophysical and in-silico tools.

3.2.1 In-silico analysis: Interaction analyses between BARD1 *wild-type* and BARD1 Gln564His mutant with CstF50 were carried out using HADDOCK server [196].

3.2.2 Gene Cloning: *bard1 ard-brct* (encoding 425-777aa) *wild-type* (cDNA of full length BARD1, a kind gift from Richard Baer, USA) was sub-cloned in pGEX-kT vector (Amersham Pharmacia). *bard1 ard-brct* (425-777aa) Glu564His substitution was generated using the mismatch primers considering *wild-type* construct as template. Different regions of CstF50 (1-431aa) and (92-431aa) (cloned in pDONR vector which was purchased from UCLA) were sub-cloned in to pGEX-kT and pET-28a vectors. The basic scheme for sub-cloning was used and oligonucleotide primers complementary to the template sequence were used in PCR reactions and their details are mentioned below in primer details.

The genes of interest were amplified using following PCR condition: 95°C denaturation (5 minutes), 95°C denaturation (45 seconds), and annealing 62°C for 35 seconds, extension 72°C at 0.5kb/min, final extension 72°C for 10 minutes, and 25 cycles. The PCR amplified products were digested with EcoR1 and BamH1 and ligated in the pGEX-kT and pET-28a vectors (pGEX-kT was a kind gift from John Ladias, BIDMC-HMS, Boston). The ligation mixture was transformed into *E.coli* DH5α cells. Colony screening were performed and positive clones were selected after the insert release from the plasmid using EcoR1 and BamH1 restriction enzymes, which was further confirmed by DNA sequencing. Sequentially correct clones were used for protein expression and purification.

3.2.3 Primer details

BARD1 ARD-BRCT domain (425-777aa) *wild-type*

FORWARD PRIMER

5'-GTCGGATCCCATATGGAGAACCTGTACTTTCAGGGTAACCATCGTGGTGAGACTTTGCTCCAT-**3'**

REVERSE PRIMER

5'-GTCGGATCCGAATTCCTATTAGCTGCTAAGAGGAAGCAACTC-**3'**

FORWARD PRIMER BARD1 Gln564His

5'-CAGTAATGAACACTGGGCACCGTAGGGATGGACC-**3'**

REVERSE PRIMER BARD1 Gln564His

5'-GGTCCATCCCTACGGTGCCCAGTGTTCATTACTG-**3'**

Primer details for CstF50 (1-431aa)

FORWARD PRIMER

5'-GTCGGATCCGAGAACCTGTACTTTCAGGGTATGTACAGAACCAAAGTGG-**3'**

REVERSE PRIMER

5'-GTCGAATTCCTATTAGTCAGTGGTCGATCTCCGG-**3'**

BARD1 ARD-BRCT (425-777aa) and Gln564His mutant were expressed in Rosetta 2(DE3) bacterial strain. For protein expression, 50 ng/µl plasmid constructs were transformed into *Rosetta 2(DE3)* cells and grown on LB agar plate containing ampicillin (100 µg/ml). CstF50 (1-431aa) was expressed in Rosetta 2(DE3) bacterial system. For protein expression, 50ng/µl plasmid constructs were transformed into *Rosetta 2(DE3)* cells and grown on LB agar plate containing kanamycin (100 µg/ml). Detailed protocol is described below.

3.2.4 Protocol for protein purification of BARD1 ARD-BRCT (425-777aa) and Gln564His mutant:

Purification buffer: 50 mM Tris (pH 6.2), 300 mM NaCl, 0.1% triton, 5% glycerol pH 6.2 (Buffer A)

FPLC buffer: 20 mM Tris (pH 7.5), 100 mM NaCl, 1 mM 0.1% triton (Buffer B)

1. **Re-suspension**: Re-suspended the pellet of BARD1 ARD-BRCT (425-777aa) and Gln564His mutant in 60 ml of buffer A; supplemented with 200 mM PMSF and 20µl of protease inhibitor.

2. **Ultra sonication**: Transferred the resuspended cell pellet into centrifuge tube and sonicated at 70 pulse rate and 70 power with 1 minutes of duty cycle. Repeated the sonication cycle 5 times with 1 min break.

3. **Centrifugation:** After sonication, the suspension was subjected to centrifugation at 18000 rpm for 45 minutes at 4^{o}C to obtain cleared lysate. The soluble protein in the supernatant was collected and cell debris was discarded.

4. **Binding**: The soluble *wild-type* and mutant protein fractions were allowed to pass from column containing pre-equilibrated affinity resin. The sepharose beads in affinity column charged with glutathione binds specifically to proteins having GST (Glutathione-S-transferase) tag. The protein bound resin ~ 10 µl was loaded onto SDS-PAGE gel to check the binding.

5. **Washing**: After binding, the protein bound column was washed with 10 column buffer A to remove any non-specific protein.

6. **Cleavage**: Added 400 µl of TEV protease (20 units), 40 µl of protease inhibitor cocktail and 100 µl of PMSF in 20 ml of buffer A and completed the cleavage step in 3 hours by passing the TEV containing buffer solution through column at an interval of 1 hour. Nearly 10 µl of protein bound beads were loaded onto 12 % SDS-PAGE get to detect the cleavage of protein.

7. **Elution**: After TEV cleavage, the protein was eluted with 30 ml of buffer A.

8. **Equilibration of Ni-NTA resin**: Provided 2 column washes with double distilled water and then 5 to 6 column washes with buffer A.

9. **Metal Ion Chelate Affinity Chromatography**: After calibration of Ni-NTA resin, passed the eluted fractions through Ni-NTA to get rid of His-tagged TEV protease contamination.

10. **Concentrating the protein:** Transferred the eluted protein in a 10 KDa pre-equilibrated centricon and concentrate the protein up to 2 ml by centrifuging at 4500 rpm for 10 minutes at 4^{o}C. Check the concentration on Nanodrop spectrophotometer (280 nm). Centrifuged for 10 minutes at 13000 rpm at 4^{o}C for removal of insoluble aggregates or precipitates.

11. **Gel filtration:** 2 ml of concentrated protein was injected in AKTA- FPLC against buffer B.

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

13. **Loading on 12% SDS-PAGE gel**: Loaded 20 µl of FPLC fractions on 12% SDS-PAGE, stained with coomassie dye, and then destained it to visualize the protein of interest.

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

Protocol for protein expression of CstF50:

1. **Inoculation:** picked a single transformed colony of pET-28a-6HIS-CstF50 from antibiotic resistant LB agar plate and inoculated it in 100 ml LB broth preinoculum containing 100 µg/ml of kanamycin. Grown at 37^{o}C overnight in shaking condition.

2. **Dilution:** Added 10 ml of pre-inoculum to 1000 ml (1:100) of autoclaved LB broth containing 100 µg/ml of kanamycin. Grown the medium on a shaker incubator at 37^{o}C until it has reached mid-log phase i.e. A_{600} between 0.6-0.8.

3. **Induction:** Cooled down the flasks and added 100 µl IPTG (stock 1M), and incubate on shaker incubator at 20^{o}C for 18 hours.

4. **Harvesting:** The culture was centrifuged for 10 minutes at 6000 rpm at 4^{o}C. The pellet was resuspended in a small volume of supernatant and centrifuged for 15 minutes at 5000 rpm, 4^{o}C.

5. **Storage:** The bacterial pellet obtained was stored at -80^{o}C for further use.

Proteins were purified by affinity chromatography followed by FPLC.

3.2.5 Protocol for purification and refolding of CstF50:

Cell lysis buffer*:* 50 mM Tris (pH 6.2), 500 mM NaCl, 0.1% triton, 5% glycerol pH 7.2, 8M urea (Buffer A)

Pellet washing buffer*:* 50 mM Tris (pH 6.2), 500 mM NaCl, 0.1% triton, 5% glycerol pH 7.2, 1% sodium do-decyl sulphate and 1% lauryl sulphate. (Buffer B)

CstF50 solubilization buffer*:* 50 mM Tris (pH 6.2), 500 mM NaCl, 0.1% triton, 5% glycerol pH 7.2, 8M urea (Buffer C)

Refolding buffer: 50 mM Tris (pH 6.2), 500 mM NaCl, 0.1% triton, 10% glycerol pH 7.2, 2mM EDTA, 700mM arginine, 100mM KCl, 50mM $MgCl_2$, 100mM $CaCl_2$, 5mM DTT and 0.2% triton (Buffer D).

FPLC buffer*:* 20 mM Tris (pH 7.5), 100 mM NaCl, 1 mM 0.1% triton, 200mM arginine (Buffer E).

1) **Re-suspension**: Re-suspended the pellet of CstF50 (1-431aa) in 100 ml of Buffer A supplemented with 200 mM PMSF and 20µl of protease inhibitor.

2) **Ultra sonication**: Transferred the resuspended cell pellet into centrifuge tube and sonicated at 70-pulse rate and 50 power with 1 minutes of duty cycle. Repeated the cycle 5 times with 1 min break.

3) **Centrifugation:** After sonication, the suspension was subjected to centrifugation at 18000 rpm for 45 minutes at 4^{o}C to obtain cleared lysate. Collected the supernatant and discarded the cell debris.

4) **Pellet washing:** The soluble fraction was discarded (CstF50 is insoluble and forms inclusion bodies) and pellet was washed with pellet buffer B (composition described above) by resuspending in the buffer B thoroughly and centrifugation at 18000r.p.m. Discarded the supernatant and washed the pellet again. Repeated the process thrice.

5) **Solubilisation in 8M urea:** Resuspended the pellet after washing in Buffer C. Resuspended the pellet in buffer C in such a way that no pellet clump should left over. Solubilized all the protein present in pellet by incubating the washed pellet in 8M urea solubilisation for four hours.

6) **Metal Ion Chelate Affinity Chromatography**: The soluble 6HIS-CstF50 in 8M urea was allowed to bind fraction obtained is brought at room temperature and then mixed with pre-equilibrated affinity resin (with Buffer C) and incubated at room temperature for 1 hour.

The Ni^{2+} in the affinity column binds specifically to those proteins having Hexa-histidine tag. Take out 40-µl bead for binding check.

7) **Washing**: After binding, the protein bound column was washed with 4-column wash buffer c to remove any non-specific protein bound to the affinity column.

8) **Elution**: Eluted the bound protein after passing buffer containing increasing concentration of imidazole (100-600 mM).

9) **Dialysis:** Eluted protein was dialyzed by using 10KDa dialysis filter bag in presence of the buffer D. After every 4 hrs buffer changes were provided to refold the protein and at every change, the concentration of arginine was reduced 100mM and five serial changes were provided**.**

10) **Concentrating the protein:** Transferred the eluted protein in a 10 KDa pre-equilibrated centricon and concentrated the protein up to 2 ml by centrifuging at 4500 rpm for 10 minutes at 4^oC. Check the concentration on nanodrop spectrophotometer (280 nm). Centrifuged for 10 minutes at 13000 rpm at 4^oC for removal of insoluble aggregates or precipitates.

11) **Gel filtration:** Injected 2 ml of concentrated protein in AKTA- FPLC against FPLC buffer.

    1. **Fraction collection**: Collected the purified protein obtained through FPLC in 1.7 ml microfuge tube at its elution volume according to gel filtration spectra profile of the sample.

12) **Loading on SDS-PAGE 10% gel**: Loaded 20 µl of FPLC fractions on SDS-PAGE, stained with coomassie dye, and then destained it to visualize the protein of interest.

13) **Concentrate the protein**: The fractions, which showed purified protein band was, concentrated as per the requirement.

3.3 Results and discussion

3.3.1 Cloning, expression and purification of BARD1 ARD-BRCT domain (425-777aa), BARD1 ARD-BRCT Gln564His mutant and functional domains of CstF50:- Selected potential clones when digested with the EcoR1 and BamH1 restriction enzymes showed the insert release of correct size **(Figure:- 3.3.1)**.

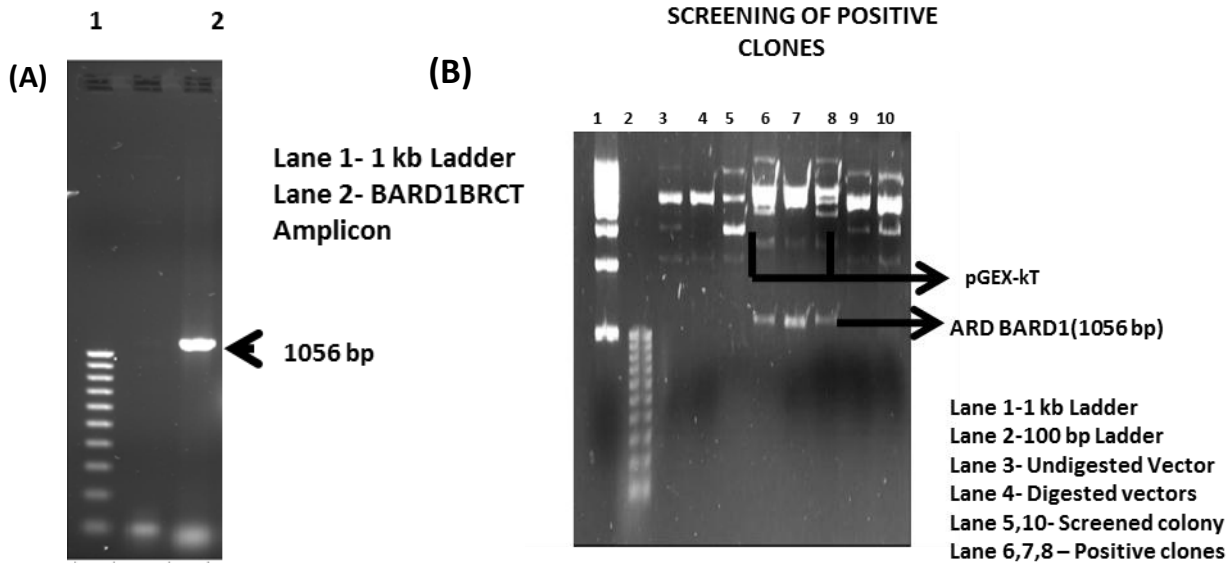


Figure 3.3.1:- **(A)** PCR and **(B)** Cloning of BARD1 ARD-BRCT *wild-type*

DNA sequencing results confirmed the presence of ligated gene of interest in the vector with desired frame of codon sequence. Purified BARD1 ARD-BRCT (425-777aa) *wild-type* and Gln564His mutant showed similar solubility in the identical buffer and pH conditions

(Figure:- 3.3.2). This indicates that Gln564His mutation does not change the solubility of the *wild-type* protein.

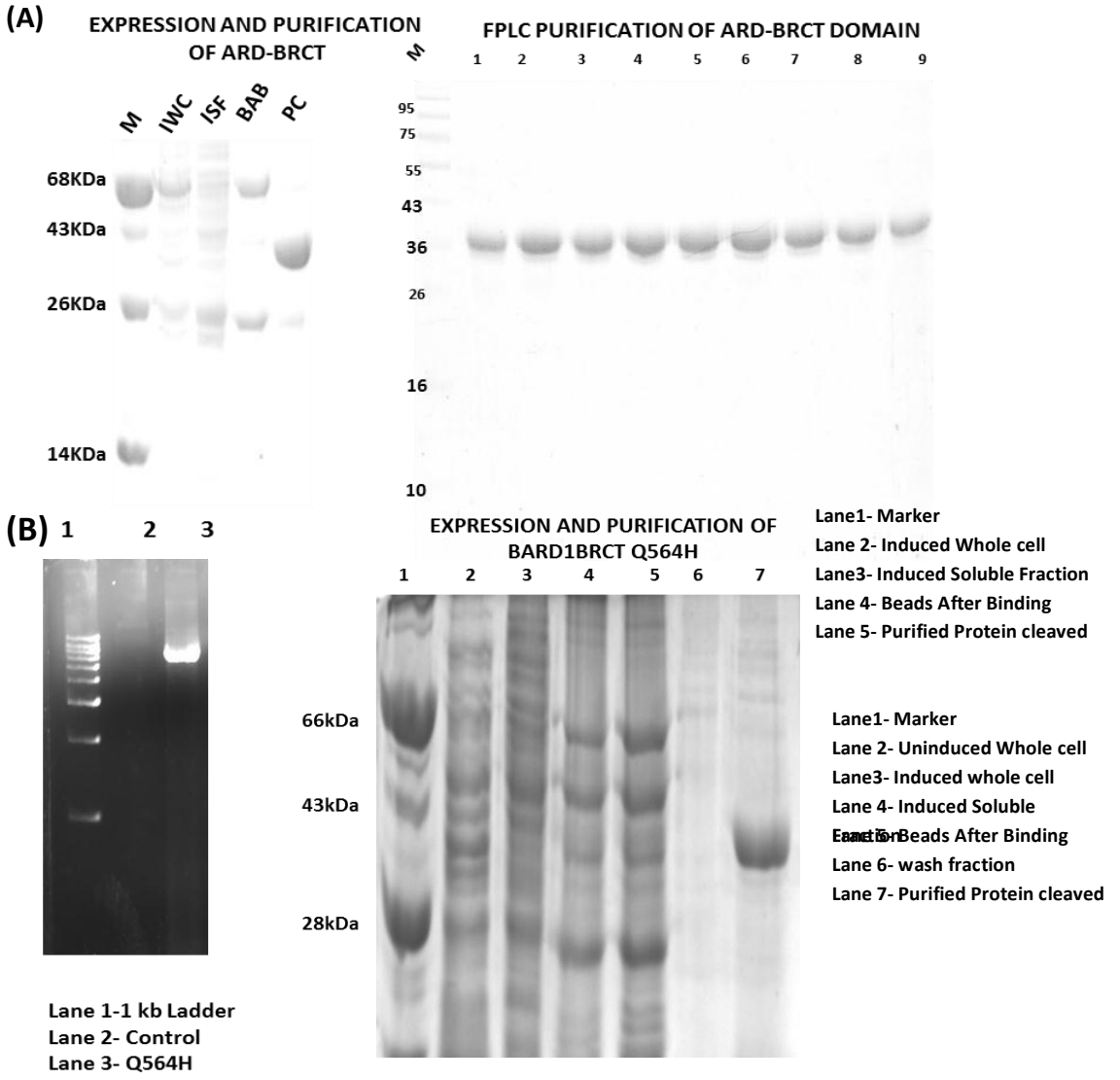


Figure 3.3.2:-Expression and purification profile of **(A)** BARD1 ARD-BRCT *wild-type* and **(B)** cloning and purification of Gln564His mutant.

3.3.2 Structural insights into BARD1 ARD-BRCT domain:-

To comprehend the thermodynamic and biophysical parameters of BARD1 ARD-BRCT *wild-type* and Gln564His mutant, gel filtration chromatography, thermal denaturation, and

chemical denaturation were performed. Molecular exclusion analytical chromatogram of BARD1 ARD-BRCT and mutant shows that BARD1 ARD-BRCT *wild-type* and mutant exist mostly in the monomeric form, and mutation has not affected the monomeric property of the *wild-type* protein. DLS performed for *wild-type* and mutant did not show any large change in the hydrodynamic size of the proteins providing evidence that mutation is not changing the oligomeric status of the *wild-type* protein (**Figure:-3.3.3).**

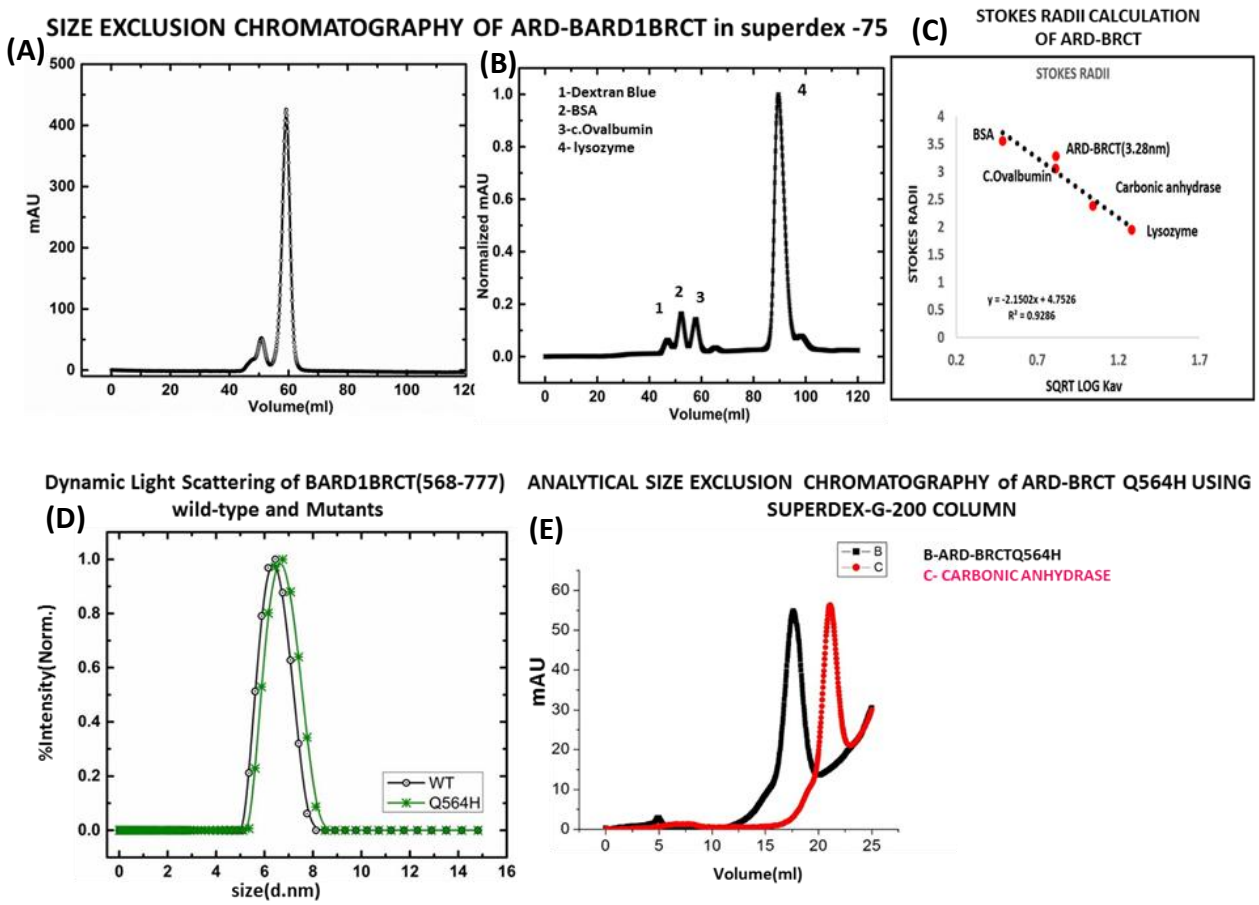


Figure 3.3.3: - In-solution oligomeric characterization of BARD1 ARD-BRCT *wild-type* and Gln564His mutant. **(A)** *Wild-type* elution profile**, (B)** Elution profile of standard proteins **(C)** Stokes radii calculation, **(D)** DLS profile of wild-type and mutant**, (E)** Elution profile of mutant.

Glutaraldehyde crosslinking experiment with BARD1 ARD-BRCT *wild-type* also yielded similar results as, there was no higher molecular weight aggregates conjugates were visible on the coomassie brilliant blue counterstained gel **(Figure:- 3.3.4)**. Secondary structural components of BARD1 ARD-BRCT and mutant were characterized with far-UV Circular-Dichroism **(Figure: - 3.3.5 A)**. To determine independent behavior of ARD domain of BARD1 from its BRCT domain, we have performed limited proteolysis in a time-dependent manner using trypsin. After treating the *wild-type* and mutant for limited time period with equal concentration of enzyme, same domain stability of *wild-type* and mutant was observed and mutation had not brought any significant changes on the structural compactness **(Figure:- 3.3.5)**.

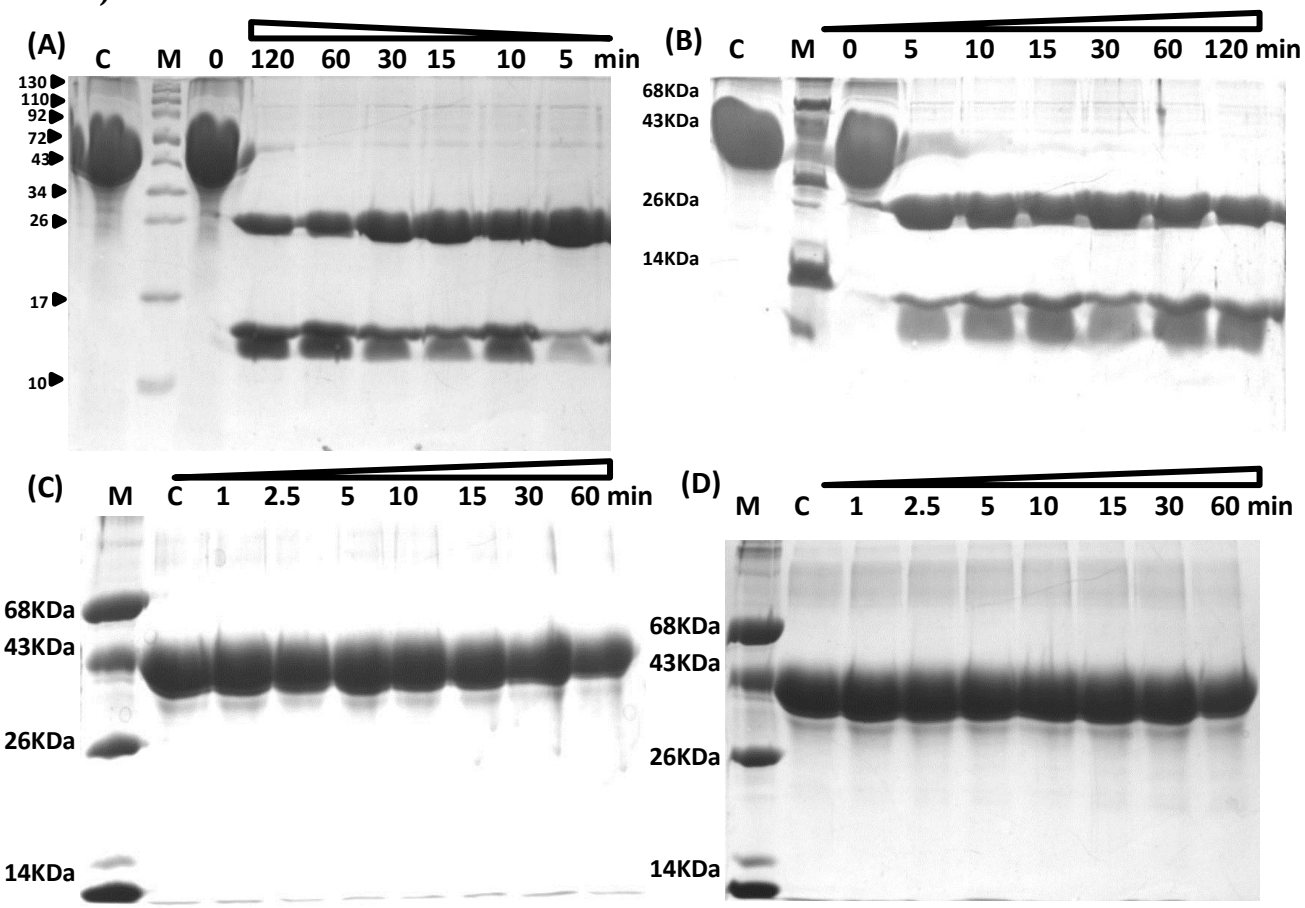


Figure 3.3.4:- **(A, B)** Limited proteolysis and **(C, D)** Chemical crosslinking of BARD1 ARD-BRCT *wild-type* and Gln564His respectively.

Comparative secondary structure characterization of BARD1 ARD-BRCT *wild-type* and Gln564His mutant protein using CD spectroscopy show similar spectra having a mixed profile of α-helices, β-sheets, turns, and random coil component in the structure.(**Figure:-3.3.5)**.

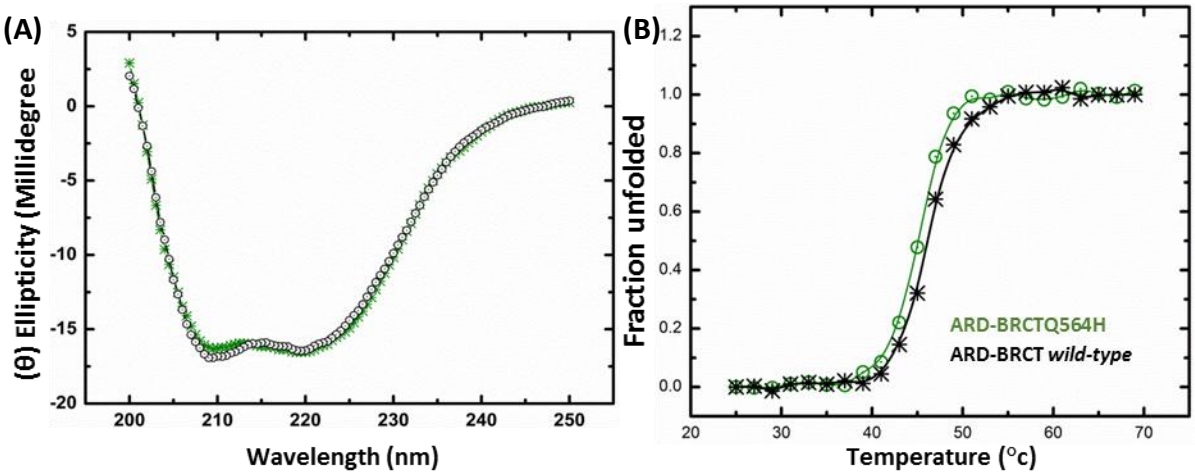


Figure:-3.3.5- **(A)** CD spectra and **(B)** thermal denaturation of BARD1 ARD-BRCT *wild-type* and Gln564His mutant protein respectively.

Thermal denaturation of *wild-type* and mutant proteins was performed using circular dichroism spectroscopy and fraction unfolded for *wild-type* and mutant proteins were calculated as function of temperature. It has been observed that *wild-type* and mutant proteins unfold via a two state pathway and both the proteins unfold completely at 55˚C. The T_m calculated for *wild-type* and mutant were 46.2±.25°C and 45.6°C±.45°C respectively. Insignificant change in the T_m indicates that the mutation does not affect the thermodynamic and unfolding pattern of the *wild-type* protein **(Figure: - 3.3.5 B)**. Chemical denaturation of *wild-type* and mutant proteins with guanidium hydrochloride (GuHcl) indicates that *wild-type* and mutant protein unfolds via intermediate formation. This observation is also supported by

blue-shift of the emission maximum in the fluorescence spectra consistent with the presence of an intermediate molten globule **(Figure: - 3.3.6)**.

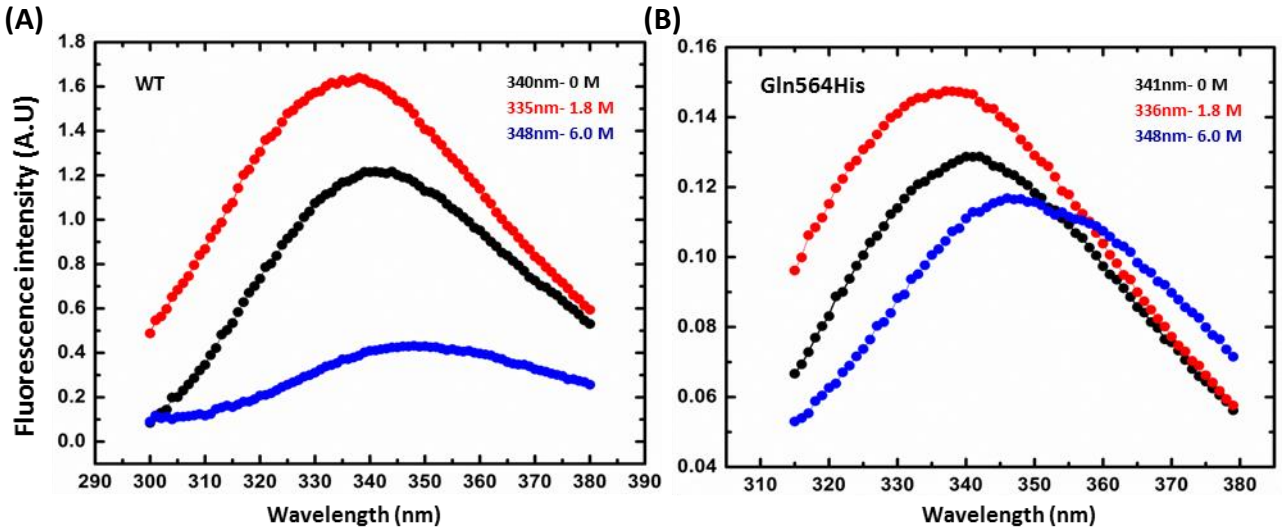


Figure 3.3.6:- Chemical denaturation of **(A)** BARD1 ARD-BRCT *wild-type* (WT) and **(B)** Gln564His mutant.

The emission maxima shifted from (λ_{max}) shifted from 341nm to 335nm in *wild-type* and from 341nm to 336nm for mutant at 1.8M GuHCl. However, the λ_{max} at 6M concentration of GuHCl was observed to be close to 350nm, thus signifying that proteins unfold via molten globule formation.

CstF50 was expressed in bacterial system using pGEX-kT and pET-28a vector **(Figure: - 3.3.6).** However, CstF50 (1-431aa) and WD40 domain (92-431aa) shows that most of the protein is insoluble in nature and forms inclusion bodies **(Figure: - 3.3.7).** CstF50 full length was cloned in pET-28a vector **(Figure: - 3.3.7)** and purified as 6HIS-CstF50 from inclusion bodies by solubilizing in 8M urea (**Figure: - 3.3.8)** and further refolded in native condition. However, no good quality CD spectra were obtained due to presence of arginine in the

refolding buffer, this interfered with the measurement. Efforts to remove arginine hydrochloride led to precipitation of the refolded protein.

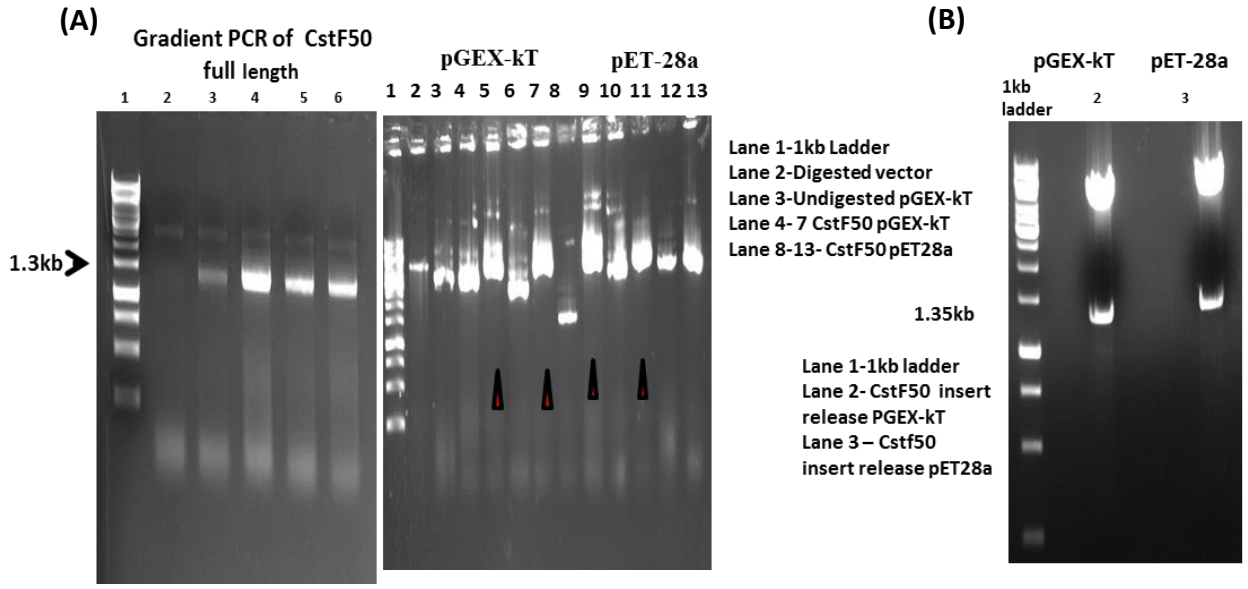


Figure 3.3.7:- **(A)** Cloning of CstF50 (**B)** pGEX-KT and pET-28a vectors

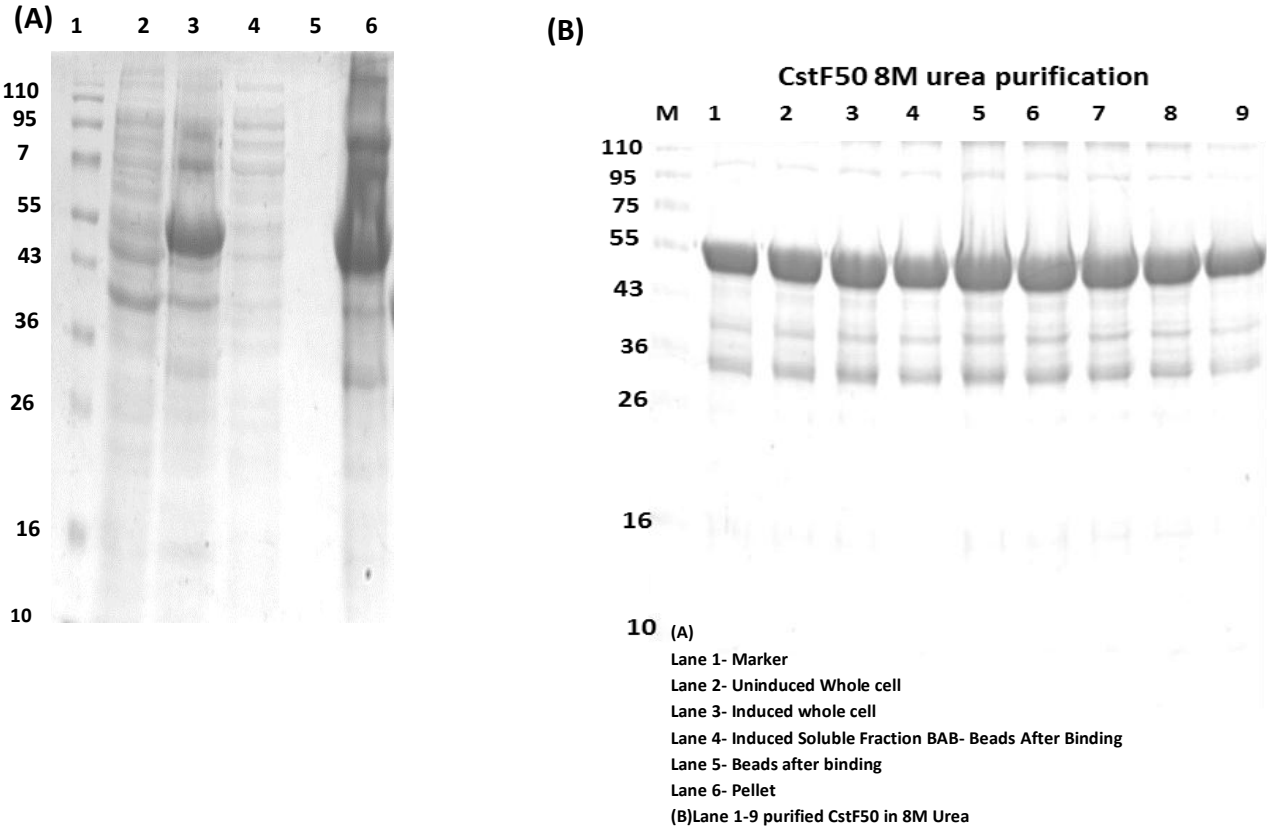


Figure 3.3.8:- Expression and purification of CstF50.

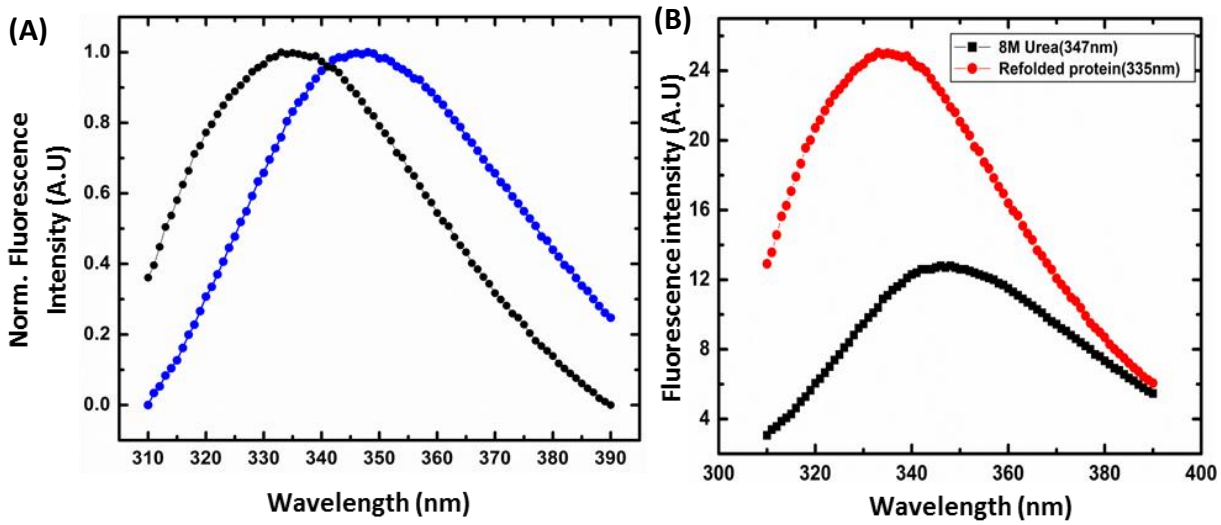


Figure 3.3.9: - Refolding profile of CstF50

To further confirm CstF50 protein folding in native condition, we have performed fluorescence spectroscopy, and the unfolded protein shows loss of fluorescence intensity and red-shift in the emission maximum (λ= 348nm) **(Figure: - 3.3.9)**. Fluorescence spectra of refolded protein shows an increase in the fluorescence intensity and blue-shift (λ= 335nm) of the emission maximum which indicates that tryptophan residues are buried in the hydrophobic core of the protein **(Figure: - 3.3.9)**.

3.3.3 Structural insight into the binding interface of BARD1 ARD-BRCT and CstF50:- BARD1 ARD-BRCT region (425-777aa) was modeled using Robetta server [37] and mutation Gln564His was incorporated using the swiss PDB viewer [40]. Furthermore, CstF50 was also modeled using Robetta server. Both the models were validated by saves (Metaserver for analyzing and validating protein structures (https://services.mbi.ucla.edu/SAVES/) [243] and Molprobity server [244, 245]. Further model was refined by Modloop server [246] and based on stereo chemistry and Ramachandran plot, best model was selected for structural analysis [247] **(Figure:- 3.3.10)**.

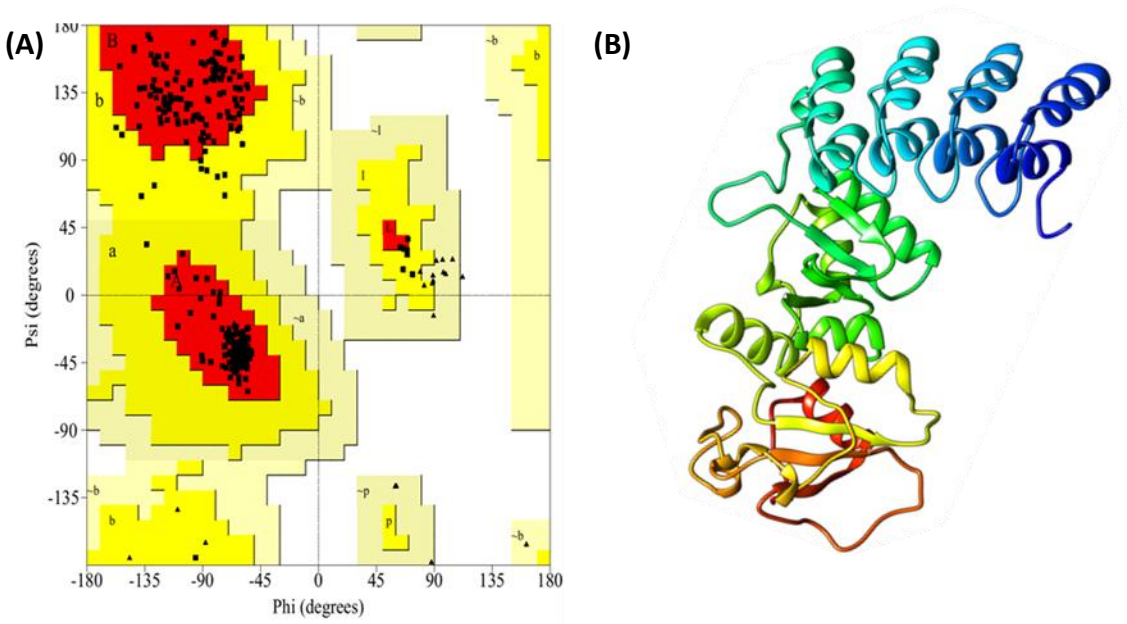


Figure 3.3.10: - **(A)** Ramachandran plot and **(B)** structural details of validated model for BARD1 ARD-BRCT (425-777aa).

Table 3.1:- Stereochemistry details BARD1 ARD-BRCT *wild-type*

Poor rotamers	2	0.64%	Goal: <1%
Ramachandran outliers	0	0.00%	Goal: <0.05%
Ramachandran favored	344	98.01%	Goal: >98%
Cβ deviations >0.25Å	1	0.30%	Goal: 0
Bad backbone bonds:	0 / 1411	0.00%	Goal: 0%
Bad backbone angles:	1 / 1762	0.06%	Goal: <0.1%

In BARD1 ARD-BRCT structure (425–777aa), ARD domain is connected to BARD1BRCT domain through a flexible linker of 14 amino acids in a head-to-tail manner. The three-dimensional model structure of ARD domain is predominantly α-helical in nature and BRCT domain shows a combination of two repeats consisting of three α–helices and β-strand connected by linkers.

Modeling BARD1 ARD-BRCT region also serves the purpose to study the extent of domain motion alone and also in complex with the CstF50. Errat evaluation [248] submits that model has an average overall quality factor of 92.151, and may have good structural resolution. 3D-1D analysis [249] indicates that 82.77% of the residues had an averaged 3D-1D score > 0.2 and never dipped below zero. MolProbity stereochemical parameters were also found satisfactory for the BARD1 ARD-BRCT (425–777aa) model and indicate that model structure parameters are in normal range **(Table 3.1)**.

However, CstF50 model is mainly α–helical at the N-terminal region and is connected to WD40 repeat region with large flexible loop. The C-terminal 7 WD40 repeats domain comprised of ~300 amino acids. WD40 repeat domain is an evolutionarily conserved domain, and it is composed of 5 to 7 repeats majorly consisting of β-strand acquiring a β propeller shape **(Figure: - 3.3.11)**. In the structural model, majority of CstF50 residues (90%) are in the most favored regions of Ramachandran plot indicating that the corresponding coordinates for the folded amino acids are satisfactory.

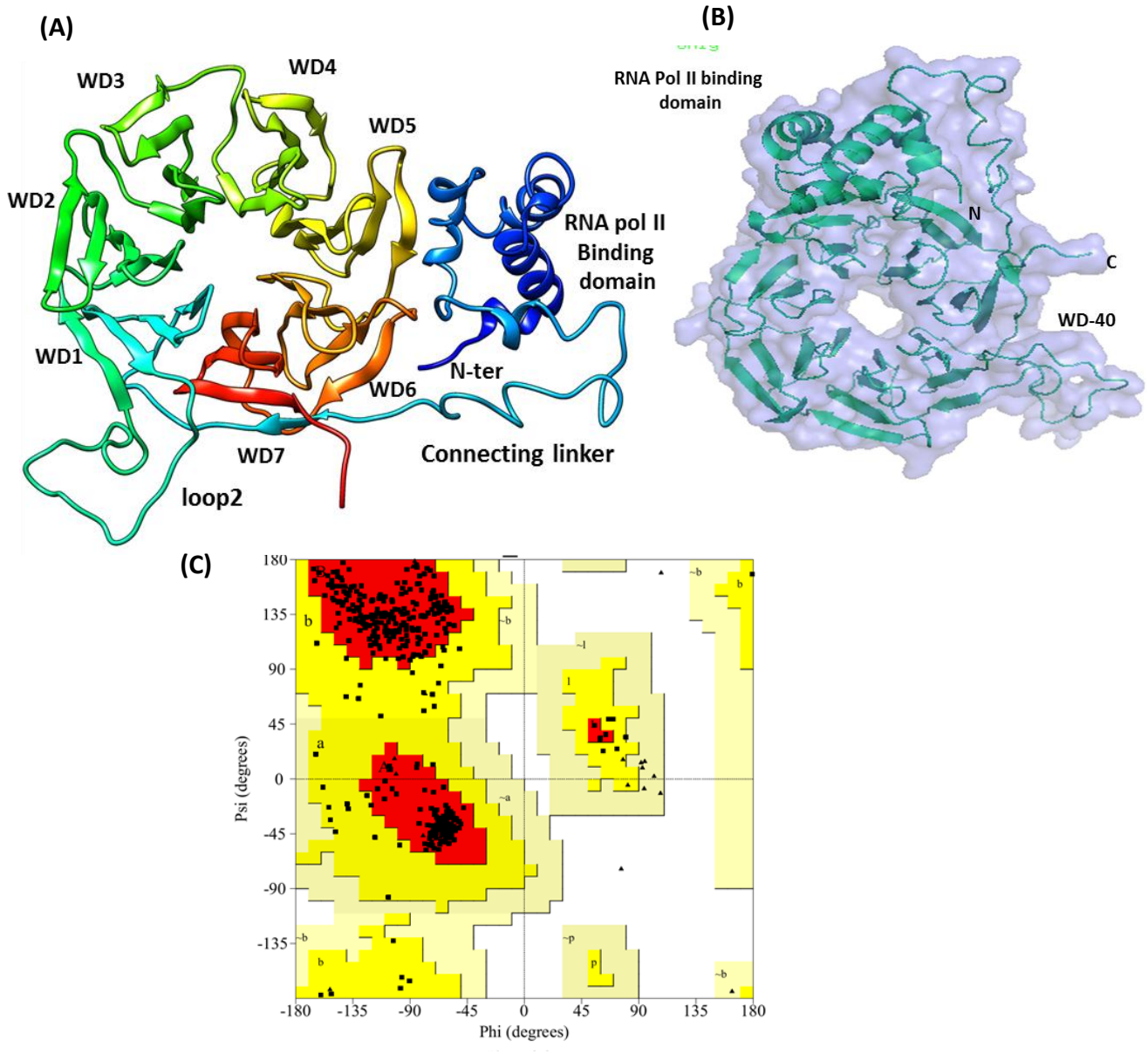


Figure: - 3.3.11:- **(A)** Model structure of CstF50, **(B)** Structural details of validated model for CstF50 and **(C)**. Ramachandran plot of cstF50.

The Verify 3D-1D profile analysis of the predicted protein model of CstF50 revealed that the average score stayed above 0.2 and never decreased below zero.

Table 3.2:- Stereochemistry details of CstF50 (1-431aa)

Poor rotamers	3	0.79%	Goal: <1%
Ramachandran outliers	0	0.00%	Goal: <0.05%
Ramachandran favored	400	93.24%	Goal: >98%
Cβ deviations >0.25Å	0	0%	Goal: 0
Bad backbone bonds:	0 / 1723	0.00%	Goal: 0%
Bad backbone angles:	2 / 2152	0.09%	Goal: <0.1%

The overall structure calculation from PROVE program [250] gave Z-score RMS of 1.53. The Z-score obtained from SAVES [251] is significant as the Z-score is not above 4.0 and below-4.0, which indicates that the hydrophobic amino acids buried inside the core are well packed. A good average overall quality factor (Errat score) is around 87% for the predicted model which is satisfactory [248]. Furthermore, for evaluation of the structural geometry of the modeled CstF50 protein, we have submitted the structure to the Molprobity server [244, 245]. As the calculated structural parameters of the CstF50 model were within the allowed range (Table 2). Therefore, we have used this model for docking and molecular dynamics simulation studies. Similarly, being a member of WD40 family of proteins, CstF50 contains seven WD40 repeat domains spanning C-terminal region (92-431aa) comprising of β-sheets. N-terminal domain (1-91aa) is homodimerization domain, which is required for interactions with RNA pol-II CTD, is predominantly comprised of the α-helices. Inter-secondary structure segments/loops were identified by DSSP and STRIDE [252, 253]. These interactions can then

form an anchor for assembly of larger protein complexes. Such interactions were first structurally characterized for heterotrimeric G-proteins where the β-subunit is a WD40 protein [59, 60].

3.3.4 Domain flexibility of ARD with respect to BRCT domain:-

In-silico Normal Mode Analysis (**NMA)** is used to study the protein dynamics of BARD1 ARD-BRCT domain and mutant, which is precisely well suited for demonstrating inter-domain motions, embedded in the structure.

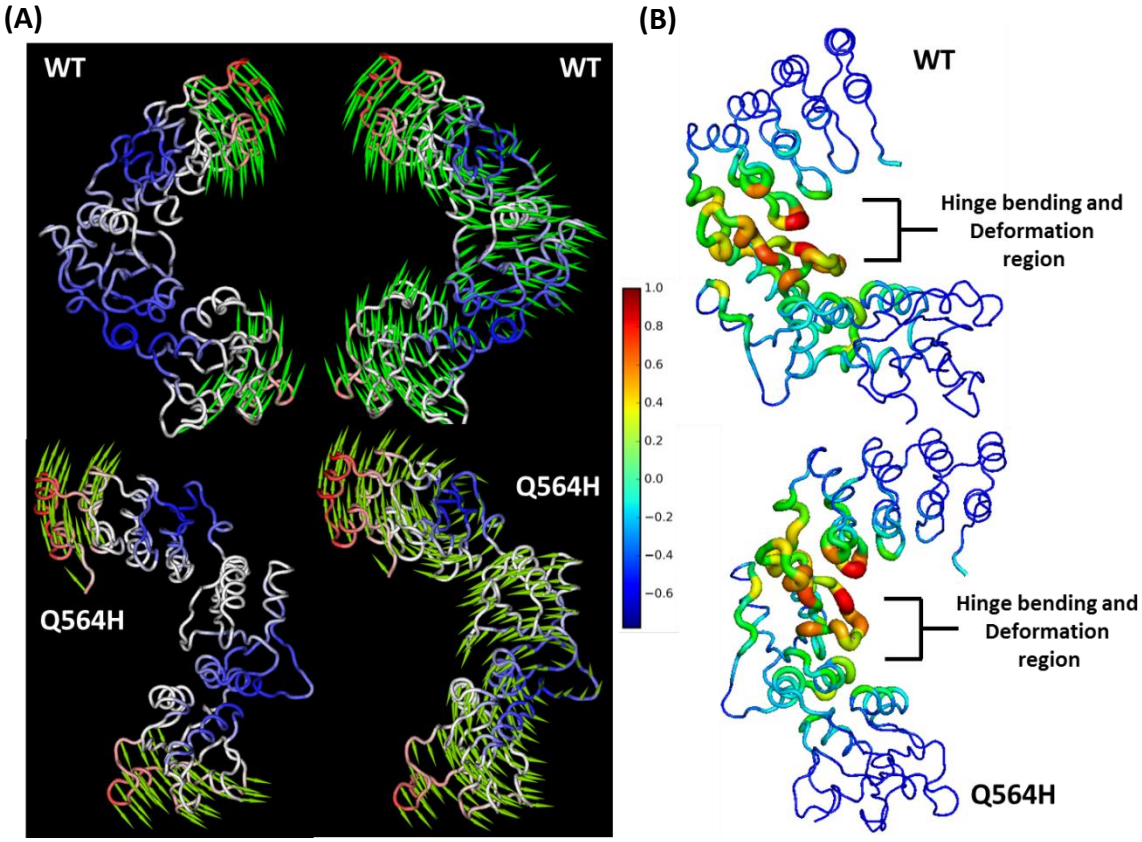


Figure 3.3.12:- Comparative NMA first mode profile of BARD1 ARD-BRCT **(A)** *wild-type* and Gln564His mutant **(B)** Comparative Deformation analysis of *wild-type* and mutant showing hinge bending.

To study the extent of domain motion in BARD1 ARD-BRCT, we have performed NMA for BARD1 ARD-BRCT *wild-type* and mutant protein structure. First few modes are the low frequency modes and represent the large motions within the protein structure **(Figure: - 3.3.12 A)**. NMA of BARD1 ARD-BRCT *wild-type* and Gln564His indicates large domain motion between the ARD and BRCT domain, and the nature of concerted motion indicates that residues in the ARD domain and BRCT domain are positively co-related while linker region shows anti-correlated motion to the ARD1 and BRCT domain respectively. First five modes show that ARD and BRCT domains fluctuate between open and closed states because of bending of hinge at the linker region. Deformation analysis indicates that linker region of *wild-type* and mutant protein has high value of deformation as the RMSF of the linker region is recorded to be the highest.

Large-scale conformational transitions in proteins can occur because of relative movement between domains. Flexibility in short segments of the protein backbone allows corresponding domain motions to occur, with only minor structural perturbations within domains [254, 255]. The collective motion of a simulated protein can be identified by analyzing both the covariance and time correlation in the positional fluctuation of its atoms [256]. It has been observed that full-length protein always has a rigid conformation than its individual domains. Here, MD simulation has been applied to investigate the conformational changes induced in the BARD1 ARD-BRCT domain. Comparative analysis of the RMSD data was performed after fitting the trajectory input of *wild-type* and mutant protein structures **(Figure: - 3.3.13)**.

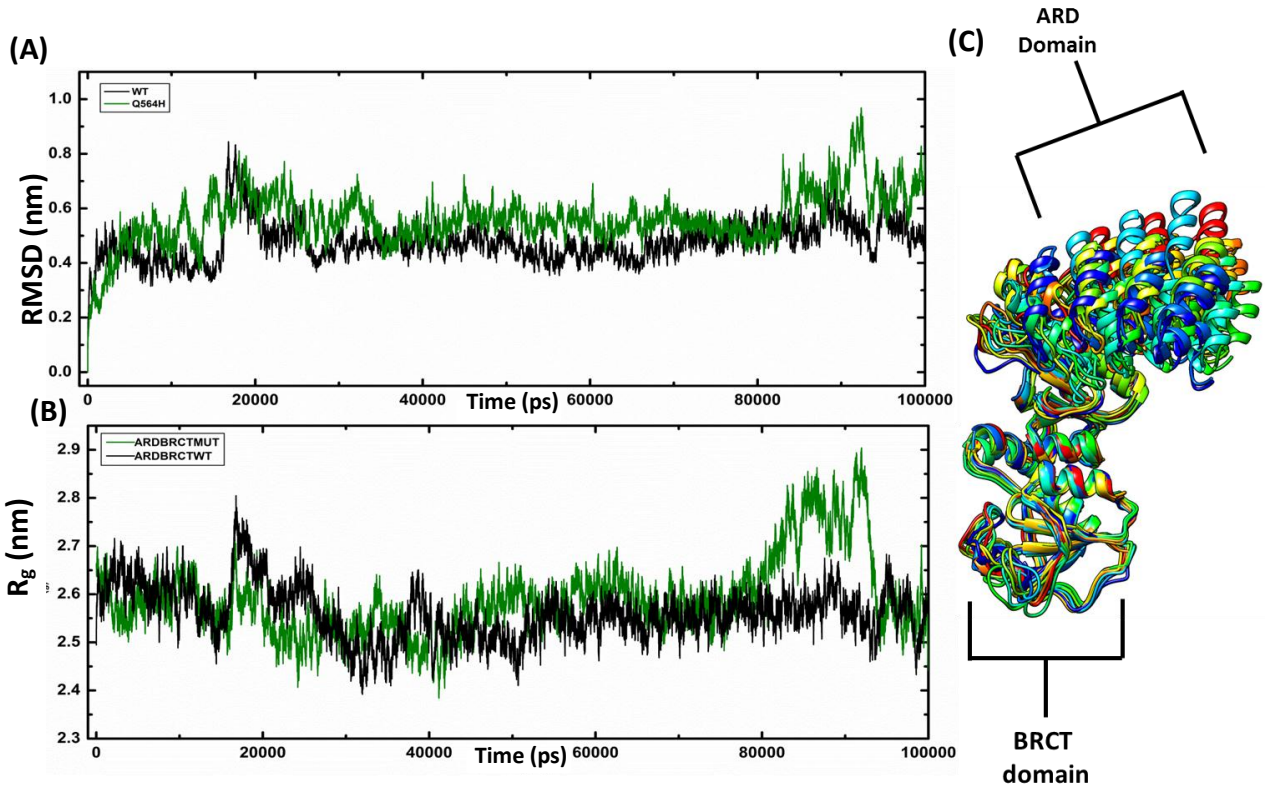


Figure 3.3.13: - Comparative **(A)** RMSD and **(B)** Rg profile of BARD1 ARD-BRCT (Black) and Gln564His (green) mutant and **(C)** Structures sampled at every 10ns for *wild-type*.

The simulation for 100ns reveals that BARD1 ARD-BRCT *wild-type* structure is comparatively more stable than the mutant. After 5ns of simulation, RMSD of both mutant and *wild type* protein increased up to 5Å but after 10ns, the *wild-type* RMSD decreases to 3.5Å, whereas mutant RMSD increases to 6.5Å. Sharp increase in the RMSD has been observed at 16ns in *wild-type* structure. Nevertheless, looking at the mutant RMSD profile, instead of sharp increase gradual change in RMSD at 18ns was observed. Sharp increase in the RMSD for mutant structure at 95ns was observed in contrast to the *wild-type* protein.

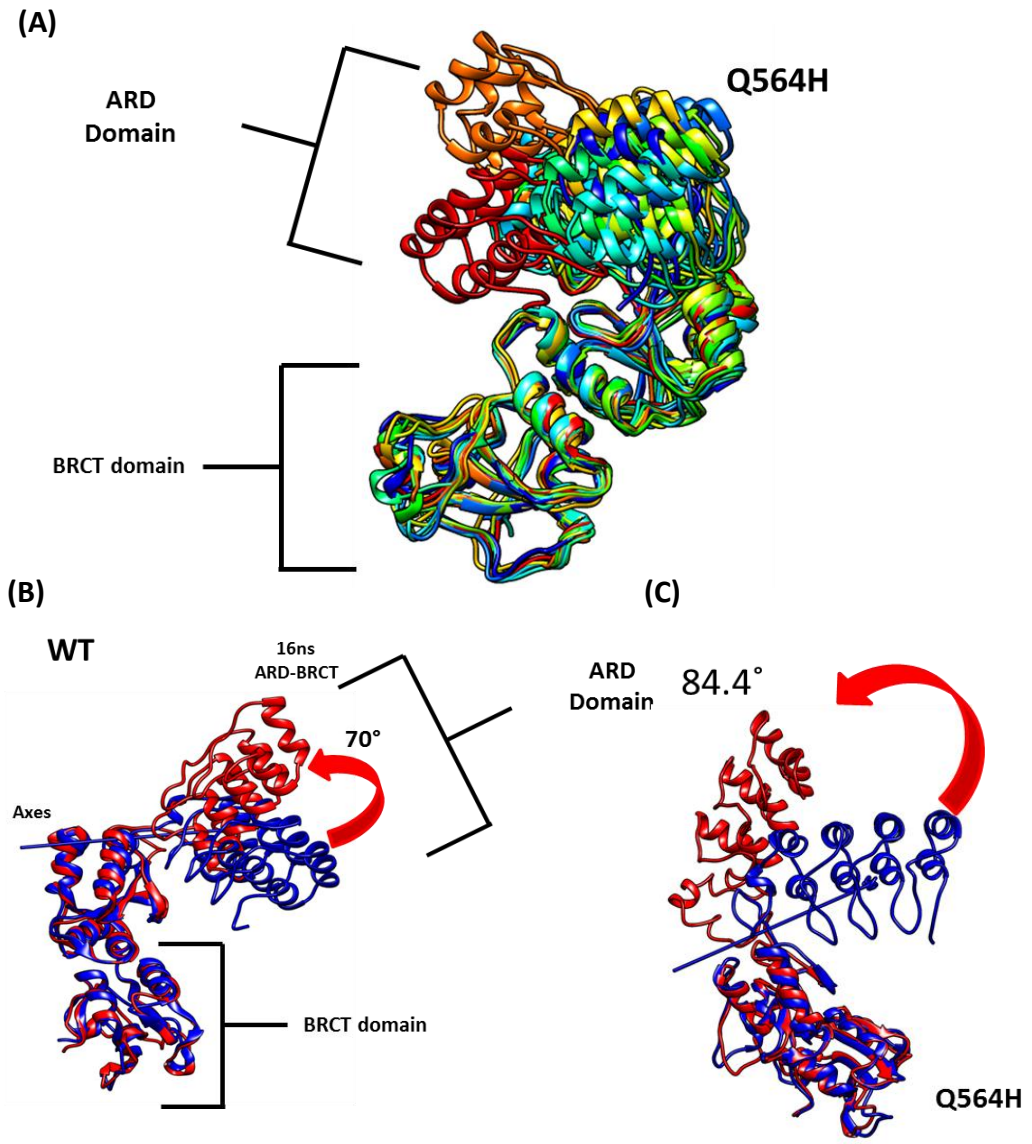


Figure 3.3.14: - **(A)** Comparative aligned structure sampled at every 10ns for mutant Gln564His, **(B)** highest RMSD and input structure aligned for *wild-type* and **(C)** mutant (Gln564His).

We have analyzed the highest RMSD extracted structure for *wild-type* and mutant protein from the trajectory. After comparing the structures of *wild-type* at 0ns and 16ns (highest RMSD frame) and for mutant 0ns and 98ns, a large shift in the ARD domain orientation was observed.

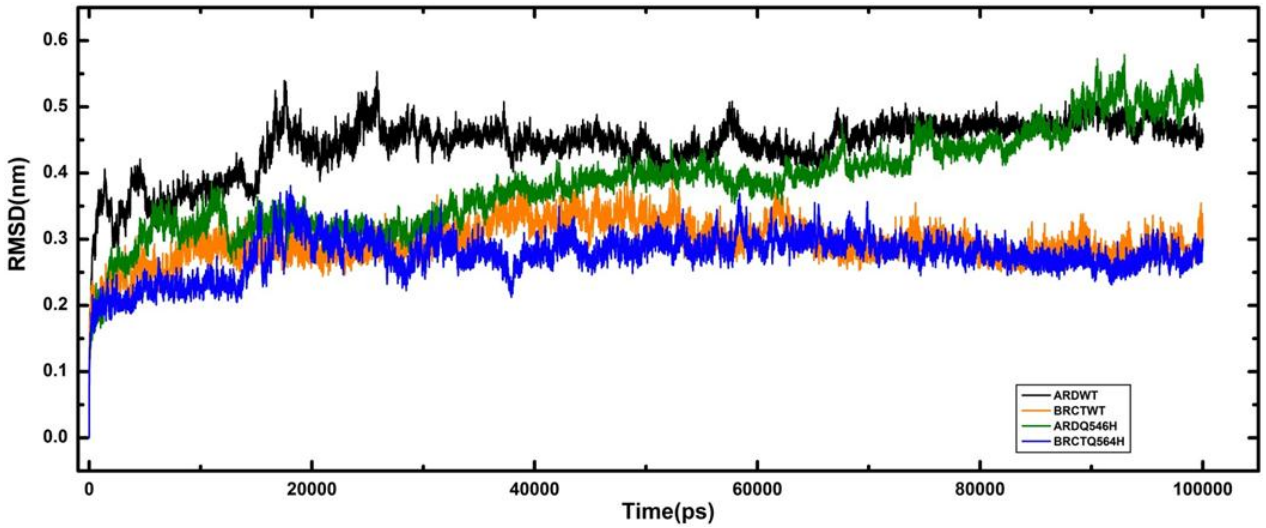


Figure 3.3.15: - Comparative RMSD representation of ARD and BRCT domain. ARD *wild-type* (Black), mutant (Green) and BRCT domain *wild-type* (yellow), mutant (Blue) respectively.

Structural alignment with input structure and highest RMSD attained structure for *wild-type* and mutant protein shows large rotation about 70˚ (RMSD=8.3 Å) in the ARD domain from the axis for *wild-type*, and 84.4˚ (RMSD=9.27 Å) for mutant protein respectively. This indicates that BRCT domain of BARD1 ARD-BRCT *wild-type* and mutant protein structures are comparatively rigid but hinge bending at the linker leads to the domain motion **(Figure:- 3.3.14, 3.3.15**).

Furthermore, domain wise comparative study of RMSD for *wild-type* and mutant protein structures shows that ARD domain is more flexible than the BRCT domain **(Figure:- 3.3.15)**. Decrease in the RMSD in both *wild-type* and mutant up to 4Å is indicative of stabilization in structural change, but the RMSD of mutant remained high as compared to *wild-type* protein. After 35ns *wild-type* structure is completely stabilized, as there was only small fluctuation in the RMSD in the *wild-type* structure. On the other hand, BARD1 ARD-

BRCT Gln564His mutant structure shows a sudden increase in the RMSD profile at 90ns, indicating an another change in the conformation leading to high average RMSD in mutant structure than the *wild-type* structure.

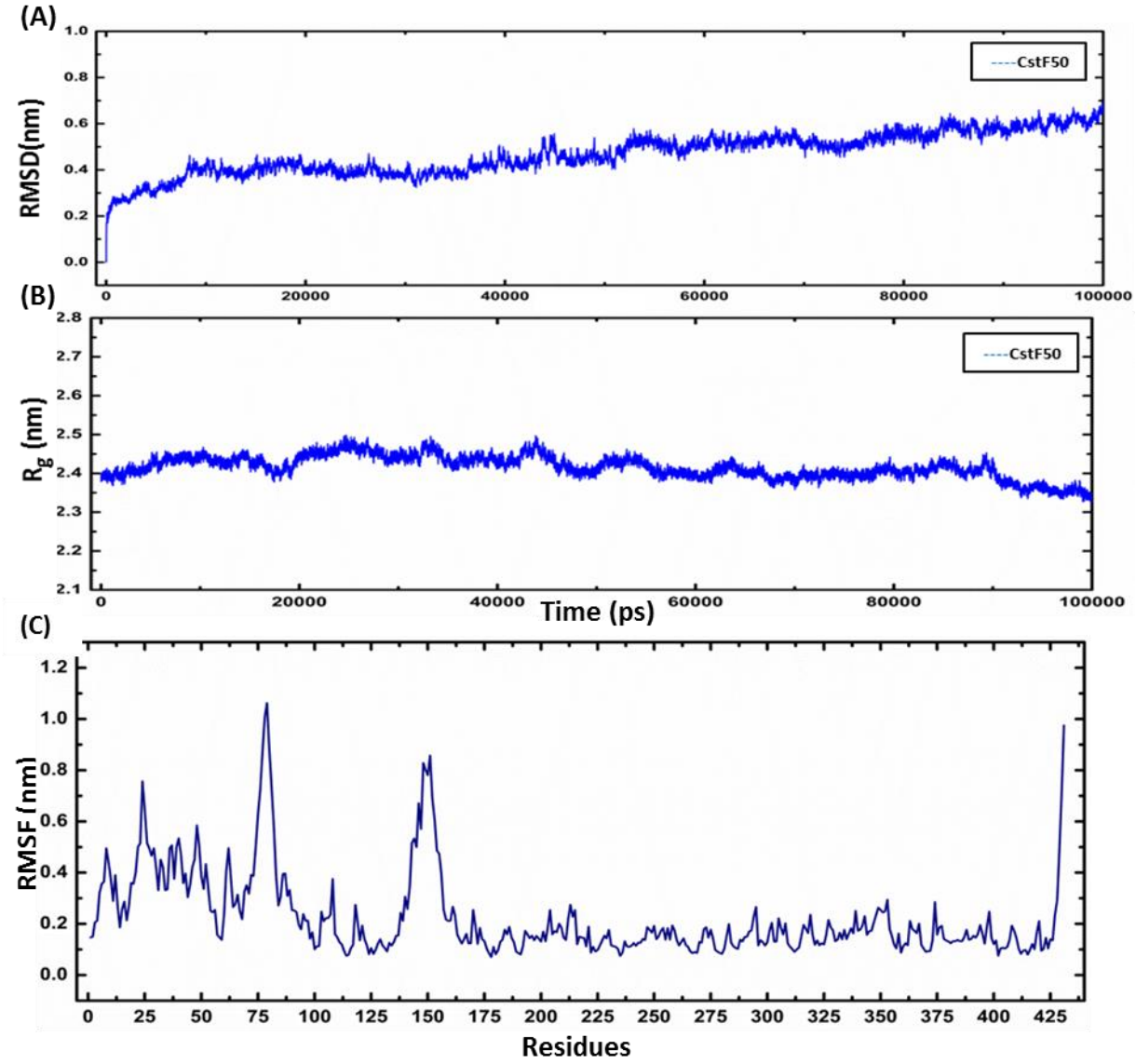


Figure 3.3.16:- **(A)** RMSD, **(B)** R_g, and **(C)** RMSF attained by CstF50 during 100ns MD production.

Moreover, CstF50 structure did not show any large RMSD, R_g deviation during 100ns simulation and residual RMSF fluctuations that contributed either from the linkers between

N-ter and WD40 domain or loops between beta strands of 1st WD40 repeat **(Figure:- 3.3.16).** The comparative radius of gyration for BARD1 ARD-BRCT, Gln564His mutant structure was also analyzed by molecular dynamics simulation.

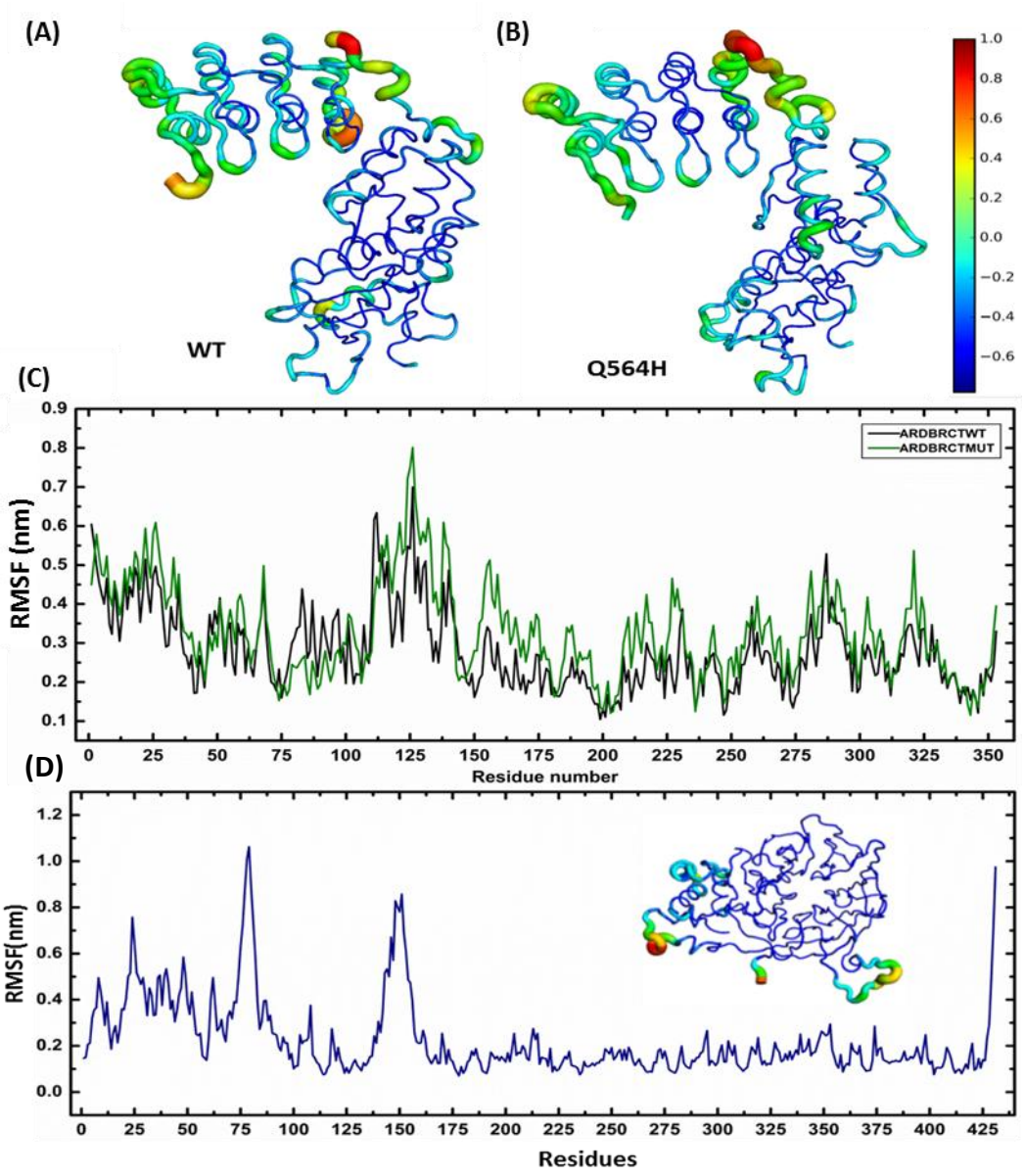


Figure 3.3.17: - (A and B) RMSF Structure representing for BARD1 ARD-BRCT *wild-type* and mutant., (**C)** RMSF for *wild-type* (green) and mutant(black) **(D)** RMSF structure and profile of CstF50 .

Change in radius of gyration during simulation indicates that the structural changes are time dependent and determines the protein structure compactness. Comparative evaluation of radius of gyration throughout the simulation of wild-*type* and mutant during the early phase of simulation doesn't show any sharp changes in the structural compactness. After 16ns, there is steep rise in the R_{gyr} in *wild-type* which is a large structural juxtapositioning of the ARD domain **(Figure: - 3.3.13B)**. However, in mutant structure, there is sharp increase in the R_{gyr} at 98ns. This may be due to bulky R-group attached to glutamine.

3.3.5 Principal Component Analysis:-

PCA was performed over the 100ns trajectory of BARD1 ARD-BRCT *wild-type*, Gln564His mutant and CstF50 to analyze the dynamics of proteins in the essential subspace. First three eigenvector was deduced by diagonalizing the covariance matrix and further projected on each other. Projection of Eigenvector 3 on 1 and 2 on 1 indicates large tertiary structural fluctuation in the conformational spaces which indicates that protein is highly dynamic. Furthermore projection of eigenvector 2 on 1 indicates that *wild-type* and mutant shows large tertiary structural changes during the simulation **(Figure: - 3.3.18 A, B)**.

Similarly, CstF50 also shows conformational fluctuations when projected for eigenvector 3 on 1 and 2 on 1. Observed large tertiary structural changes may be because of the two large interconnecting loops between N-terminal domain and WD40 domain and second in loop between 1^{st} β-strands of WD40 repeat **(Figure:- 3.3.18 C)**. The Cross-correlation for PCA of *wild-type* and mutant indicates that differential residual correlation in both the domains. Compared to mutant *wild-type* shows more positive correlation within the residues of ARD domain and BRCT domain but mutant show comparatively reduced positive correlation in both the domains **(Figure:- 3.3.19A, B)**.

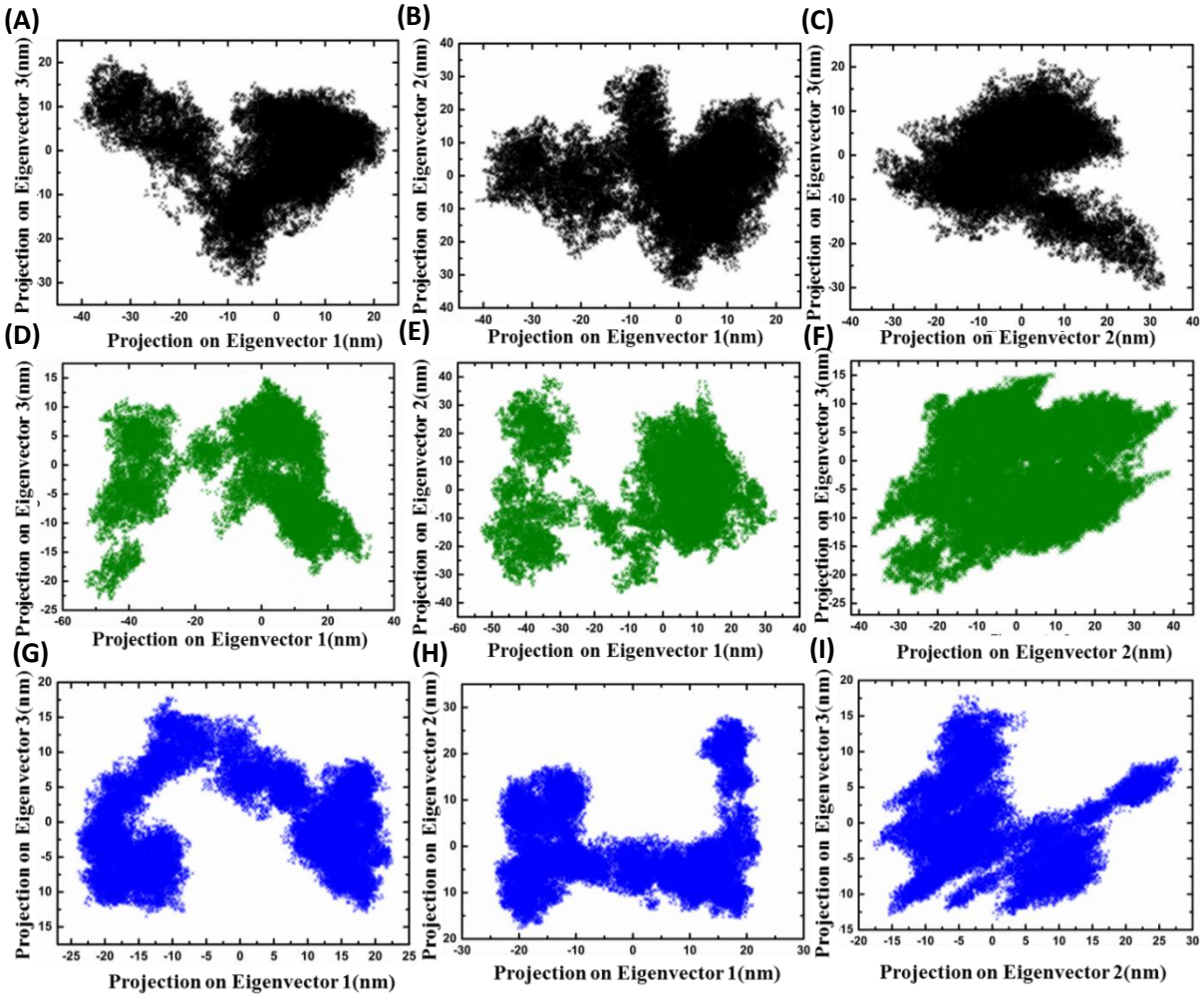


Figure 3.3.18: - Projection of first three eigenvectors for the structure **(A, B, C)** BARD1 ARD-BRCT *wild-type*, **(D, E, F)** mutant and **(G, H, I)** CstF50.

CstF50 structure shows high value of positive correlation in WD40 repeat domain and comparatively low positively correlated motion within the N-ter domain **(Figure:- 3.3.19C)**. To understand comparative residual displacements in the subspaces spanned by the first two eigenvectors were performed to understand the fluctuation at the residual level **(Figure:- 3.3.18)**.

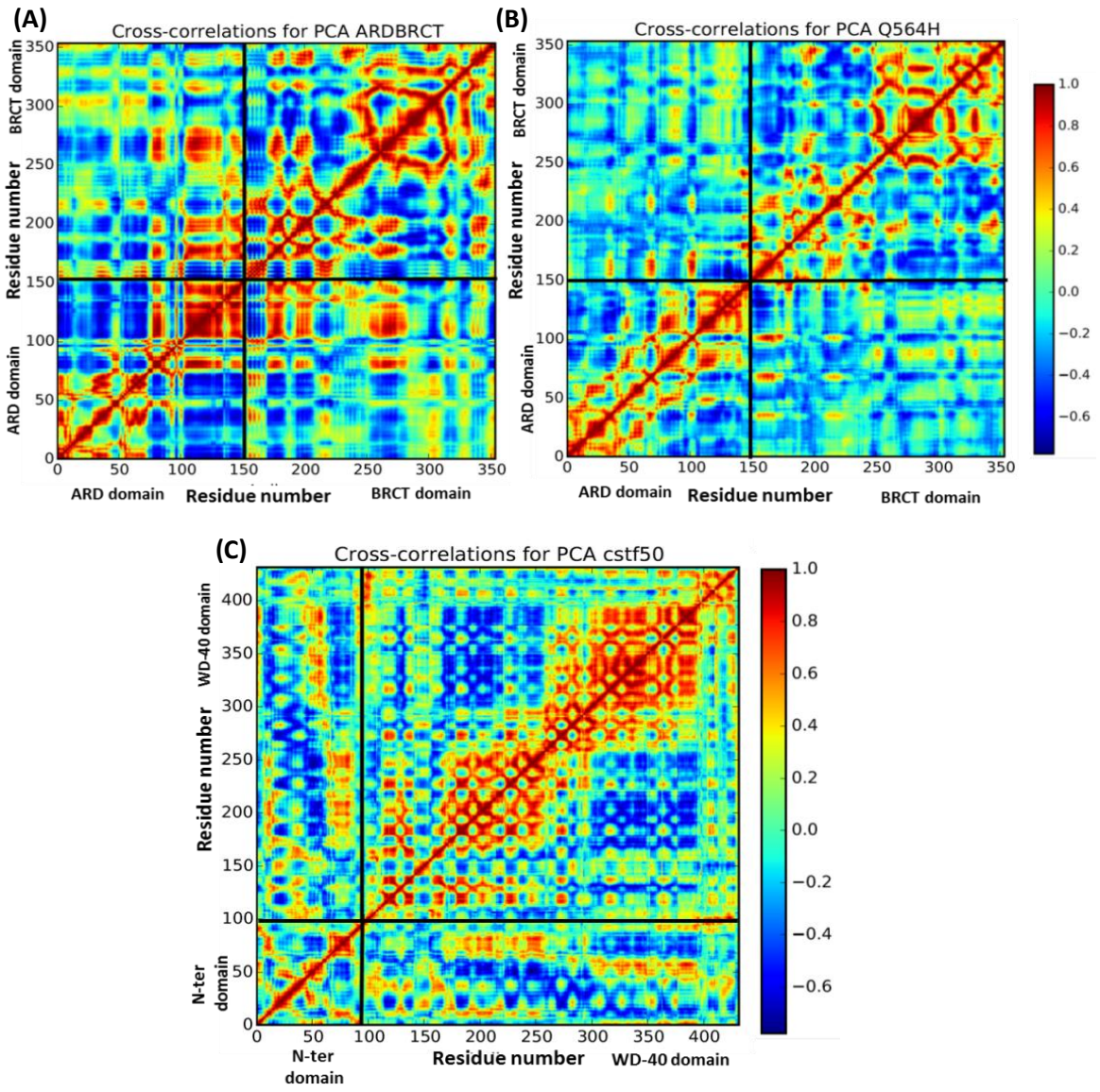


Figure 3.3.19: - Cross-correlation for the structure **(A)** BARD1 ARD-BRCT *wild-type* **(A)**, **(B)** mutant and **(C)** CstF50.

Projection of eigenvector 1 on residue of *wild-type,* mutant protein N-terminal, and linker region shows comparatively high fluctuations as compared to the BRCT domain **(Figure:-3.3.20)**. Similarly, projection of eigenvector 2 on residues shows a similar pattern of fluctuation in second eigenvector values, which shows the concerted character of residual

fluctuation in the ARD domain and linker region of the *wild-type* and mutant protein, substantiating results obtained from the NMA and RMSF from MD simulations **(Figure:- 3.3.20 and 3.3.17)**.

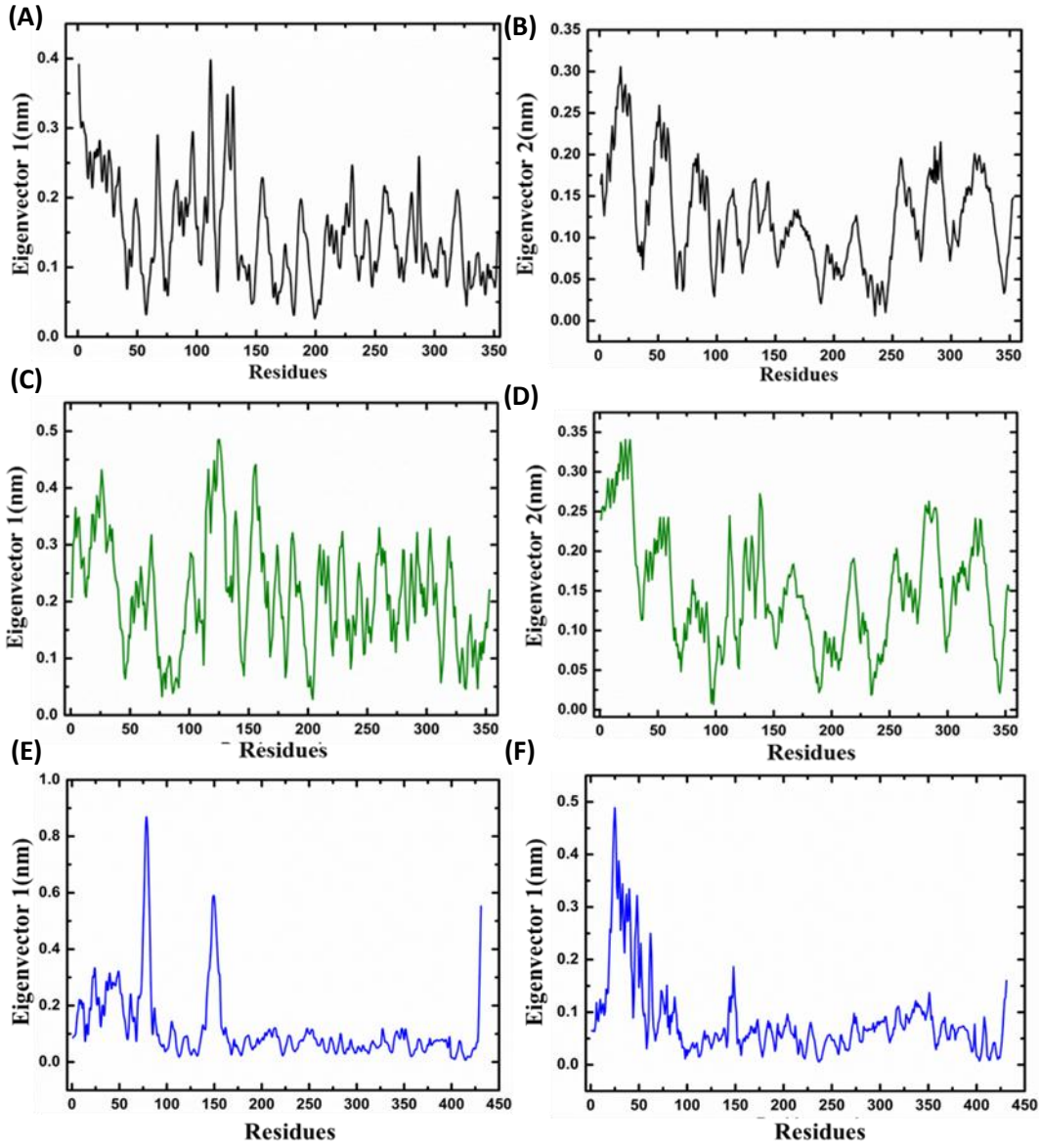


Figure 3.3.20: - Residual fluctuation against first two eigenvectors for the structure **(A, B)** BARD1 ARD-BRCT *wild-type*, **(C, D)** mutant and **(E, F)** CstF50.

Moreover, CstF50 shows large fluctuation in eigenvector 1 against the residual region related to the N-terminal connecting linker and 1st WD40 interconnecting loop. Eigenvector 2 projection on residues indicates flexible N-terminal domain may be due to the flexible linker and stable WD40 domain **(Figure: - 3.3.20 E and F)**.

3.3.6 Binding mechanism of BARD1 ARD-BRCT and CstF50:-

We have used HADDOCK server to understand the binding interactions between CstF50 and BARD1. It is well established that 7th WD40 repeat domain from 395-431 amino acids of CstF50 is required to establish the functional complex with BARD1 [6]. PDBsum software was used to analyze the interactions between BARD1 ARD-BRCT and CstF50 complex [257].

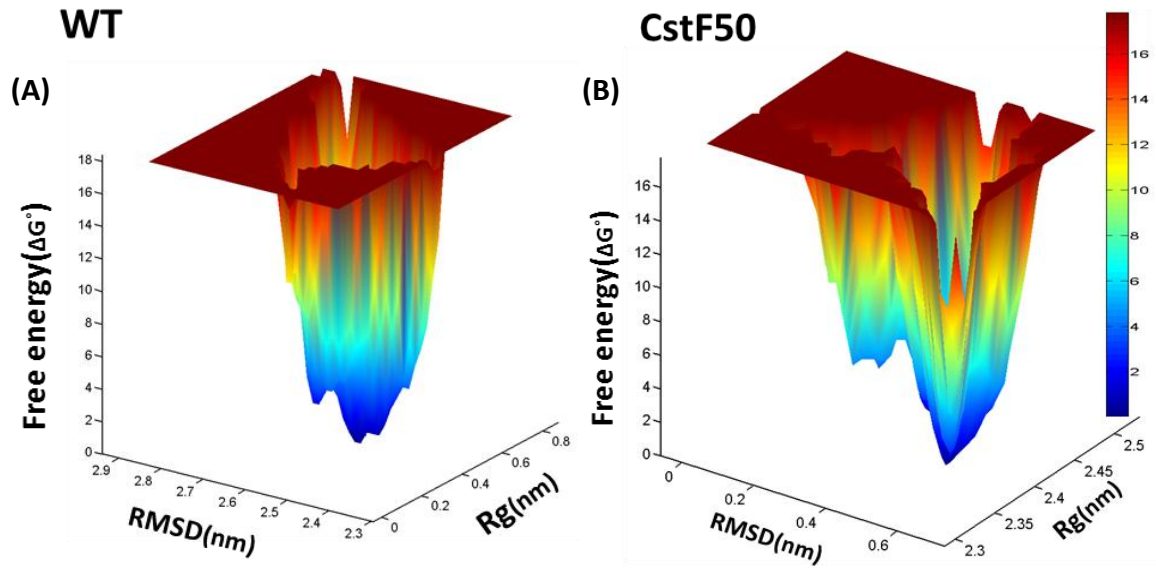


Figure 3.3.21: - Free-energy calculation for the structure **(A)** BARD1 ARD-BRCT *wild-type* and **(B)** CstF50.

As Gln564His mutation leads to loss in CstF50 binding to BARD1 [258]. Best docked complex was selected, and PDB sum was used to understand the residues involved in the

binding interface. Proteins were simulated for 100ns and minimum energy structure was selected for docking studies **(Figure: - 3.3.21)**.

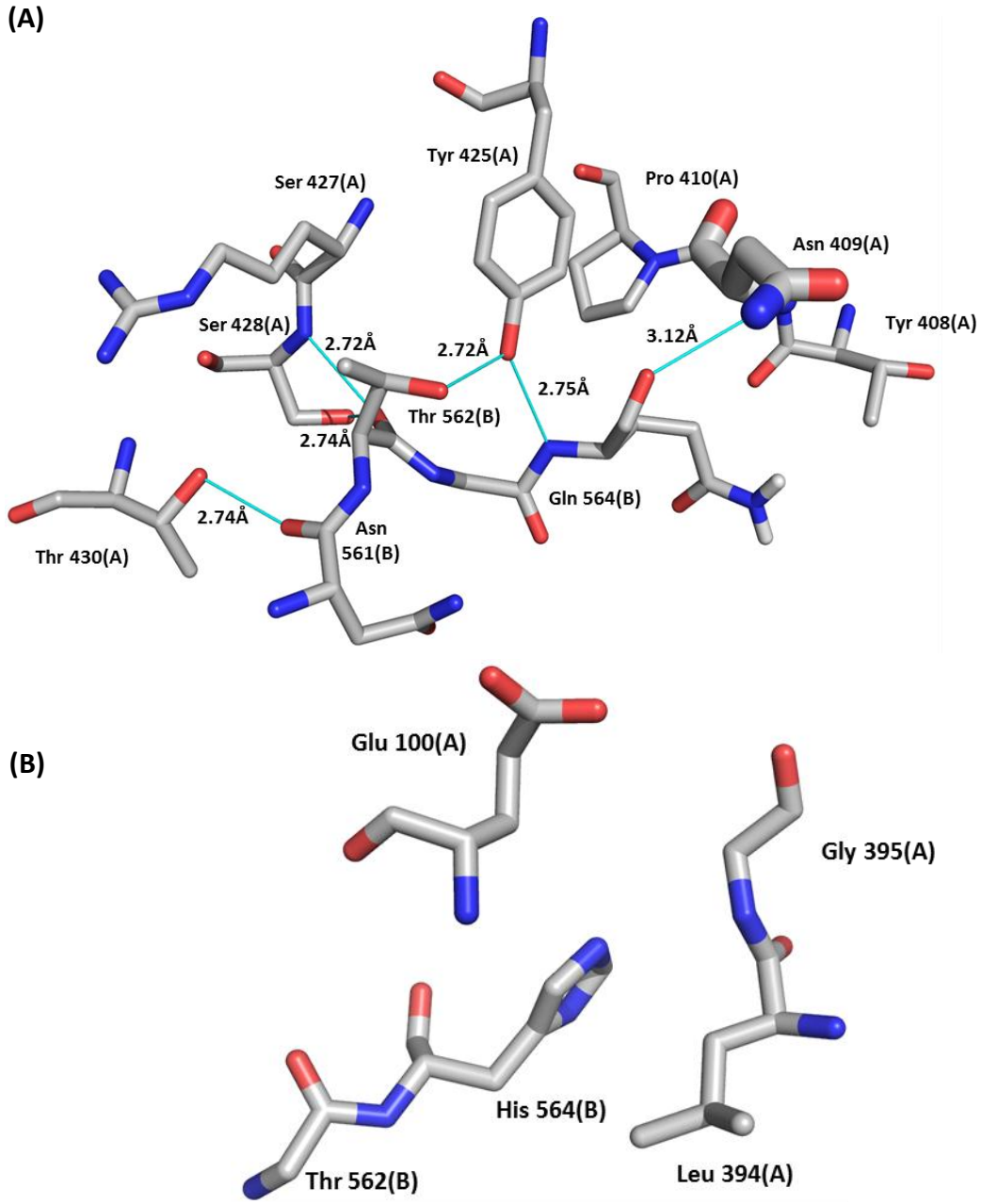


Figure 3.3.22: - Binding interface of **(A)** BARD1 ARD-BRCT *wild-type* and CstF50 complex, **(B)** BARD1Gln564His and cstF50 complex.

Docking was performed between minimum energy structure of BARD1 ARD-BRCT, mutant and CstF50 **(Figure: - 3.3.22)**. Similar, docking parameters were used for BARD1 ARD-BRCT mutant and CstF50 docking, and results were analyzed using PDB sum. BARD1 Gln at 564 position forms hydrogen bond with Asn409, Tyr425 and participates in non-hydrogen bonding interaction with Asn409, Tyr425, Pro410, and Thr 408. While BARD1 Thr 562 facilitates other major interactions between BARD1 and CstF50, and forms hydrogen bonding interaction with Tyr425 and non-hydrogen bonding interaction with Arg427, Ser428. Thr430 of CstF50 supports the interaction by forming hydrogen and non-hydrogen bonding interactions with linker residues **(Figure: - 3.3.22 A)**.

CstF50 Thr 430 forms non–hydrogen bonding interactions with BARD1 linker residues like Asn561. From the mutant's structure, it has been observed that BARD1 His564 has lost all hydrogen and non-hydrogen bonding interactions with the CstF50, 7^{th} WD40 repeat region residues and form non hydrogen bonding interaction with CstF50 Glu100, Asp394 and Leu394 **(Figure:- 3.3.22 B)**. To determine the stability of the docked complex and interaction we simulated the complex for 100 ns **(Figure: - 3.3.23)**. Further simulation of complex structure shows no large changes in RMSD, R_g and RMSF (**Figure: - 3.3.23)**. Hence, it can be concluded that binding of CstF50 binding to BARD1 ARD-BRCT stabilizes the ARD domain flexibility. Hydrogen bonding analysis between the BARD1 ARD-BRCT domain and CstF50 shows that linker region has intact hydrogen bonding interactions with CstF50 C-terminal throughout the simulation **(Figure: - 3.3.23 C)**.

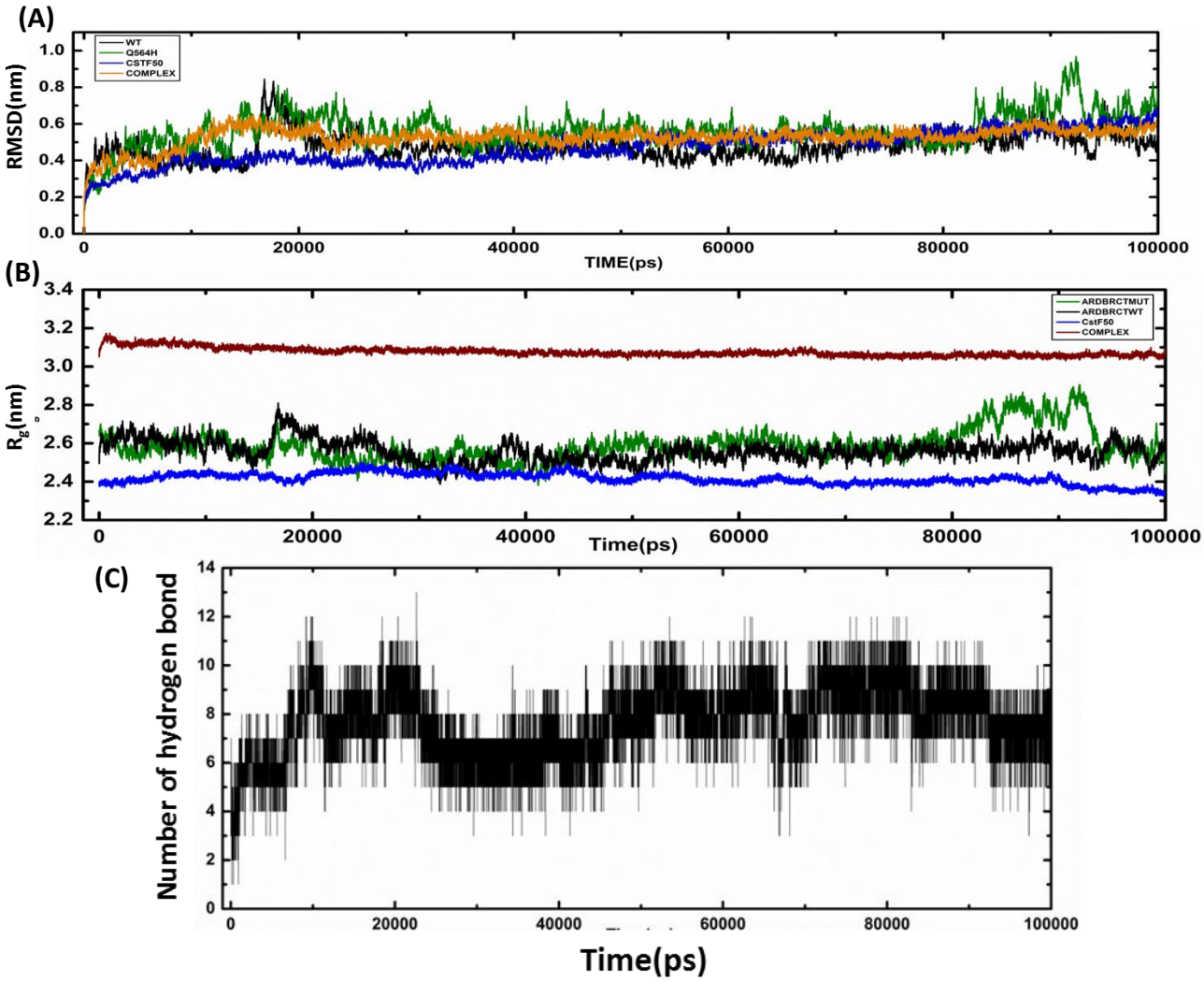


Figure 3.3.23: - **(A and B)** Comparative RMSD and R_g of BARD1 ARD-BRCT *wild-type* (Black), Gln564His (green), CstF50 (blue) and complex (yellow) respectively, and **(C)** Hydrogen bonding profile for BARD1 ARD-BRCT-CstF50 complex.

Furthermore, cross-correlation of PCA for BARD1 ARD-BRCT, CstF50 and complex indicates that positive correlation has increased within the N-terminal region of CstF50; it shows increased positive residual correlation within the WD40 repeat and N-terminal residues in CstF50 in the complex.

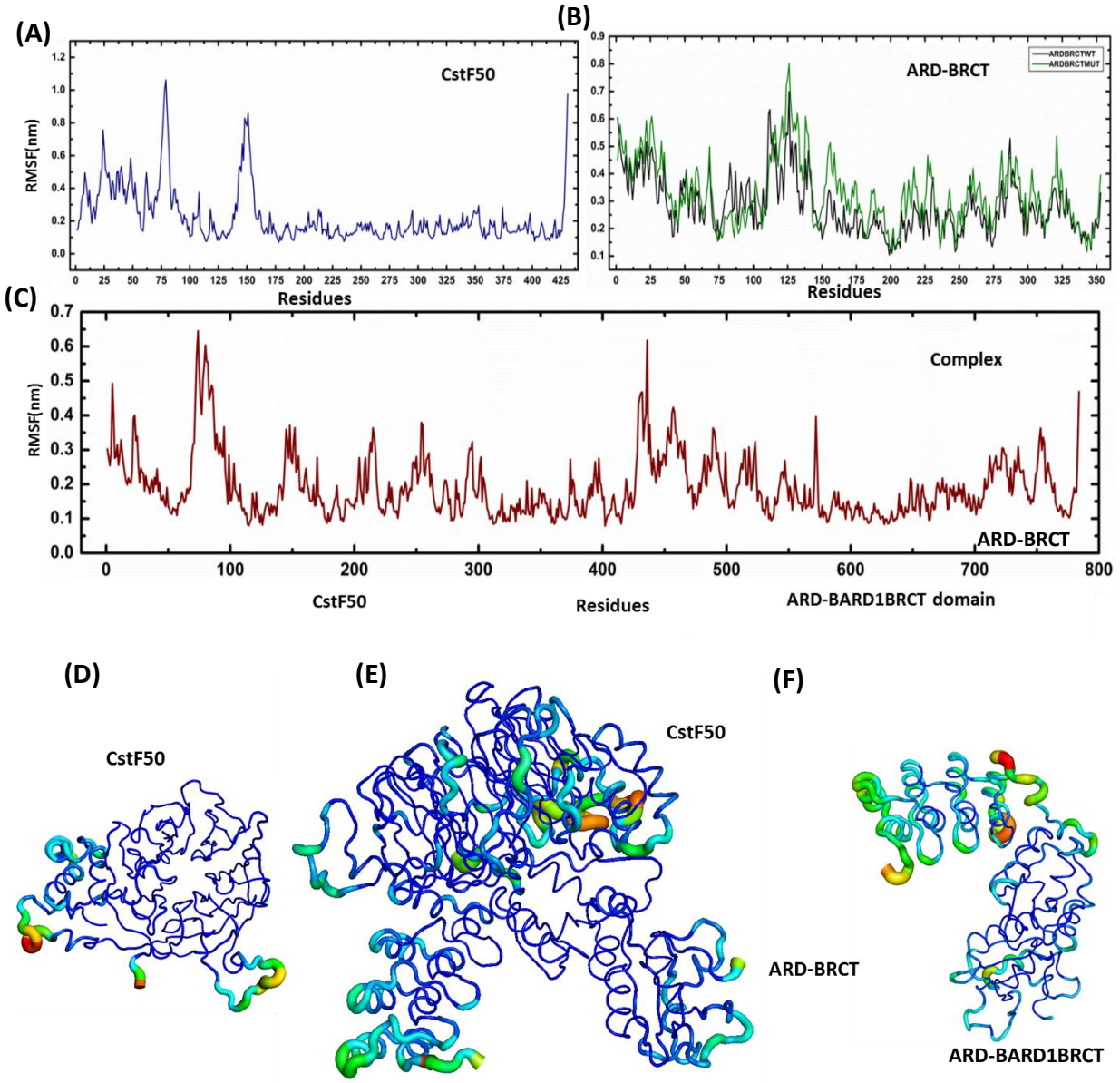


Figure 3.3.24: - Comparative RMSF **(A, B and C)** of BARD1 ARD-BRCT *wild-type* (Black), Gln564His (green), CstF50 (blue) and BARD1 ARD-BRCT *wild-type*-CstF50 complex (Red) respectively, **(D)** RMSF structure for CstF50, **(E)** complex, and **(F)** BARD1 ARD-BRCT.

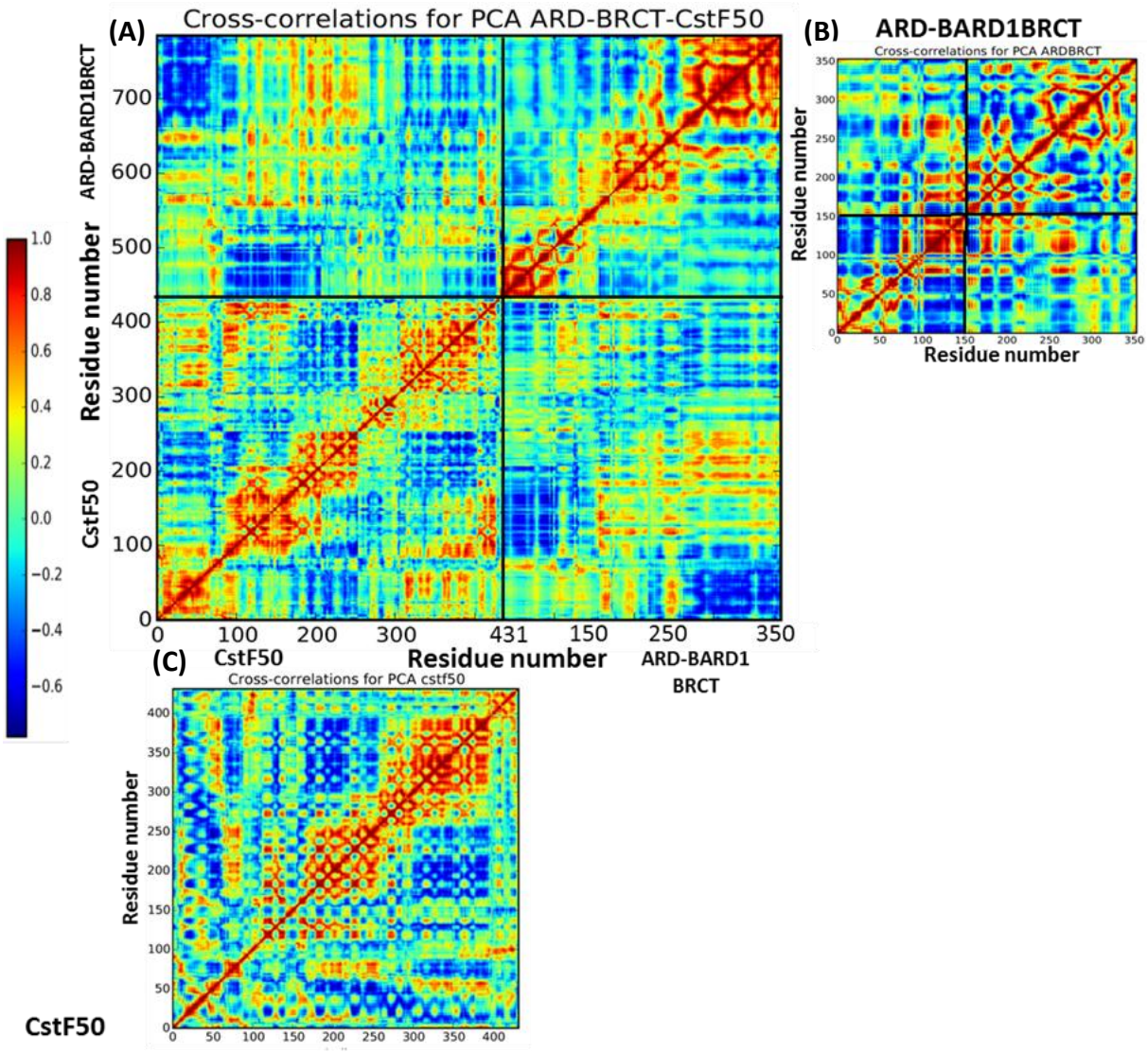


Figure 3.3.25:- Comparative cross-correlation for PCA of **(A)** BARD1 ARD-BARD1BRCT-CstF50 complex, **(B)** BARD1 ARD-BRCT *wild-type* and **(C)** CstF50.

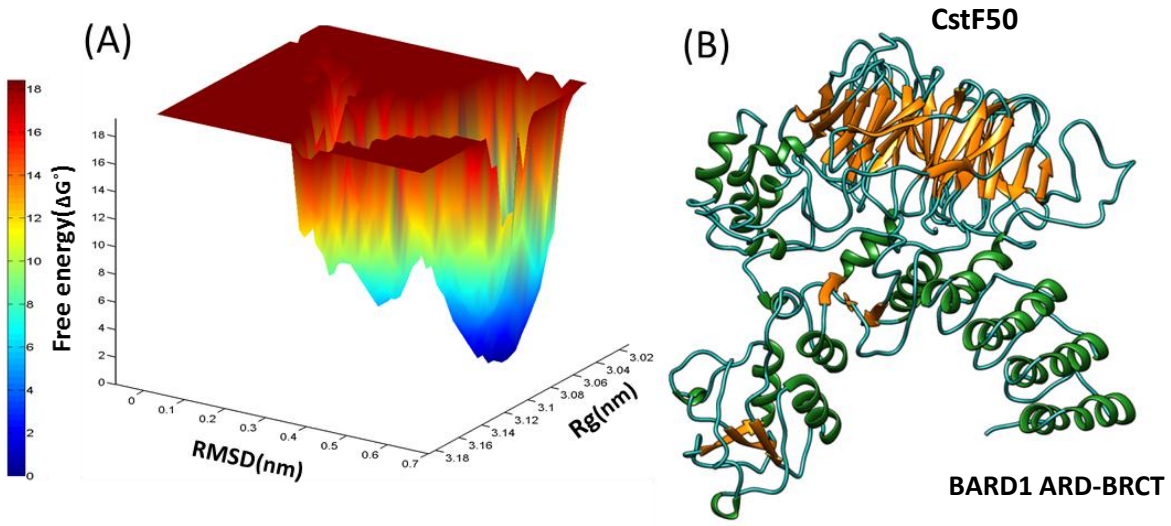


Figure 3.3.26: - Free-energy calculation for the ARD-BRCT-CstF50 complex. **(A)** Free-energy landscape of BARD1 ARD-BRCT *wild-type* CstF50 complex and **(B)** Model of structure of the BARD1 ARD-BRCT-CstF50 complex.

Moreover, cross-correlation of PCA for BARD1 ARD-BRCT in complex also shows significant increase in the positively correlated motion within the ARD-domain and BRCT domain **(Figure: - 3.3.23)**. The representative complex structure derived after free-energy calculation shows CstF50 binds to the BARD1 ARD-BRCT domain in such a way that it prevents the motion of ARD domain and keep it in a compact conformation **(Figure:- 3.3.26)**. Further distance calculated between ARD domain and BRCT domain for *wild-type* and complex structure indicates that CstF50 binding has reduced the domain dynamics of BARD1 ARD-BRCT *wild-type* protein structure (**Figure: - 3.3.27)**.

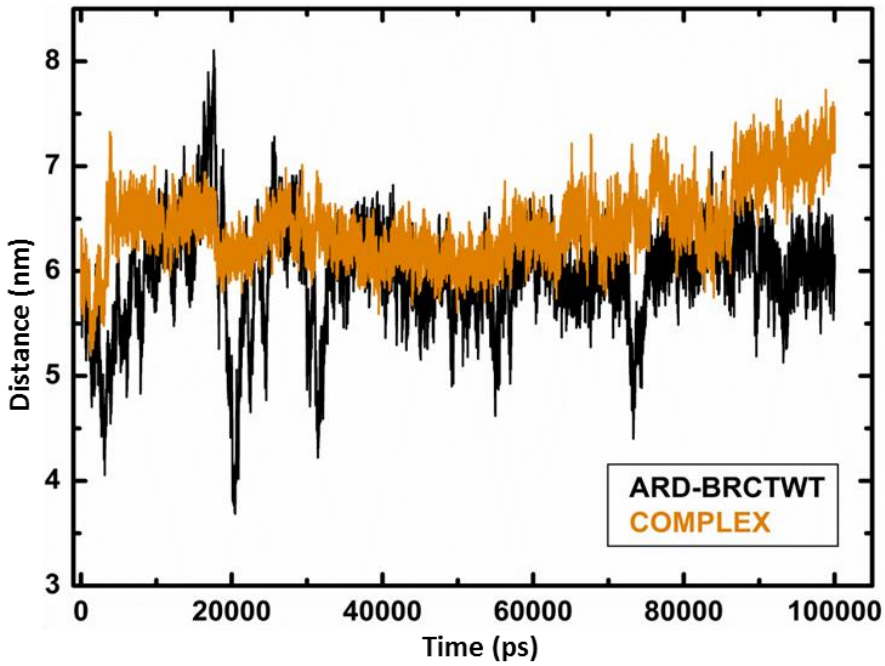


Figure 3.3.27: - Distance calculation between ARD domain and BRCT domain for the structure BARD1 ARD-BRCT *wild-type*- CstF50 complex

In BARD1 ARD-BRCT *wild-type* ARD and BRCT domain shows a fluctuation of 5.5nm, when alone. However, distance fluctuation calculated for ARD and BRCT domain in complex has drastically reduced demonstrating that CstF50 binding restricts the domain dynamics of ARD domain.

3.4 Crystallization of BARD1 ARD-BRCT

An attempt was made to crystallize the BARD1 ARD-BRCT protein from 7 mg/ml of protein concentration using vapor diffusion method. The protein and mother liquor solution was mixed in 1:1 ratio (1 µl+1µl) and allowed to crystallize at 22°C with 500-µl reservoir solution in vibration free incubator. A clear drop or light precipitation was detected in most of the crystallization conditions. No attempt was made to crystallize CstF50 as refolding process

yielded very a low concentration of protein and furthermore, it precipitated during concentration.

3.5 Conclusions:-

It has been observed that BARD1 ARD-BRCT *wild-type* and mutant Gln564His are monomeric in nature and mutation does not affect the biophysical property of *wild-type* protein. CstF50 was found insoluble in bacterial system, therefore efforts were made to purify from inclusion body. NMA analysis and deformation analysis predicted large domain fluctuation between ARD and BRCT domain of BARD1. Molecular dynamics simulation data shows high correlation with small angle X-ray scattering (SAXS) of BARD1 ARD-BRCT domain (425-777aa) [147]. Docking studies indicate that linker region residue Gln564 plays important role in complex formation between BARD1 ARD-BRCT domain and CstF50. Binding of CstF50 to BARD1 ARD-BRCT domain provides rigidity to the structure and moderates the inter-domain motion between ARD and BRCT domain **(Figure:- 3.3.28)**.

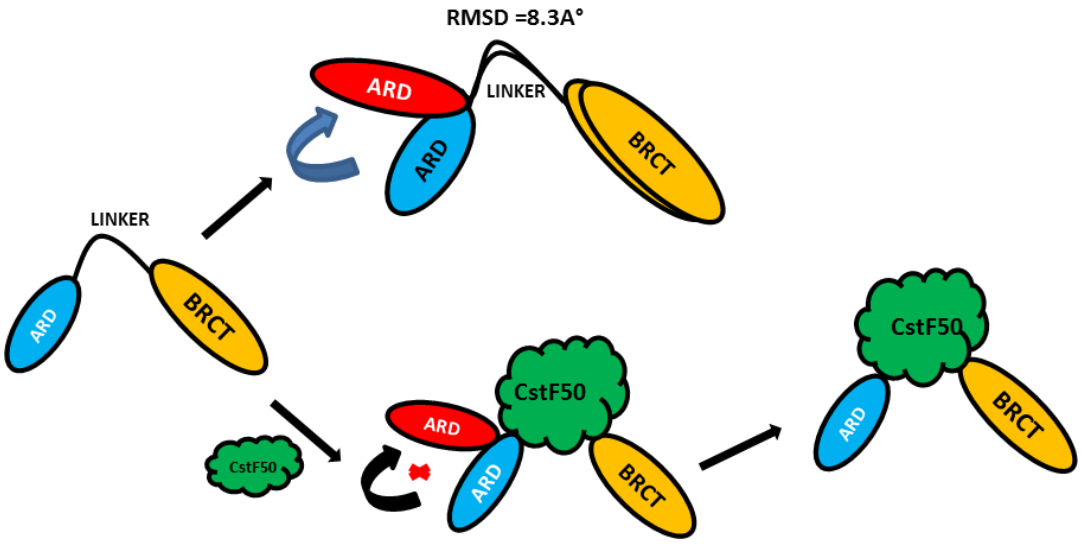


Figure 3.3.28: - Model for BARD1 ARD-BRCT and CstF50 complex formation.

Chapter

4

Biophysical Characterization and *In-Silico* studies of Cancer Predisposing Mutations Discovered In BARD1BRCT Domain

4.1 Introduction

Mutations in the *brca1* and *brca2* genes are responsible for breast and ovarian cancer predisposition [141, 259, 260]. BRCA1 share a great deal of similarity with tumor suppressor protein BARD1 in domain organization, as both harbors N-terminal RING domain and two tandems BRCT repeat motifs at C-terminus [59]. However, BARD1 contains four ankyrin repeats in the region from 425-550 amino acids, which is absent in BRCA1 [59, 64, 237]. BRCA1 & BARD1 heterodimerizes through N-terminal RING-RING domain interaction, and co-localize to distinct nuclear sites at during the cell-cycle progression [72, 73]. Furthermore, *in-vivo* BRCA1-BARD1 heterodimer is an active E3-ubiquitin ligase complex, which plays a significant role in its tumor suppression function [21, 63, 64]. Mis-sense mutations within the BRCA1 RING domain lead to loss of heterodimerization, and reduces E3 ubiquitin ligase activity [63]. It has been reported that exon 11-deficient BRCA1 splice variants which only retain the RING and BRCT domain are targeted to IR-induced foci, suggesting that RING domain and BRCT are required for foci formation [261]. Furthermore, tandem BRCT domain when transfected in U2OS cell, foci formation was restored, and this reversal clearly indicates that BRCA1 BRCT alone is needed for DNA damage foci after IR mediated DNA damage [30].

The BRCT domain is established as a phosphopeptide recognition motif, and is conserved in many other multiple DNA damage-response (DDR) proteins [142, 143]. Several proteins are found to be harboring BRCT domain and having role in DNA damage repair such as base excision response scaffold protein XRCC1, DNA ligase IV, BRCA1, MDC1, p53-binding protein 53BP1 and BARD1 [142, 143, 161]. Different DDR proteins have diverse repeats of BRCTs [262]. It has been reported that BARD1BRCT is required for

the early recruitment of BRCA1 at the DNA damage site [18]. Furthermore, BRCA1 Cys61Gly mutation in the RING domain abrogates the BRCA1-BARD1 heterodimer formation and early recruitment of BRCA1 at the DNA damage foci, leading to the conclusion that BARD1BRCT is required for the BRCA1 recruitment and tumor suppression [18]. Recently it has been identified that ADP-ribosylated proteins are new interacting partners of the BARD1BRCT domain [263]. However, no binding was observed during Isothermal Titration Calorimetry with ADP-ribose and PAR with other $BRCT_s$ like BRCA1 [18].

Noting the role of BARD1 in genomic integrity, we have decided to explore functional consequences of genetic alterations in the BRCT region. In this study, we have used *In-vitro,* biophysical and *in-silico* approaches to understand the folding pattern of BARD1BRCT *wild-type* and cancer predisposing mutants BARD1 Cys645Arg, Val695Leu, Arg658Cys, Ile738Val and Ser761Asn. Cancer predisposing mutations, Arg658Cys have been detected in Caucasian, African, Finnish population and Ile738 Val in Polish and Belgian families with unknown functional consequence [27, 32, 264-266]. The families were also tested negative for both BRCA1 and BRCA2 inherited mutations in sporadic and BRCA1 associated breast cancers [33, 267]. BARD1 Ser761Asn missense mutation is reported in uterine and breast cancer, Cys645Arg mutation in breast and ovarian cancer, but Val695Leu mutation only in breast cancer [27, 33]. The mutants Cys645Arg and Val695Leu failed to show any interaction either with ADP-Ribose or with its polymeric form PAR, and also failed in early recruitment of BRCA1 to the DNA damage site [18]. Therefore, *In-silico* and *In-Vitro* approach was applied to investigate the structural changes acquired due to these point mutations, and also to assess the effect of protein folding patterns on their ability to interact with phospho-peptide

pSer-X-X-Phe, and could posed pathogenic implications for breast cancer risk. Molecular dynamics simulation (MDS) was carried out to explore alterations in the structure of *wild-type* and mutant proteins. To our conclusion, Val695Leu, Arg658Cys Ile738Val and Ser761Asn mutants except BARD1 Cys645Arg show loss of thermodynamic stability. However, BARD1 Ile738Val mutant protein structure showing higher structural flexibility.

4.2 Material and Methods

A comparative structural and biophysical characterization of BARD1BRCT (568-777aa) *wild-type* and mutants were performed. BRCT domain of BARD1 was sub-cloned into pET28a vector (Amersham Pharmacia) using BARD1 full length cDNA (a kind gift from Richard Baer, USA). The basic scheme for sub-cloning was to use oligonucleotide primer complementary to the template sequence for PCR. The mutants were generated using primer mis-match method and further confirmed by DNA sequencing.

The gene of interest was amplified using following PCR condition: 98°C denaturation (5 minutes), 98°C denaturation (45 seconds), and annealing 57°C for 35 seconds, extension 72°C at 0.5kb/min, final extension 72°C for 10 minutes, and 25 cycles. The PCR amplified product was digested with EcoR1 and BamH1 and ligated in the pGEX-kT vector (a kind gift from J. Ladias, Harvard Medical School-BIDMC, Boston US). The ligation mixture was transformed into *E.coli* DH5α cells. The final ligated product was confirmed by DNA sequencing.

Primer details:-

FORWARD PRIMER C645R 5'-GTCTACGAAGAAAAGTACGTGAACAGGAAGAAAAG- **3'**

REVERSE PRIMER C645R 5'-CTTTTCTTCCTGTTCACGTACTTTTCTTCGTAGAC-**3'**

FORWARD PRIMER V695L 5'-CCTTATTAAGCTCCTCACTGCAGGTGGGGGCC-**3'**

REVERSE PRIMER V695L 5'-GGCCCCCACCTGCAGTGAGGAGCTTAATAAGG-**3'**

FORWARD PRIMER S761N 5'-GGAAGGCTCCTTCGAACTGGTTTATAGACTG-**3'**

REVERSE PRIMER S761N 5'-CAGTCTATAAACCAGTTCGAAGGAGCCTTCC-**3'**

FORWARD PRIMER R658C 5'-GAAATTCCTGAAGGTCCATGCAGAAGCAGGCTCAACAG-**3'**

REVERSE PRIMER R658C 5'-CTGTTGAGCCTGCTTCTGCATGGACCTTCAGGAATTTC-**3**'

Forward primer I738V 5'-CTGCACACAGTATATCGTCTATGAAGATTTGTGT-**3'**

Reverse primer I738V 5'-ACACAAATCTTCATAGACGATATACTGTGTGCAG-**3'**

BARD1BRCT (568-777aa) *wild-type* and mutants were expressed in Rosetta 2(DE3) bacterial system.

Protocol for protein expression:-

1. **Inoculation:** A single colony from antibiotic resistant LB agar plate was inoculated in 100 ml LB broth preinoculum containing 100 µg/ml of ampicillin. Further, grown at 37^{o}C overnight in shaking incubator at 200 rpm.
2. **Dilution:** Added 10 ml of pre inoculum to 1000 ml (1: 100) of autoclaved LB broth containing 100 µg/ml of ampicillin. Grown medium on shaker incubator at 37^{o}C until it reached mid-log phase i.e. A_{600} between 0.6-0.8.
3. **Induction:** Cooled down the flasks, added 400 µl IPTG (stock 1M), and incubated on shaker incubator at 24^{o}C for 16 hours.
4. **Harvesting:** The culture was centrifuged for 10 minutes at 6000 rpm at 4^{o}C. The pellet was resuspended in a small volume of supernatant and centrifuged for 15 minutes at 5000 rpm, 4^{o}C.

5. **Storage:** The bacterial pellet obtained was stored at -80^{o}C for further use.

Protocol for purification of BARD1BRCT and mutants:-

Purification buffer*:* 50 mM HEPES (pH 7.5), 300 mM NaCl, 2mM EDTA, 5% glycerol pH 7.5 (buffer A)

FPLC buffer*:* 50 mM HEPES (pH 7.5), 300 mM NaCl (buffer B)

Protocol for protein purification:-

1. **Re-suspension**: Re-suspended the induced bacterial pellet of BARD1BRCT and mutants in 40 ml of buffer A; supplemented with 200 mM PMSF and 20µl of protease inhibitor.
2. **Ultra sonication**: Transferred the resuspended cell pellet into centrifuge tube and sonicated at 70-pulse rate and 50 power with 1 minutes of duty cycle. Repeated the cycle 5 times with 1 min break.
3. **Centrifugation:** After sonication, the suspension was subjected to centrifugation at 18000 rpm for 45 minutes at 4^{o}C to obtain cleared lysate. Collected the supernatant and discard the cell debris.
4. **Binding**: The soluble protein fraction obtained was brought at room temperature, then mixed with pre-equilibrated affinity resin and incubated at room temperature for 1 hour. The sepharose beads in affinity column charged with Ni^{2+} binds specifically to those proteins having HIS (Histidine) tag. Take out 40 µl bead for binding check.
5. **Washing**: After binding, the protein bound column was washed with 4 column volume of buffer A to remove non-specific bound proteins.

6. **Cleavage**: Added 200 µl (20 units) of TEV protease, 40 µl of protease inhibitor cocktail and 100 µl of PMSF in 20 ml of buffer A and completed the cleavage step in 3 hours by passing the solution through column at interval of 1 hour. Take out 10 µl of resin to detect the cleavage of protein.

7. **Elution**: After TEV cleavage, the protein was eluted with 30 ml of buffer A**.**

8. **Equilibration of Ni-NTA resin**: Provided 2 column washes with double distilled water and then 5 to 6 column washes with buffer A.

9. **Metal Ion Chelate Affinity Chromatography**: After calibration of Ni-NTA resin, passed the eluted fractions through them to get rid of His-tagged TEV protease contamination.

10. **Concentrating the protein:** Transferred the eluted protein in a 10 KDa pre-equilibrated centricon, and concentrate the protein up to 2 ml by centrifuging at 4500 rpm for 10 minutes at 4^{o}C. Check the concentration on Nanodrop spectrophotometer (280nm). Centrifuge for additional 10 minutes at 13000 rpm at 4^{o}C to removal insoluble aggregates or precipitates.

11. **Gel filtration:** Injected 2 ml (or as per availability of FPLC loop) of concentrated protein in AKTA-FPLC against Buffer B.

12. **Fraction collection**: Collected the purified protein eluted through FPLC in 1.7 ml microfuge tube at its elution volume according to gel filtration spectra profile of the sample.

13. **Loading on SDS-PAGE 12% gel**: Loaded 20 µl of FPLC fractions on SDS-PAGE, stained with coomassie dye, and then destained it to visualize the protein of interest.

14. **Concentration of protein**: The fractions, which showed purified protein band was concentrated as per the 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, purification and characterization of BARD1BRCT and Mutants:-

Selected clones were digested with the EcoR1 and BamH1 restriction enzymes, and insert release was observed at appropriate size on agarose gel. DNA sequencing results have confirmed the presence of the gene of interest in the vector with desired frame of codon sequence. Biophysical and computational characterization of BARD1BRCT *wild-type* and mutants were performed to investigate the conformational changes due to the mutations. Biophysical characterizations were performed using CD spectroscopy, fluorescence spectroscopy, chemical crosslinking and limited proteolysis. Whereas, computational techniques included molecular dynamics simulation and PCA (Principal component analysis). Oligomeric character of purified BARD1BRCT and mutants were studied with respect to proteins of known molecular weight by injecting the proteins in superdex-75 column. In AKTA-FPLC, both *wild-type* and mutant proteins eluted at the 65 ml of the column volume in superdex-75, which suggests the insignificant effect of mutation on the monomeric nature of the BARD1BRCT (**Figure:- 4.1**). Glutaraldehyde cross-linking was also performed to characterize the oligomeric property, and it has been observed that there is no higher molecular weight oligomers formed in *wild-type* and mutant proteins.

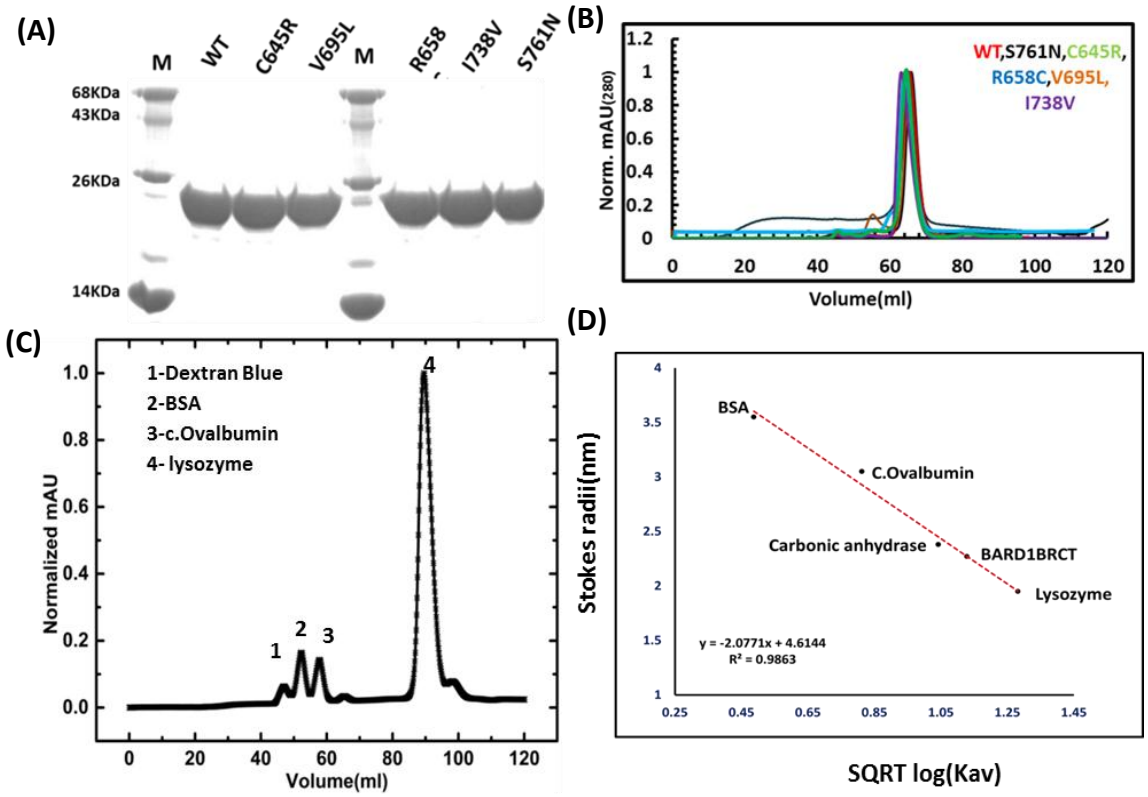


Figure 4.3.1:- **(A)** Expression and purification of BARD1BRCT and mutant proteins, **(B)** FPLC purification profile of *wild-type* and mutants,**(C)** Elution profile of standard known molecular weight proteins and **(D)** Stokes radii calculation of BARD1BRCT *wild-type* protein.

Hence, it can be concluded that mutations do not change the monomeric property of the BARD1BRCT (**Figure: - 4.3.2**). CD spectroscopy was performed in far-UV and near-UV range to study the secondary and tertiary structure of the proteins **(Figure: - 4.3.3)**. Purified monomeric *wild-type* and mutants do not show any large change in CD spectroscopy profile, which suggests that mutations did not, affects the secondary and tertiary structure of the proteins **(Figure:- 4.3.3)**.

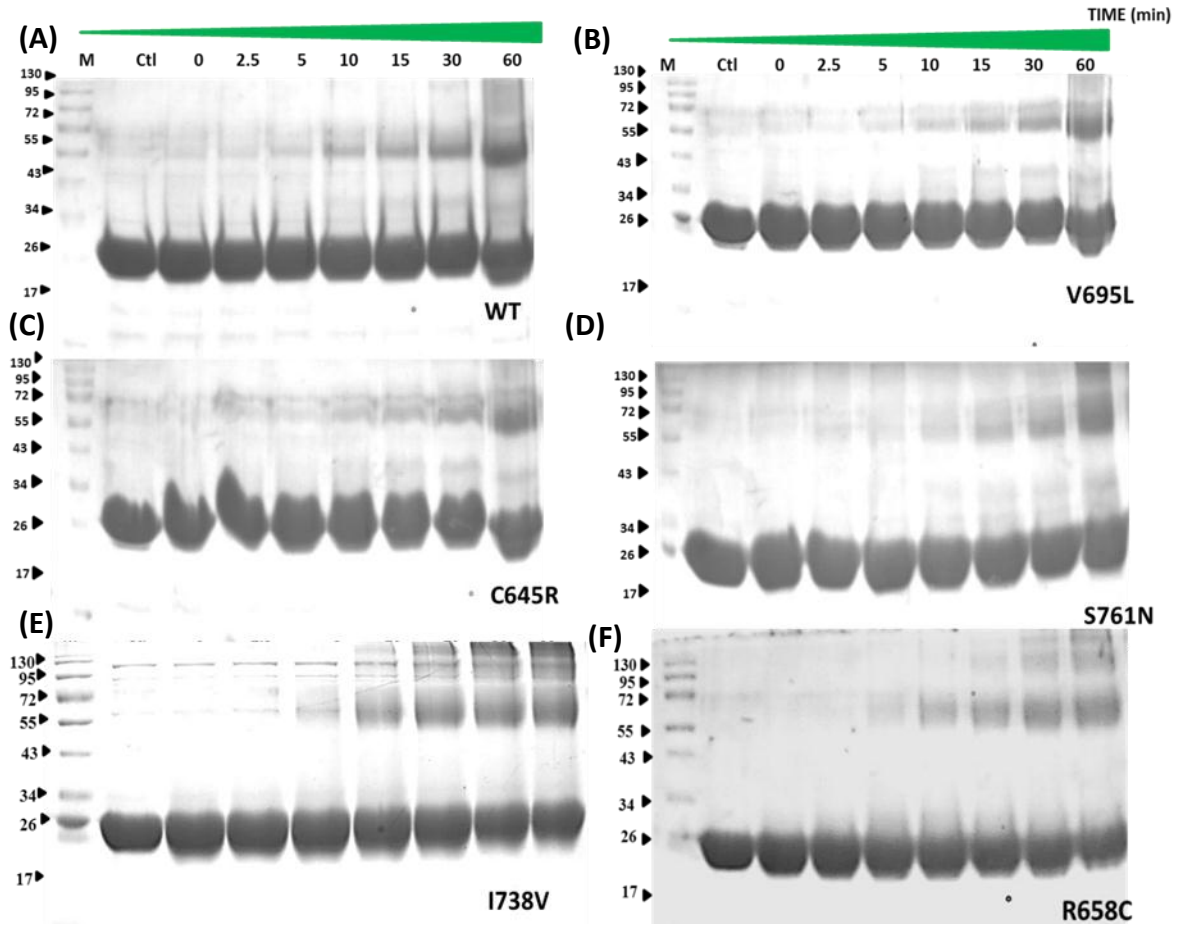


Figure 4.3.2:- Comparative chemical crosslinking profiles of BARD1BRCT and mutant proteins.

Nevertheless, local changes induced in the structure, which is beyond the sensitivity of CD cannot be ignored. To investigate the overall 3D packing, chemical denaturation using fluorescence spectroscopy was performed for *wild-type* and mutant proteins. *Wild-type* and mutant proteins show an emission maximum for folded protein in the range λ=332-336nm indicating the complete burial of tryptophan residue inside the hydrophobic core, hence signifying that mutations do not affect the compactness of BARD1BRCT.

4.3.2 Structural characterization and thermal stability

To study the effect of mutations on thermodynamic stability, thermal denaturation and chemical denaturation using GuHcl was performed.

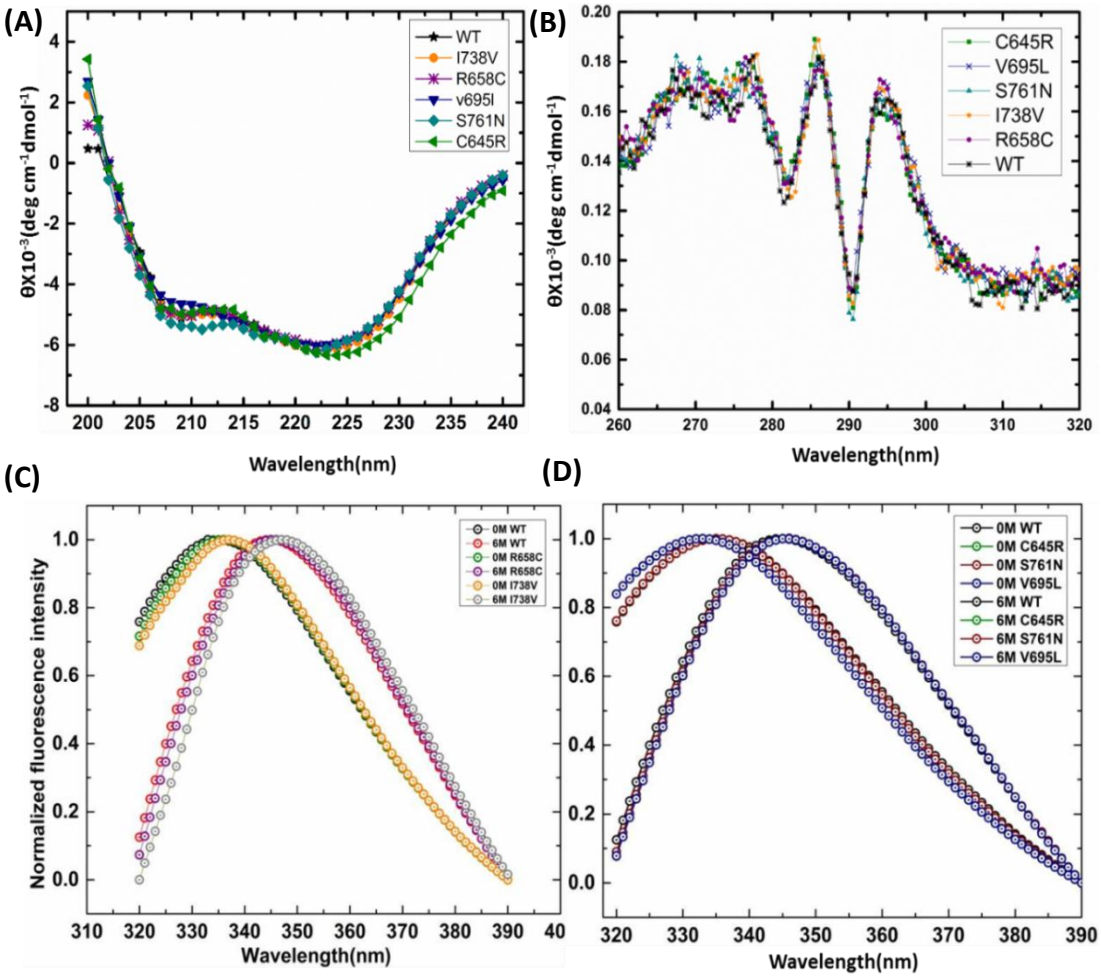


Figure 4.3.3:- **(A and B)** secondary and tertiary structure characterization of BARD1BRCT domain and mutant protein using CD spectroscopy. **(C and D)** Chemical denaturation profile of BARD1BRCT w*ild-type* and mutants using fluorescence spectroscopy.

Thermal denaturation spectra of BARD1BRCT *wild-type* and mutants in far-UV range and near-UV range suggest that except BARD1 Cys645Arg, mutant proteins Val695Leu, Arg658Cys, Ile738Val, Ser761Asn have lost thermodynamic stability and all mutant proteins unfold via a two state pathway (**Figure:- 4.3.4A)**.

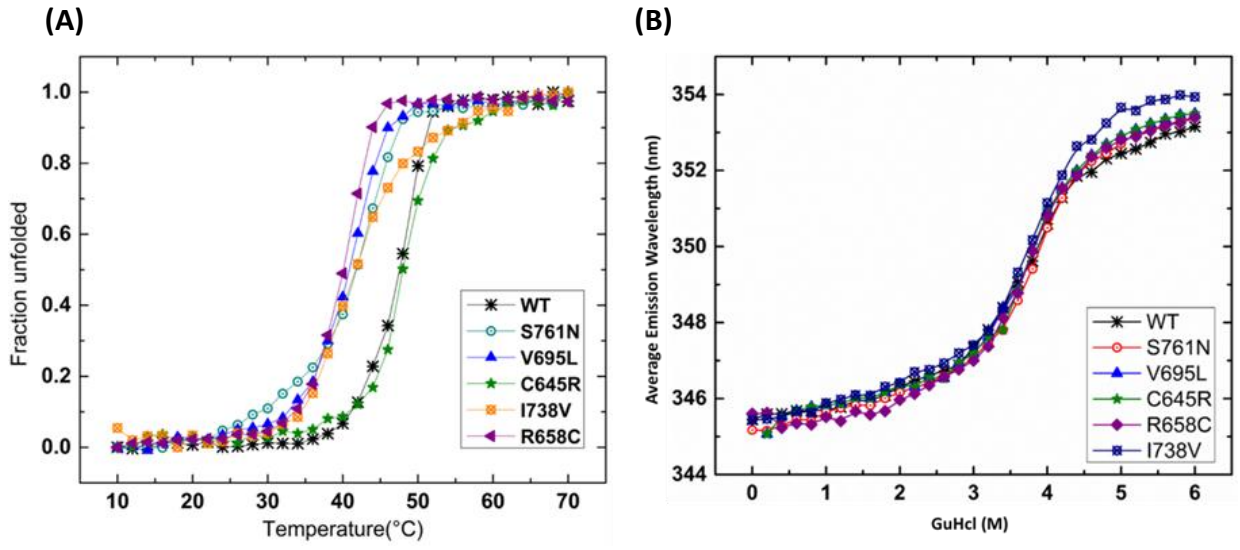


Figure 4.3.4:- **(A)** Thermal and chemical denaturation of *wild-type* and mutant proteins using CD and **(B)** fluorescence spectroscopy.

Near-UV range thermal denaturation also has good corroboration with far-UV range data, except BARD1 Cys645Arg, Val695Leu, Ser761 Asn, Arg658Cys and Ile738Val mutants have lost their overall packing at 45˚C, and completely unfold at 60˚C (**Figure:- 4.3.5)**. T_m calculated for BARD1BRCT *wild-type*, Cys645Arg, Val695Leu, and Ser761Asn mutant proteins were 47.7±.85°c, 47.6±.65°c, 42.3±.78°c, and 41.1±.42°c respectively. ΔG calculated for BARD1BRCT *wild-type*, Cys645Arg, Val695Leu, Ser761Asn, Arg658Cys and Ile738Val mutant proteins were 9.8±0.39kcal/mol, 9.6±0.11kcal/mol, 7.4±0.69kcal/mol, 7.1±0.12kcal/mol, 7.3±0.12 kcal/mol, and 7.6±.42 kcal/mol respectively. Chemical denaturation study shows that Val695Leu and Ser761Asn, Arg658Cys and Ile738Val mutants have lost thermodynamic stability compared to *wild-type* protein and Cys645Arg mutant proteins and unfold via a three state pathway **(Figure:- 4.3.4B)**. ΔG calculated for BARD1BRCT *wild-type*, Cys645Arg, Val695Leu, Ser761Asn, Arg658Cys and Ile738Val

mutant proteins were 7.19±.36 kcal/mol, 6.28±0.82kcal/mol, 6.2±0.45kcal/mol, 6.11±0.18kcal/mol, 6.29 ± .41 kcal/mol, and 6.8± .21 kcal/mol respectively.

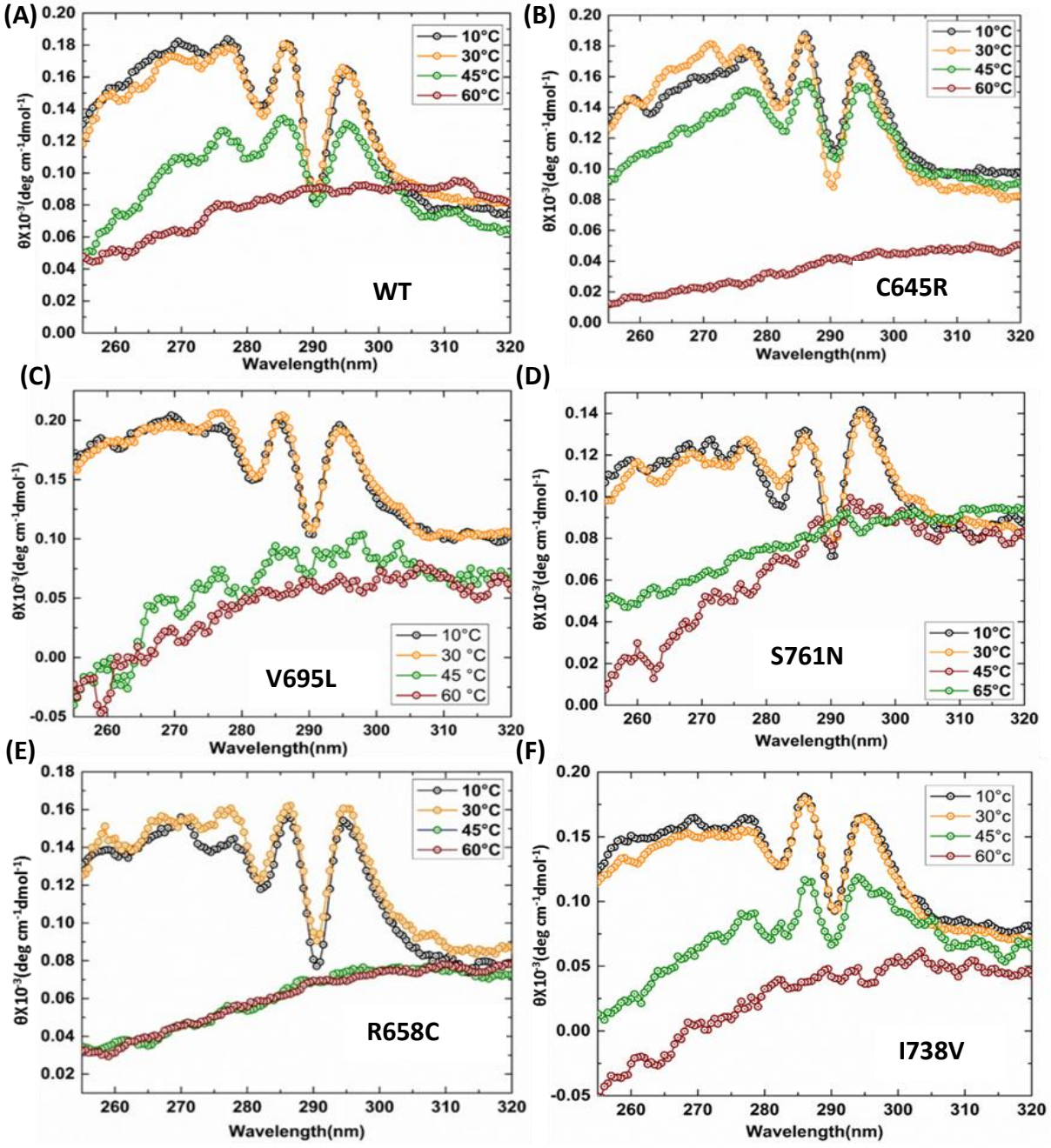


Figure 4.5:- **(A-F)** Thermal denaturation of BARD1BRCT *wild-type* and mutant protein using CD spectroscopy in near- UV range (λ=350-250nm).

To confirm BARD1BRCT *wild-type* and mutants unfold via a three state pathway thermal denaturation studies of *wild-type* and mutants using fluorescence spectroscopy was performed.

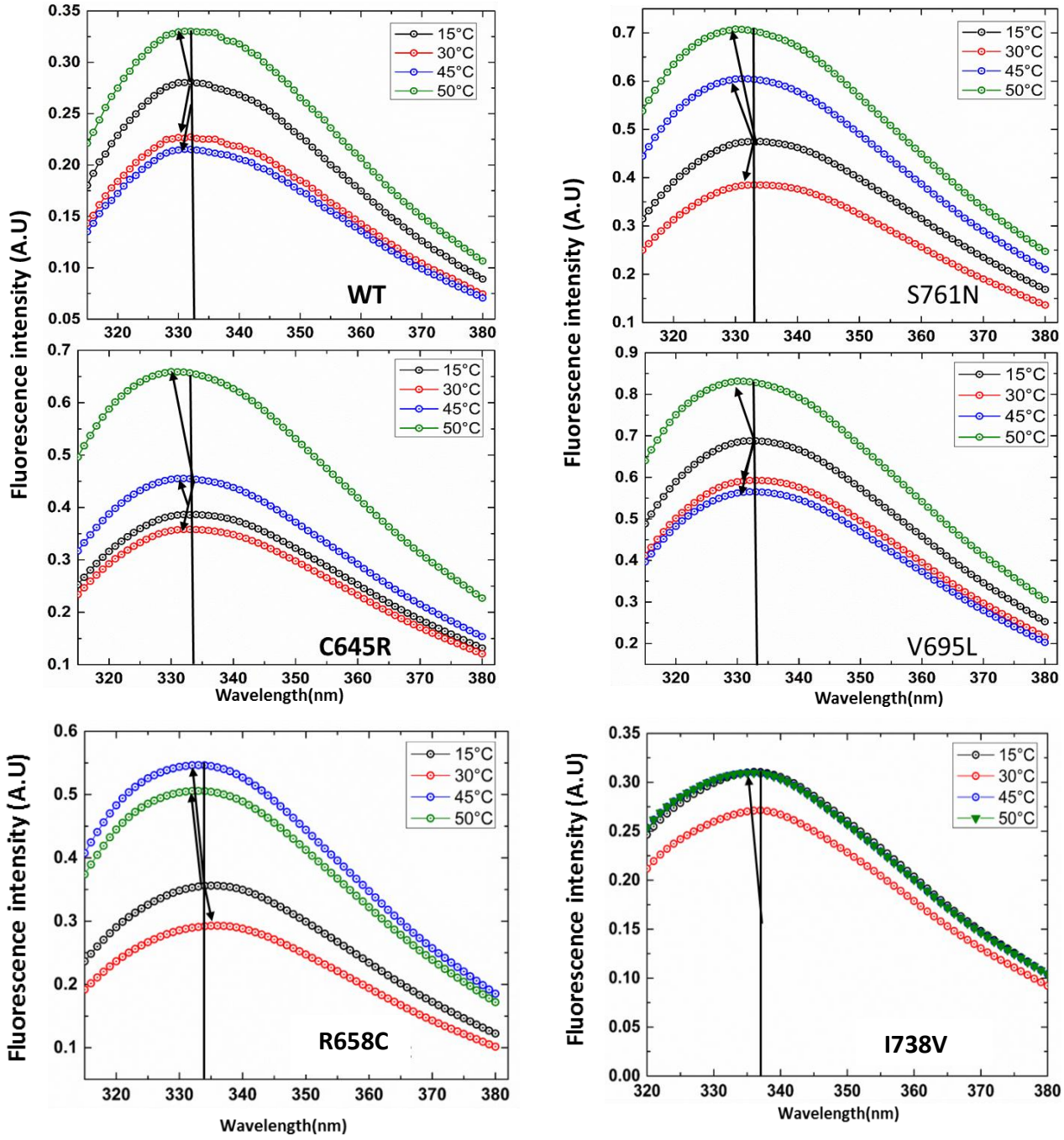


Figure 4.6: - Thermal denaturation of BARD1BRCT *wild-type* and mutant proteins using fluorescence spectroscopy.

Wild-type and mutant proteins show a non-monotonic shift in the intensity of emission maximum. While the fluorescence emission intensity for *wild-type* and mutants first gradually decreases from 15˚C to 45˚C, and further it shows a sharp rise in intensity when temperature increased to 50˚C. These observations are consistent with the presence of a molten globule intermediate during the three-step denaturation process. A similar blue shift in the emission maximum with an increase in the temperature indicating unfolding through molten globule formation is consistent with the conclusions drawn by chemical denaturation studies as well **(Figure:- 4.3.6 and Table 4.1)**.

Table 4.1: - Comparative emission maximum attained by BARD1BRCT *wild-type* and mutant protein during thermal denaturation using fluorescence spectroscopy

	Wild-type	**Ser761 Asn**	**Val695 Leu**	**Cys645Arg**	**Arg658Cys**	**Ile738 Val**
Tm	**45.8°C**	**40°C**	**39.7°C**	**45.2°C**	**40.6°C**	**41.7°C**
Emission maxima recorded at different temperatures for BARD1BRCT *wild-type* and Mutants						
	WT	**S761N**	**V695L**	**C645R**	**R658C**	**I738V**
15^0C	**332 nm**	**335 nm**	**333 nm**	**333 nm**	**334nm**	**336nm**
30^0C	**330 nm**	**333 nm**	**333 nm**	**333 nm**	**336nm**	**335nm**
45^0C	**328 nm**	**331 nm**	**331 nm**	**331 nm**	**333nm**	**335nm**
50^0C	**329 nm**	**329 nm**	**329 nm**	**330 nm**	**331nm**	**334nm**

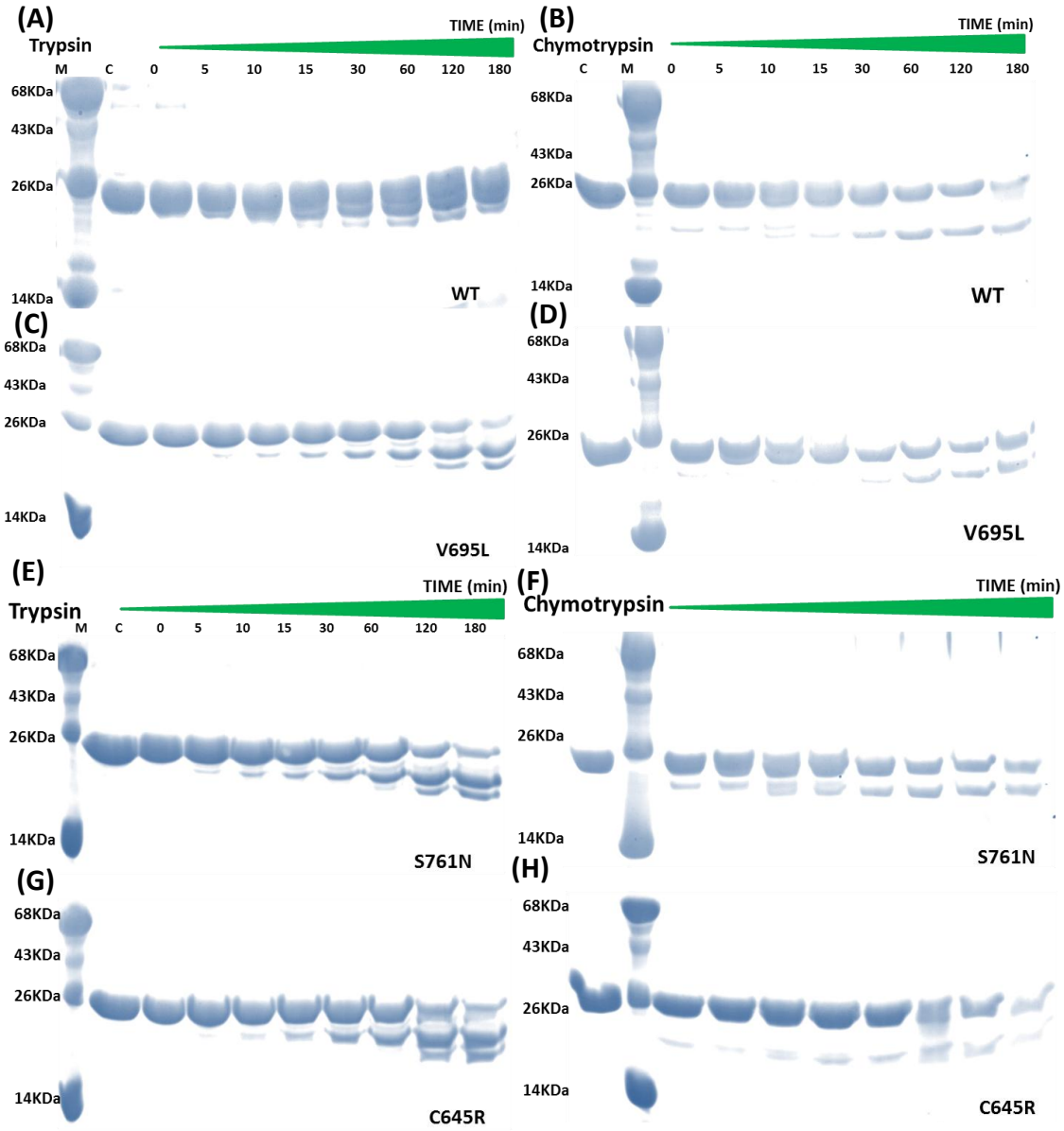


Figure 4.3.7: - Comparative limited proteolytic profile of BARD1BRCT domain and mutant protein using trypsin and chymotrypsin **(A and B)** *wild-type*, **(C and D)** Val695Leu, (**E and F)** Ser761Asn and **(G and H)** Cys645Arg respectively.

BARD1BRCT and mutants were subjected limited proteolysis using trypsin and chymotrypsin to study the effect of the mutation on overall folding of protein. Trypsin digestion profile of *wild-type* and mutants concluded that proteins are resistant to trypsin digestion and mutation has not brought about any drastic change in domain organization (**Figure: - 4.3.7)**.

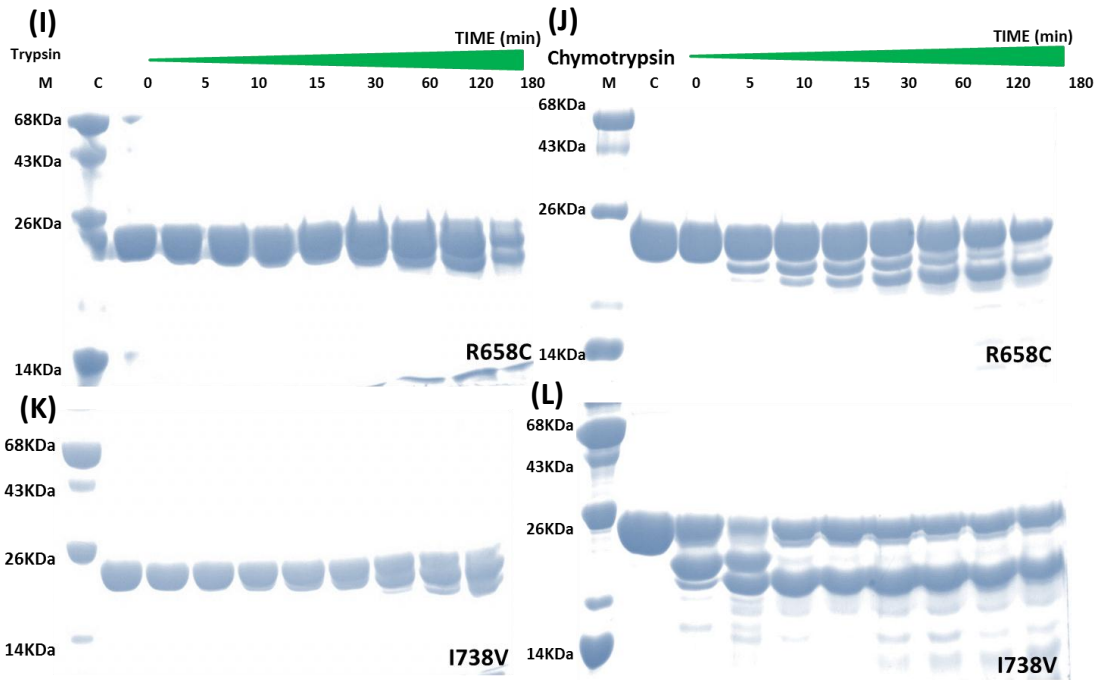
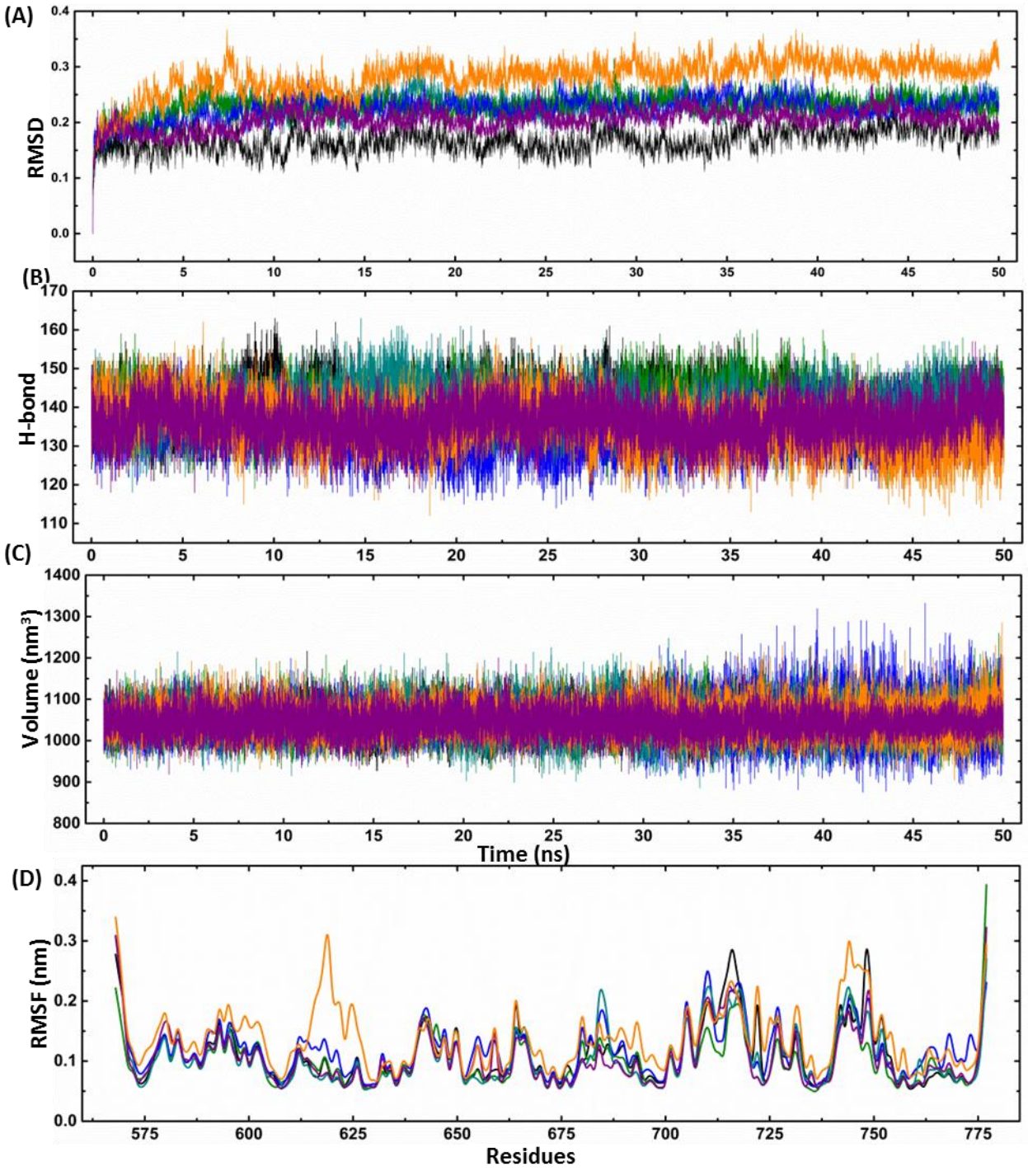


Figure 4.3.8:- (I and J) Limited proteolytic profile of BARD1 Arg658Cys and **(K and L)** BARD1 Ile738Val respectively using trypsin and chymotrypsin.

Similarly, chymotrypsin digestion profile shows that, except BARD1 Ile738Val mutant protein, *wild-type* and other mutant proteins Cys645Arg, Val695Leu and Ser761 Asn and Arg658Cys show similar resistivity profile in a time-dependent manner. Hence, it can be concluded that BARD1 Ile738Val has lost overall packing locally **(Figure: - 4.3.8 L)** as compared to *wild-type* and mutant proteins Cys645Arg, Val695Leu, Ser761Asn and Arg658Cys **(Figure: - 4.3.7).**

4.3.3 Structural characterization of BARD1BRCT and mutants:- A Molecular dynamics simulation approach

To explore the effect of mutations on the BARD1BRCT structure, molecular dynamics simulations was performed on *wild-type* and mutants structures for a period of 50ns.



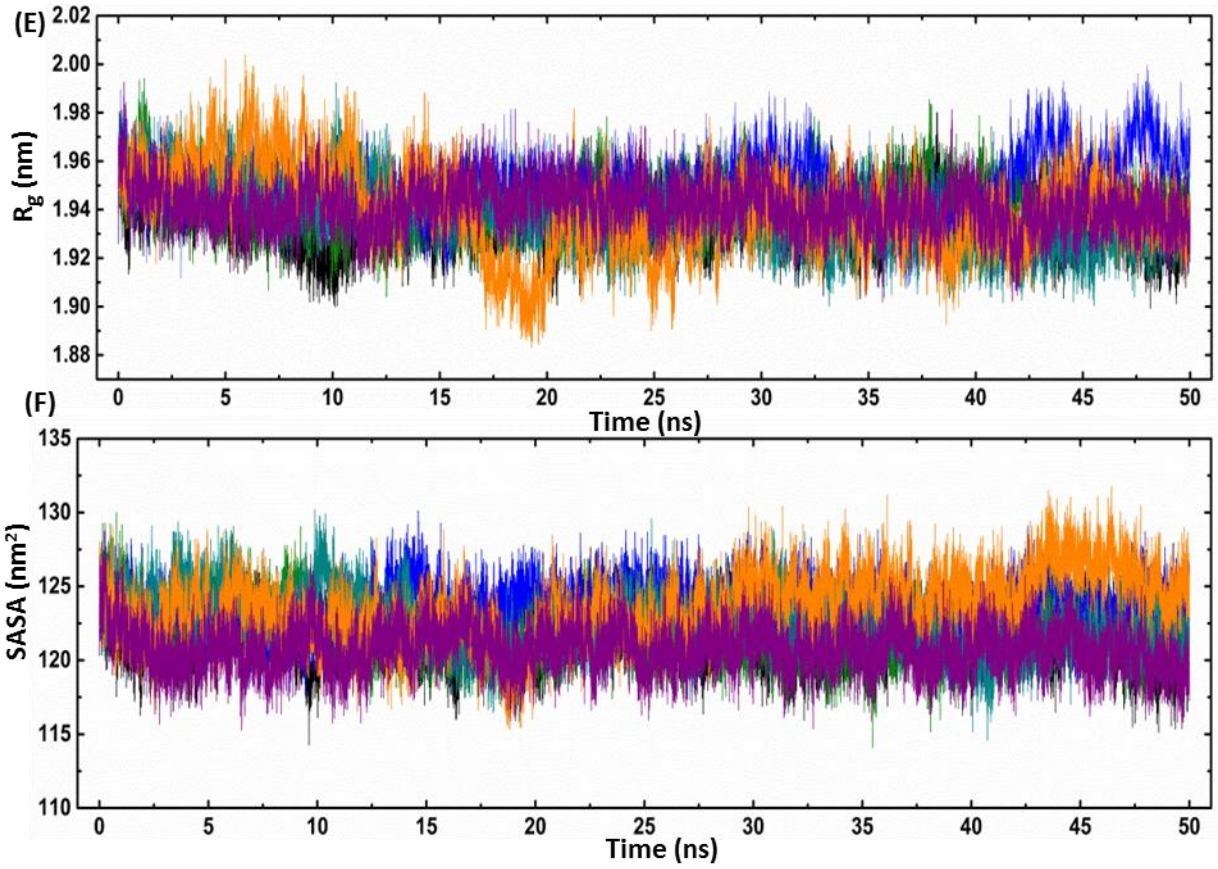


Figure 4.3.9:- Comparative representation of **(A)** RMSD, **(B)** RMSF, **(C)** H-bond, **(D)** Rg, **(E)** Volume, and **(F)** SASA. The *wild-type* is represented as black, Cys645Arg as green, Val695Leu as blue, Ser761Asn as cyan, Arg658Cys as purple and Ile738Val yellow respectively.

The RMSD Fluctuations for BARD1 Ile738Val protein structure were observed to be comparatively higher than *wild-type* and mutant proteins Cys645Arg, Val695Leu and Ser761 Asn and Arg658Cys. This observation suggests that Ile738Val mutation destabilizes the hydrophobic core of the protein that is also reflected in the fluctuating values of radius of gyration attained by the Ile738Val mutant structure **(Figure: - 4.3.9)**.

The destabilizing effect of BARD1 Ile738Val mutation is also very well revealed in the RMSF profiles of Ile738Val mutant, as overall RMSF value for Ile738Val is high as compared with the *wild-type* and mutant proteins Cys645Arg, Val695Leu and Ser761 Asn and Arg658Cys. Mutant proteins Val695Leu, Ser761Asn, Ile738Val and Arg658Cys except Cys645Arg mutant structure show reduced number of intra-molecular hydrogen bonds compared to *wild-type* protein structure. Average H-bond formed during different periods of simulation suggests that except BARD1 Cys645Arg mutant proteins, Val695Leu and Ser761 Asn, Ile738Val and Arg658Cys mutants show significant decrease in the number of hydrogen bond **(Figure:- 4.3.10)**. Furthermore, BARD1 Ile738Val and Val695Leu show increase in volume and solvent accessible surface area, which is consistent with the reduced number of hydrogen compared to *wild-type*, Cys645Arg, Arg658Cys and Ser761Asn mutant proteins.

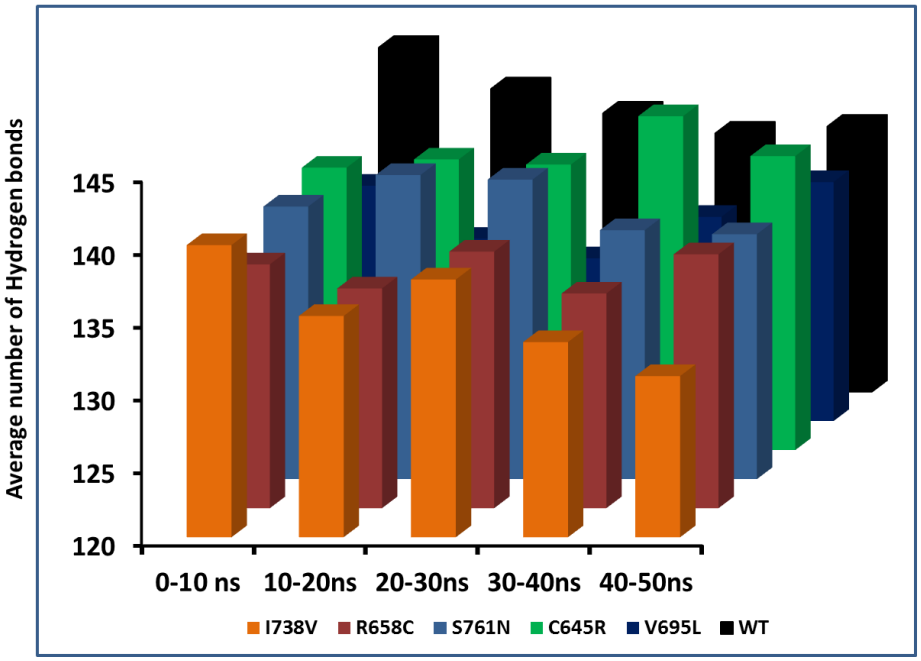


Figure 4.3.10: - Comparative average intra-molecular hydrogen bond formed by the *wild-type* and mutant structures.

4.3.4 PCA of BARD1BRCT *wild-type* and mutant proteins:-

Principal component analysis was performed over the *wild-type* and mutant trajectory to study the comparative effect of mutation on the concerted motion of the protein.

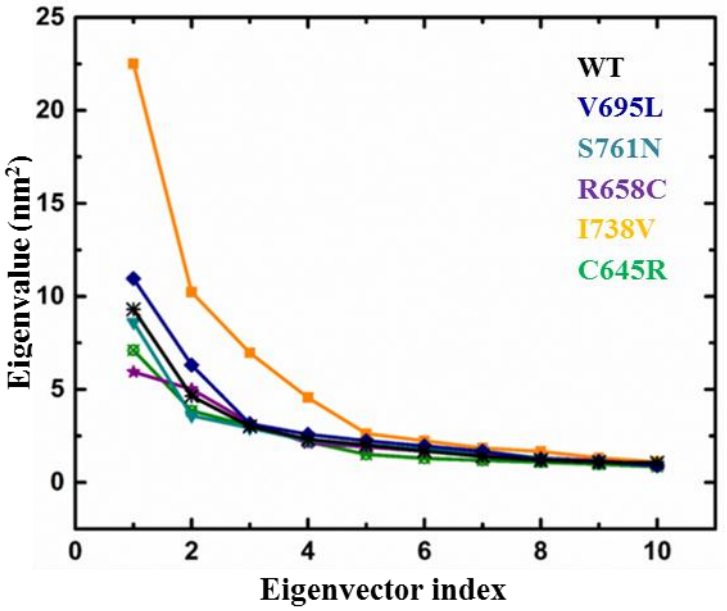


Figure 4.3.11:- Scree plot of mutant and *wild-type* structure.

Eigenvector and eigenvalues were calculated after diagonalization of the covariance matrix for the *wild-type* and mutant proteins Val695Leu, Ser761 Asn, Ile738Val, Arg658Cys and Cys645Arg. Scree plot for *wild-type* and mutants shows the magnitude of Eigen value for each eigenvector and it is evident that first few eigenvector describes more than 90% of the protein dynamics. It may be seen that Ile738Val mutant has the highest eigenvalue for first few eigenvector as compared to *wild-type* and mutant protein Val695Leu, Ser761 Asn, Ile738Val and Arg658Cys and Cys645Arg mutant structures **(Figure:- 4.3.11)**. Trace of covariance matrix which is sum of all the eigenvalues provides important information illustrating the dynamics of the protein. Trace of co-variance matrix calculated for Ile738Val

was highest compared to *wild-type* and mutant proteins Val695Leu, Ser761Asn, Arg658Cys and Cys645Arg suggesting that mutation Ile738Val structure is more flexible than the *wild-type* and mutants proteins structures. Further, projections of 1st three eigenvectors were performed for *wild-type* and mutant protein structures. As, the projection of eigenvector 3 on 1 and 2 on 1 for Ile738Val has large periodic transition in the tertiary structure than the *wild-type* and Val695Leu and Ser761Asn, Arg658Cys and Cys645Arg mutant proteins.

Table 4.2:- Comparative tabulated values of trace of covariance matrix attained by BARD1BRCT *wild-type* and mutant

BARD1BRCT and Mutants	**Trace of covariance matrix**
WT	**53.89nm^2**
Cys645Arg	**57.22nm^2**
Arg658Cys	**56.35nm^2**
Val695Leu	**64.97nm^2**
Ser761AsnN	**58.49nm^2**
Ile738Val	**92.53nm^2**

These observations are suggestive of larger and varied motion for BARD1 Ile738Val mutant compared to *wild-type*, Val695Leu, Ser761Asn, Arg658Cys and Cys645Arg mutant proteins. Projection of eigenvector 3 on 1 and 2 on 1 shows Ile738Val mutant has highest value of trace

of co-variance matrix attained as compared to the *wild-type* and Val695Leu, Ser761Asn, Arg658Cys and Cys645Arg mutant proteins, signifying that Ile738Val structure has lost

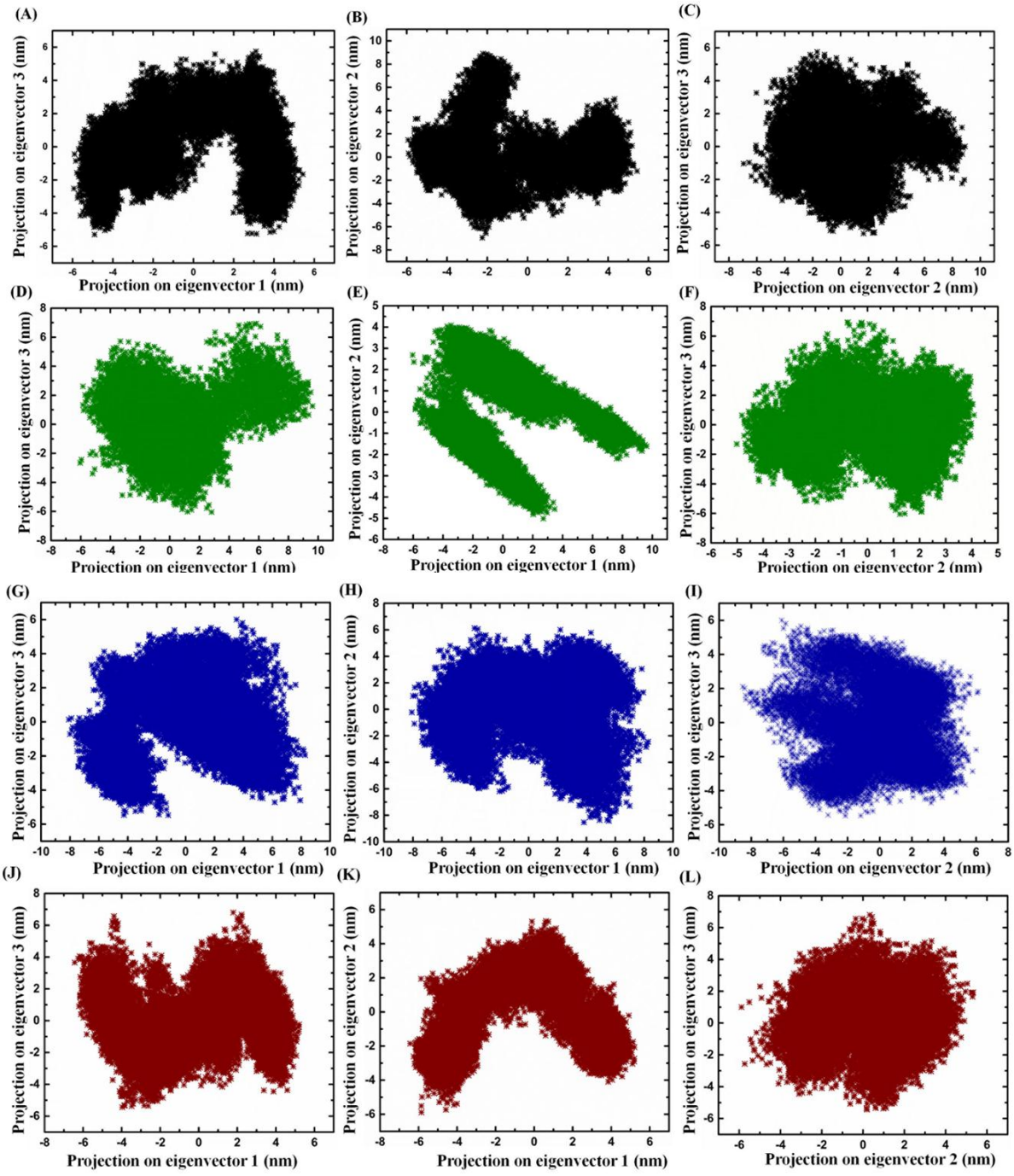


Figure 4.3.12: - Projection of first three eigenvectors for the structure **(A, B, C)** BARD1BRCT *wild-type*, **(D, E, F)** Cys645Arg, **(G, H, I)** Val695Leu and **(J, K, L)** Ser761Asn.

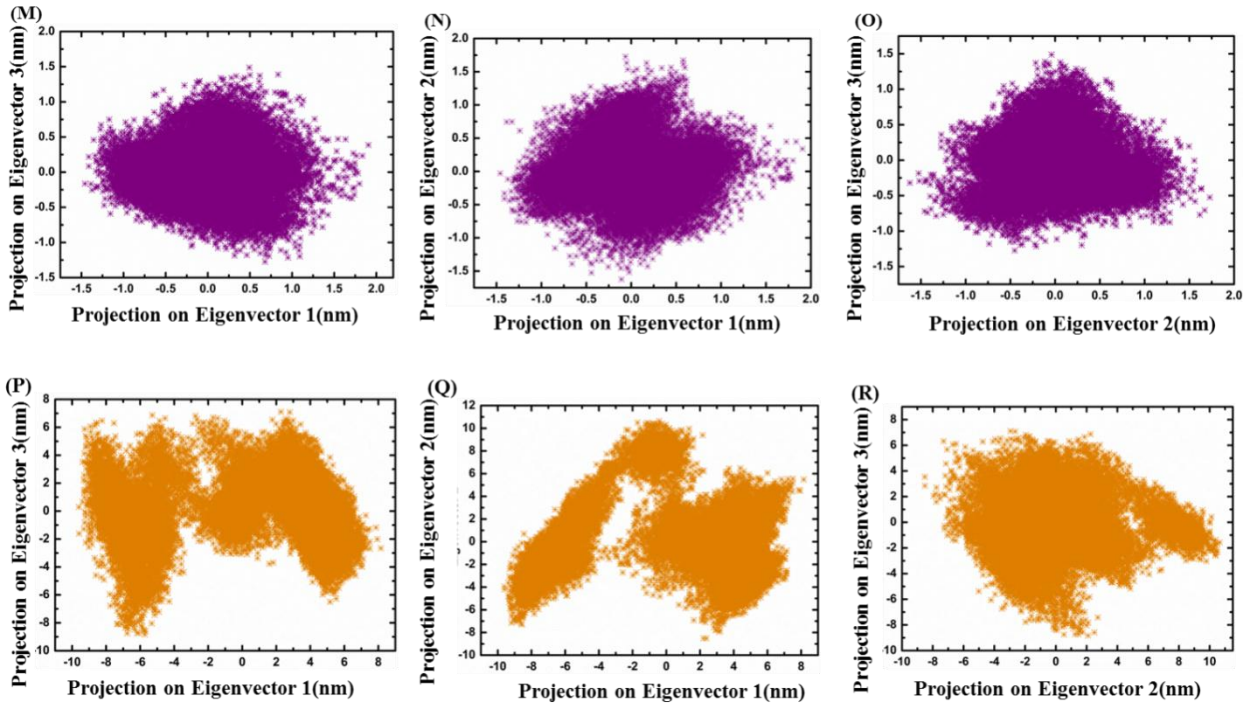


Figure 4.3.13: - (M, N, O) Projection of first three eigenvectors for the BARD1BRCT Arg658Cys and **(P, Q, R)** Ile738Val mutant protein.

structural rigidity and compactness. The eigenvector projection for Ile738Val also shows that the number of tertiary structure conformers attained by Ile738Val mutant is higher than the *wild-type* and other mutant proteins. Similarly, eigenvector projections for Val695Leu mutant structure also show relatively high trace of covariance matrix attained by the Cys645Arg, Arg658Cys and Ser761Asn mutant proteins as compared to the *wild-type* protein, indicating that Val695Leu mutation has mild effect on the structural rigidity of the protein (**Figure:-4.3.12**). Furthermore, Arg658 Cys, Ser761Asn mutant protein structure does not show significant fluctuation in their respective structures but Val695Leu and Ile738Val shows three major groups of tertiary structure conformations as compared to the *wild-type* and mutant protein structure.

4.3.5 Comparative residual displacements in the subspaces spanned by the first two eigenvectors:-

To study the overall effect of mutations at the residual level, we have projected first two eigenvector (which describes the 80% of protein motion) on the c-α residues for *wild-type* and mutant protein structures. Projection of eigenvector 1 on residues for *wild-type* structure shows no large fluctuation at N-terminal region of the structure but the eigenvector 1 and 2 captured the concerted fluctuation in the c-terminal loop region of the protein structure **(Figure:- 4.3.14 A, B)**.

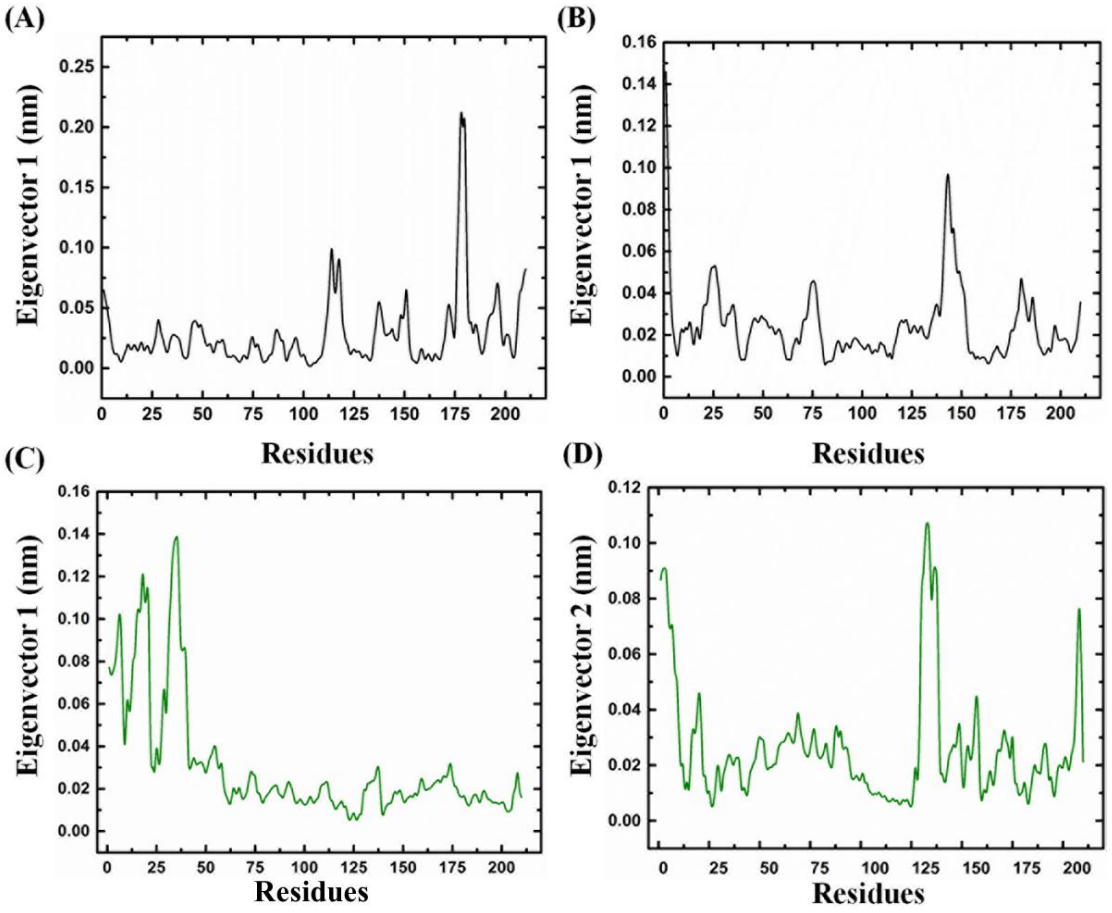


Figure 4.3.14: - Residual fluctuation against first two eigenvectors for the structure **(A, B)** BARD1BRCT *wild-type*, **(C, D)** Cys645Arg mutant.

Projection of eigenvector 1 on residues for Cys645Arg mutant shows significant residual fluctuation in the N-terminal region of the protein structure suggesting that mutation is destabilizing the N-terminal region of the protein structure **(Figure:- 4.3.14 C, D)**.

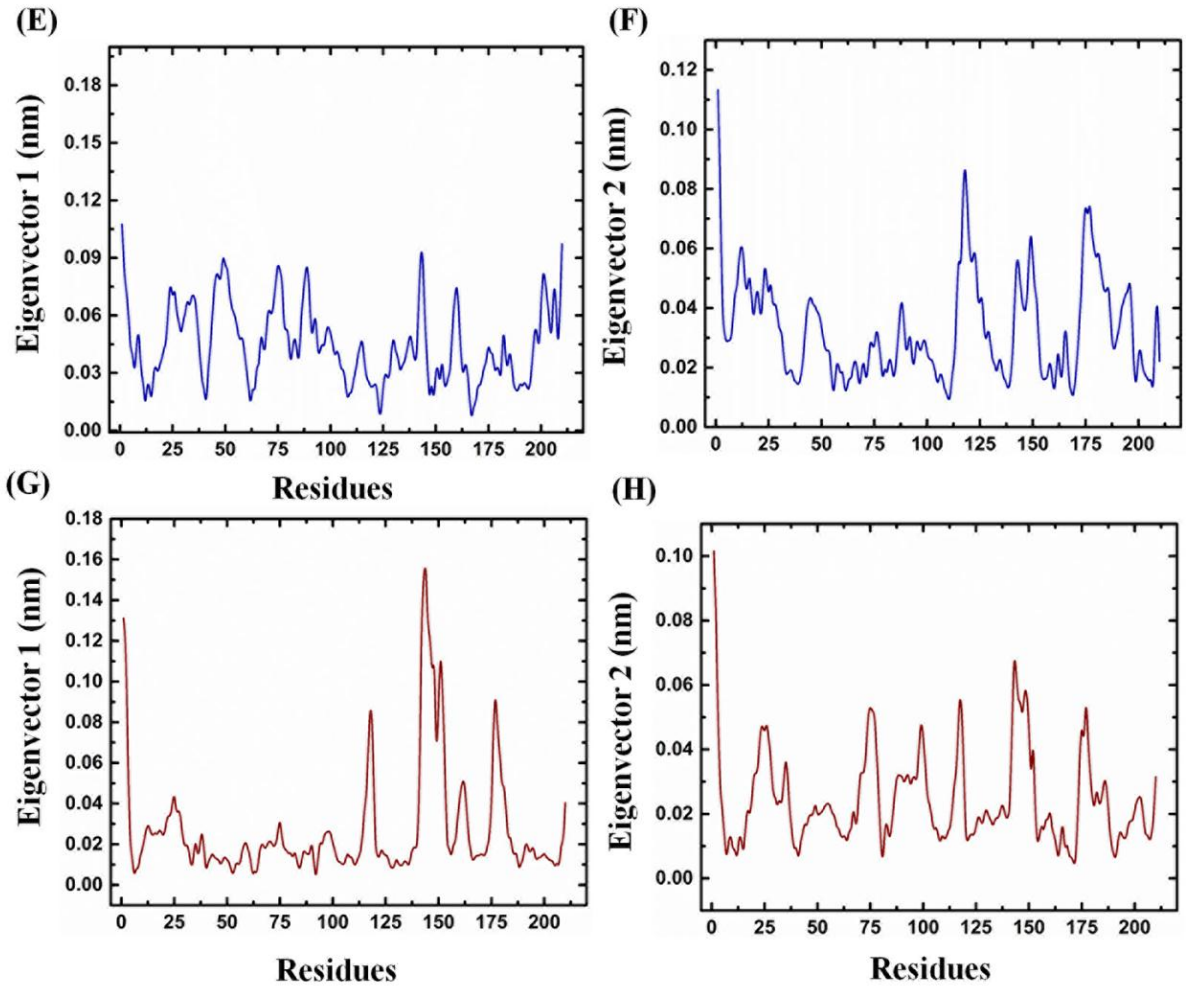


Figure 4.3.15: - Residual fluctuation against first two eigenvectors for the mutant structures **(E, F)** BARD1 Val695Leu and **(G, H)** Ser761Asn mutant.

Projection of eigenvector 2 on residues for Cys645Arg mutant structure also shows local increase in fluctuation in the connecting loop between N-ter BRCT and C-ter BRCT where the mutation site is located, indicating the local increase in the conformational entropy due to the mutation.

Eigenvector 2 projection indicates fluctuation in the β'2 bridging α'2 and β'2**(Figure:- 4.3.14 C, D and 4.3.17).**

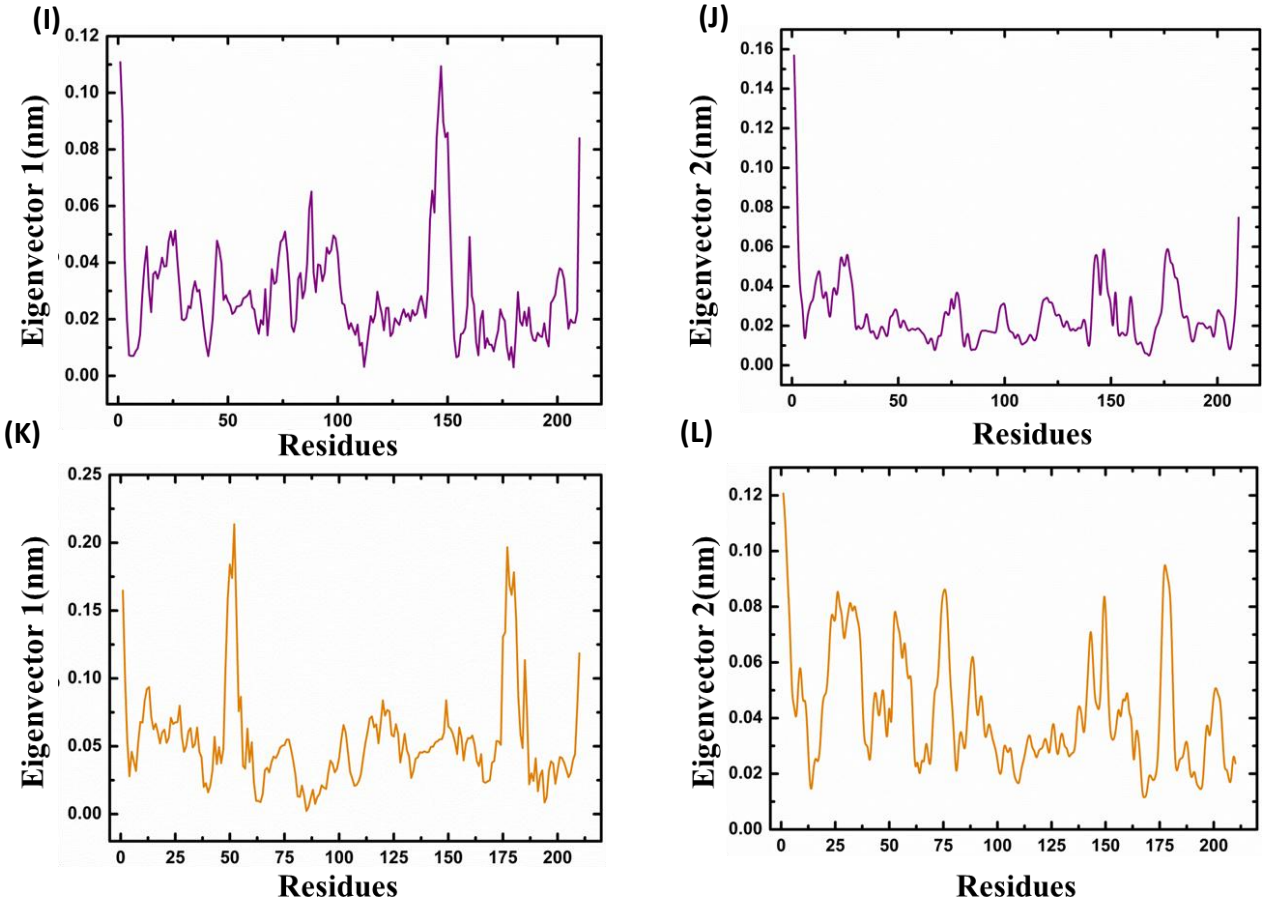


Figure 4.3.16: - Residual fluctuation against first two eigenvectors for the mutant structures **(I, J)** BARD1 Arg658Cys and **(K, L)** Ile738Val.

Projection of eigenvector 1 and 2 on residues also shows good correlation with the RMSF structure for Cys645Arg **(Figure: - 4.3.18)**. As projection of eigenvector 1 against residue for Cys645Arg indicates that N-terminal region showing high fluctuation at region from 568-610 amino acids. Residual amino acid region from 740-750 is also showing high fluctuation of eigenvector 2, which corresponds to β'2 and β'3 connecting loop **(Figure:- 4.3.17)**. Residual projection against eigenvector 2 also suggests that dynamic region between α'2 and β'2 which is between 710-721 amino acids in Cys645Arg mutant structure. Results obtained from

projection of eigenvector fluctuation on residues have very good corroboration with the Cys645Arg RMSF structure **(Figure: - 4.3.18)**.

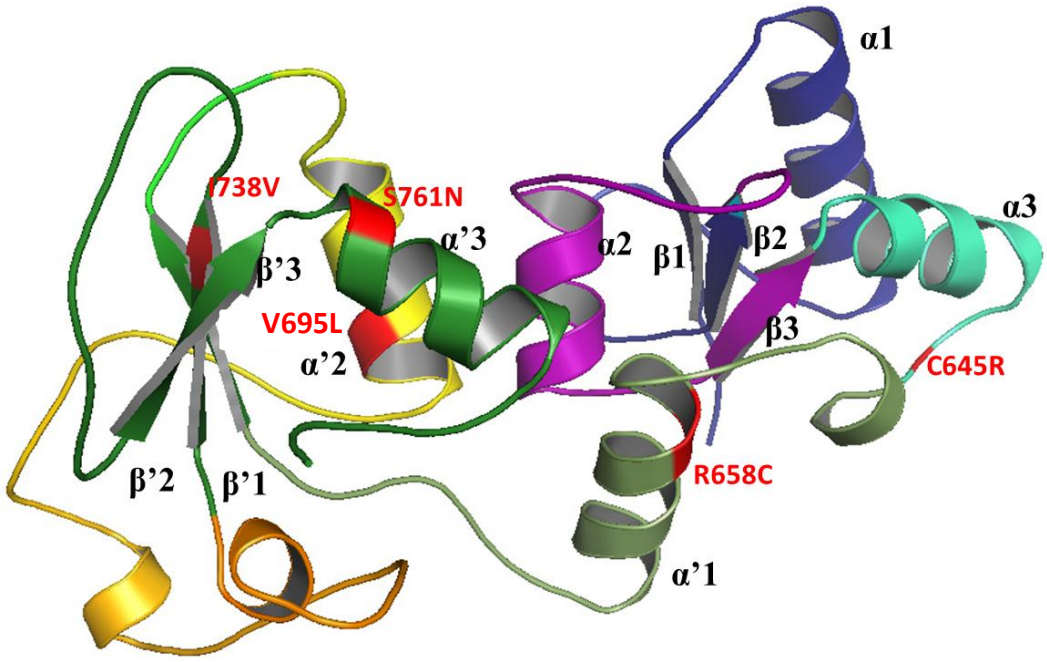


Figure 4.3.17: - Cancer predisposing mutations in different regions of BARD1BRCT domain.

Furthermore, projection of eigenvector 1 and 2 on residues for Val695Leu specifies that mutational effect on the motion is not cantered at local level. As the fluctuation of eigenvector, 1 and 2 on residues is affecting the overall structure of the protein. RMSF structure for Val695Leu indicates that mutation has affected the structure at a global level **(Figure: - 4.3.18)**. Projection of eigenvector 1 on residues for Ser761Asn structure reveals that connecting loop between β'1-α'2, random region between α'2-β'2, and β'2-β'3 is showing fluctuation which has consistency with the RMSF structure **(Figure:- 4.3.15, and 4.3.17)**. Interestingly projection of eigenvector 2 on residue for Ser761Asn structure indicates that effect of amino acid substitution on overall structure in a cumulative way (**Figure:- 4.3.15, and 4.3.17**). As projection of eigenvector 1 against residue for Arg658Cys indicates that N-terminal region showing low fluctuation at N-terminal and low RMSF (blue) regions

shows residual terminal region but has large fluctuation random region between α'2-β'2 and β'2-β'3 connecting loop. Moreover, projection of eigenvector2 does not show any major fluctuation in the structure **(Figure: - 4.3.15, 4.3.17 and 4.3.18)**. Projection of eigenvector 1 on BARD1 Arg658Cys mutant structure shows fluctuation due to the linker region between α'1-β'1 and α'2-β'1. Eigenvector 2 on residues for Arg658Cys shows insignificant fluctuation in the structure, which may be due to the superficial location of mutation in the protein structure **(Figure: - 4.3.16, 4.3.17 and 4.3.18)**.

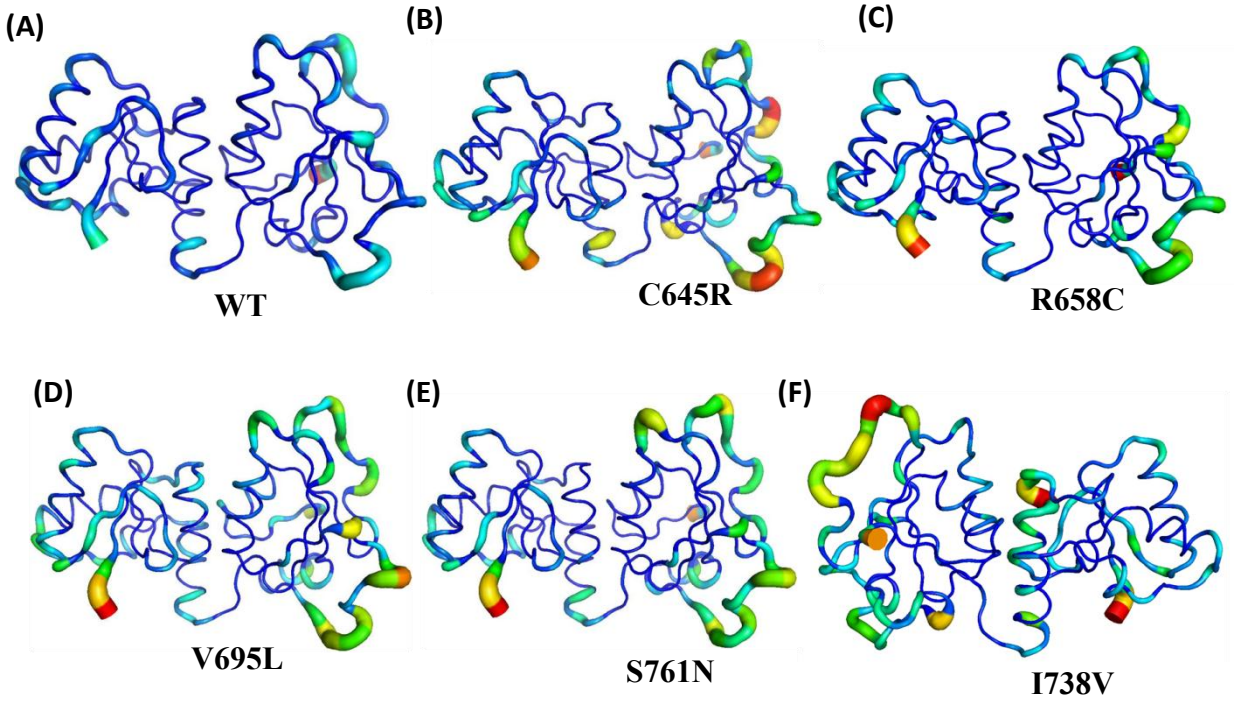


Figure 4.18: - Comparative representation of RMSF structure of *wild-type* and different mutants of BARD1. Red show regions of highest fluctuation and blue regions indicate areas of minimum fluctuations.

Projection of eigenvector 1 on BARD1 Ile738Val mutant structure shows fluctuation due to the linker region between α1-β2 and α'2-β'1. Eigenvector 2 on residues for Ile738Val shows overall fluctuation in the structure, which may be due to the destabilization of the

hydrophobic core of the protein. Increase in the fluctuation in the eigenvector 2 at residues at α1, β2 is evident when eigenvector 1 residual projection, which indicates the concerted nature of motion described by these two eigenvectors **(Figure: - 4.3.16, 4.3.17 and 4.3.18)**.

4.3.6 Comparative of Cross-correlation of PCA for BARD1BRCT *wild-type* and Mutants:-

To understand the effect of mutation on protein dynamics, residual cross-correlation analysis for *wild-type* and mutants protein were performed. The trajectory of *wild-type* and mutants were transformed in to covariance matrix, and cross-correlation for PCA was generated to analyze the dynamics of the *wild-type* and mutants structures. Cross-correlation for PCA heat map for *wild-type* indicates that N-terminal BRCT domain shows high positive correlation but C-terminal BRCT residue shows residual correlation within the domain **(Figure: - 4.3.19 A)**.

BARD1 Cys645Arg, Arg658Cys, Val695Leu, Ile738Val and Ser761Asn mutants show similar pattern as they have poor positive correlation within the C-terminal residue. Moreover, Ile738Val has lost the positive correlation within N-terminal BRCT as well C-terminal BRCT residues, which may be due to loss of stabilizing contacts within the structure **(Figure: - 4.3.19)**.

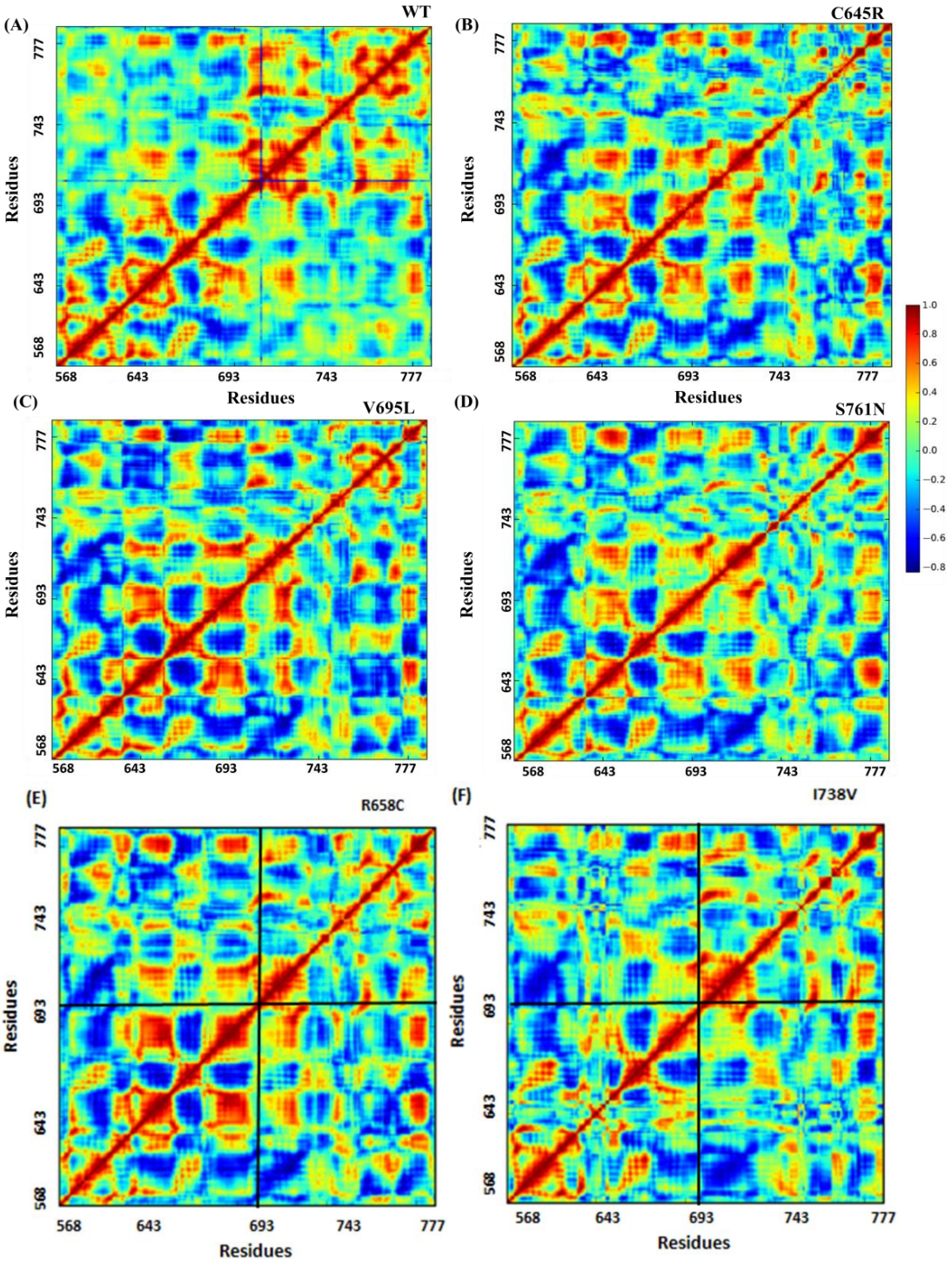


Figure 4.3.19: - Comparative cross-correlation for PCA of **(A)** *wild-type*, **(B)** Cys645Arg, **(C)** Arg 658 Cys, **(D)** Val695Leu, **(E)** Ser761Asn and **(F)** Ile738Val mutant.

4.3.7 Intra-molecular interaction analysis:-

To study, the effect of mutation on intra-molecular interactions minimum energy structure was calculated using the RMSD and R_g after free energy landscape preparation (**Figure: - 4.3.20**). Comparative residual interaction between *wild-type* and Cys645Arg mutant structure indicates that in the *wild-type* structure Cys645 is involved in non-hydrogen bonding interaction with Leu590 and Leu570, but in the mutant structure residue the Arg645 formed hydrogen with Leu592 carbonyl group and non-hydrogen bonding interaction with Leu570 (**Figure:- 4.3.21**).

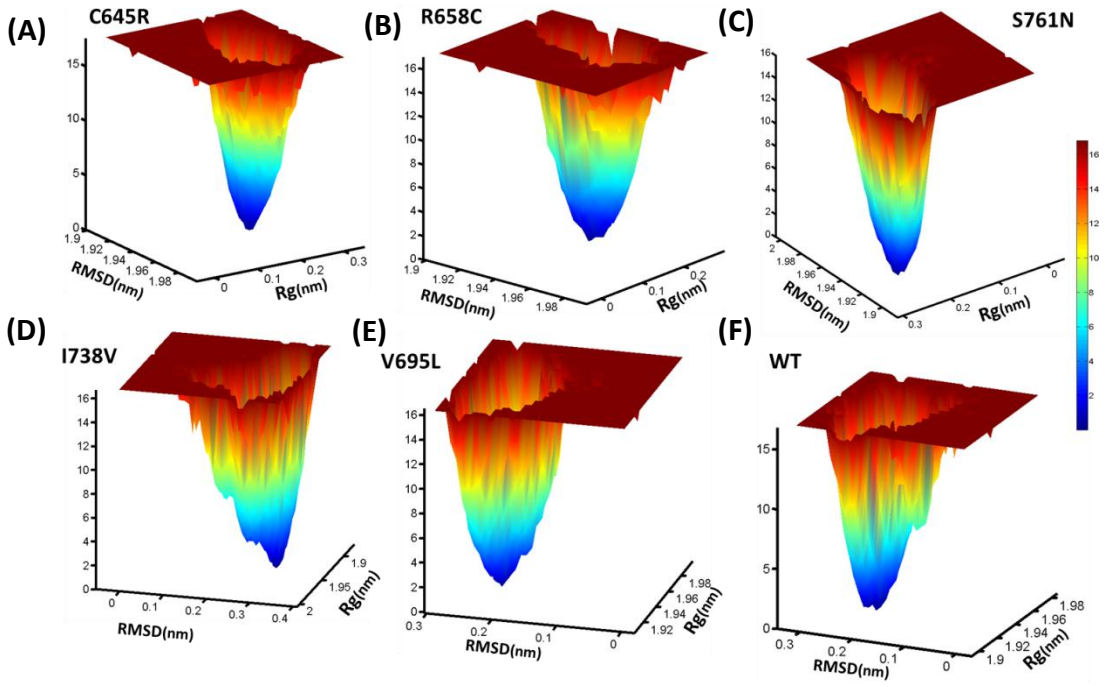


Figure 4.3.20: - Free-energy calculation for the structure BARD1BRCT *wild-type* and mutant structures.

Similarly intra-molecular interactions environment in *wild-type* Val695 structure illustrates hydrogen-bonding interaction with Gly698, Thr696, Ile692, and Leu691. Furthermore, BARD1 Val695 is also found involved in the non-covalent interaction with Phe577, Phe763, Gly698, Lys693, Ala697. However, Leu695 in mutant structure formed hydrogen bonding interactions with Gly699, Gly700, and Leu691 and has lost one hydrogen bond. However, similar kind of non-covalent bonding pattern in *wild-type* and mutant were observed **(Figure: - 4.3.21)**.

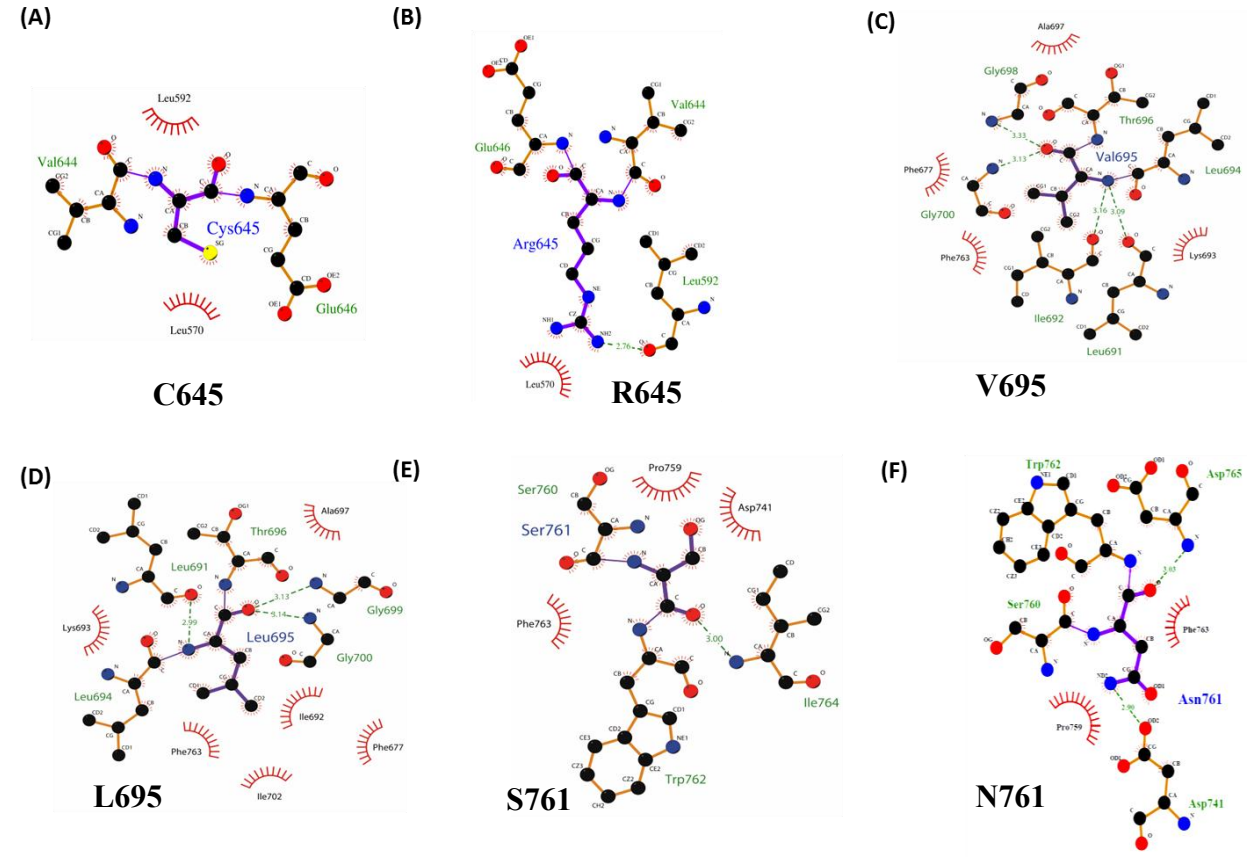


Figure 4.3.21: - Intra-molecular interaction analysis for the mutant structures **(A, B)** Cys645Arg, **(C, D)** Val695Leu and **(E, F)** Ser761Asn.

In BARD1 *wild-type* structure Ser at 761 position establishes contact with very few residues in its vicinity. It forms hydrogen bond with Ile764 and non-covalent interactions with Pro759, Asp741, and Phe763. But BARD1 Ser761Asn, hydrogen bond with Asp741 and non-covalent interactions with Pro759 and Phe763 were observed.

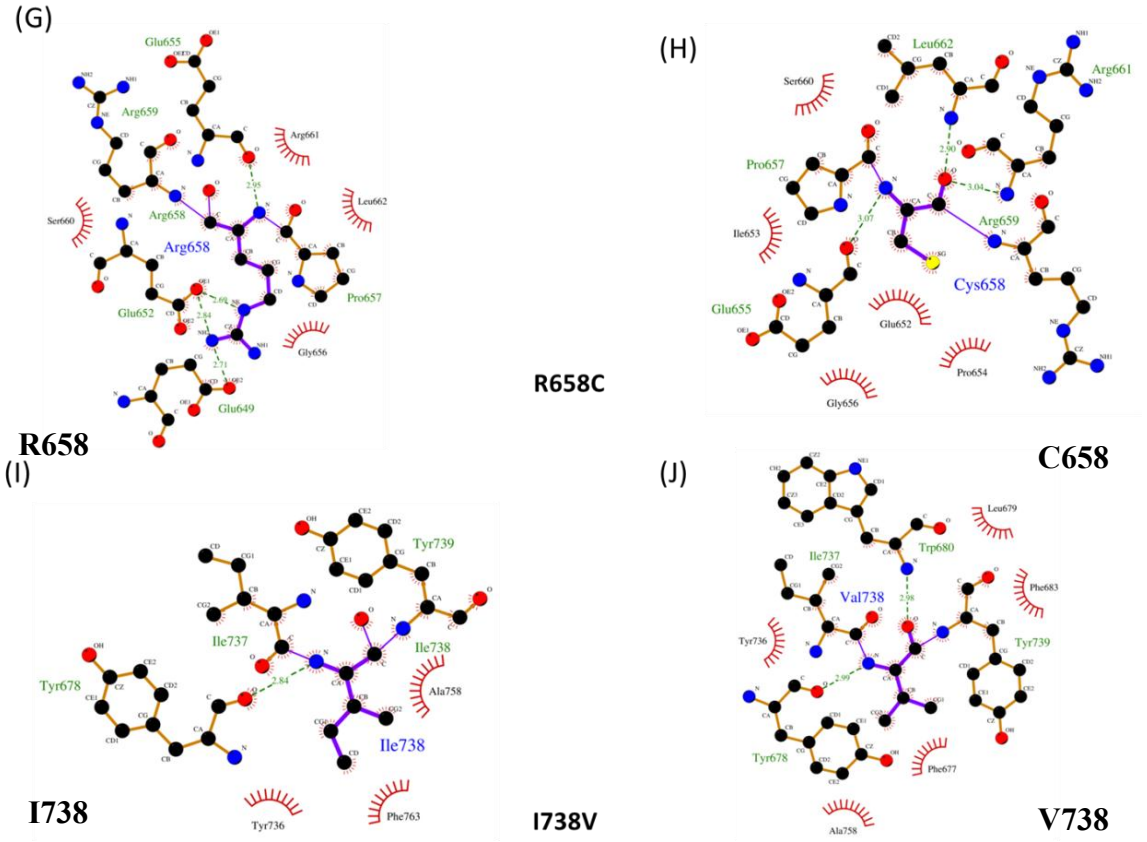


Figure 4.3.22: - Comparative intra-molecular interaction analysis for the mutant structures **(G, H)** BARD1BRCT Arg658Cys and **(I, J)** Ile738Val.

BARD1 *wild-type* Arg658 forms three hydrogen bonds with Glu649, Glu655, Glu652 and non-hydrogen bonding interactions with Ser 660, Gly656 and Leu662. On the other hand, in mutant protein structure Cys658 shows more hydrophobic interactions in the local milieu with Ser660, Gly656, Ile653, Pro654, Glu652, and hydrogen bonding interactions with Arg661, Glu655 and Leu661 **(Figure:- 4.3.22)**. Comparative intra-molecular interactions of BARD1

Ile738 in *wild-type* and BARD1 Ile738Val indicates that Ile738 shows only hydrogen bonding interaction with Tyr678 and non-hydrogen bonding interactions with Tyr736, Ala758, Phe763 and Tyr736. As compared to Ile738 mutant Val738 forms two hydrogen bond with Trp680, Tyr678 and non-hydrogen bonding interactions with Tyr736, Ala758, Phe677, Phe683 and leu679 showing additional interactions than *wild-type* **(Figure:- 4.3.22)**. The additional interactions established in mutant may be due to the void space and conformational entropy created for neighbouring residues side chain, which leads to infiltration of side chains of neighbouring amino acid near the valine establishing new contacts.

4.4 Conclusions:-

Biophysical characterization of BARD1BRCT *wild-type* and mutants shows that mutations do not have drastic effect on the monomeric property and secondary structures of the proteins, however, local structural changes cannot be denied. Thermal and chemical denaturation using CD spectroscopy indicated that except BARD1 Cys645Arg mutant, other reported BARD1BRCT mutants Arg658Cys, Val695Leu, Ile738Val and Ser761Asn have lost thermodynamic stability. Molecular dynamics simulation studies with *wild-type* and mutant structure show significant loss in intra-molecular hydrogen bonding in mutants than the *wild-type*. PCA shows BARD1BRCT Ile738Val and Val695Leu mutants have lost compactness than the BARD1 *wild-type* and BARD1 Cys645Arg, Arg658Cys, and Ser761Asn mutant protein structures. Intra-molecular interactions analysis shows that BARD1BRCT Ser761Asn mutation leads to loss of hydrogen bonding with Ile764 and Ser761, which is an important residue suspected to interact with the phosphopeptide P_s-x-x-F. Hence, it may lead to loss of hydrophobic pocket architecture, which can inhibit phosphopeptide docking pocket for DNA damage repair proteins.

Chapter

5

Molecular Basis of Interactions of

Platinum Drugs with

BRCA1-BARD1

Ring domain

Complex

5.1 Introduction:-

BRCA1 and BARD1 have wide role in tumor suppression, as both the genes are frequently found mutated in breast and ovarian cancer [3, 27]. BRCA1 is a, 1863 amino acids large protein comprising of N-terminal RING domain, central DNA binding domain, transactivation domain and the C-terminal phosphopeptide interacting BRCT domain. BARD1 comprises of 777 amino acids shows structural homology in domain organization with BRCA1. BARD1 and BRCA1 heterodimerizes *in-vitro* by RING-RING domain interactions at the N-terminus [21, 69, 70], and heterodimer displays E3 ubiquitin ligase activity which have a wide role in tumor suppression [59, 76, 85]. The BRCA1–BARD1 E3 ubiquitin ligase activity is down-regulated by the CDK2–cyclin A1/E1 and CDK2–cyclin B1 complex in a phosphorylation dependent manner [97].

BRCA1-BARD1 heterodimer plays an important role in cell-cycle regulation, DNA repair and apoptosis [13, 94, 96]. BRCA1–BARD1 ubiquitin ligase mediates ubiquitination and degradation of RNA Polymerase II [13, 26], cell-cycle arrest, γ-tubulin ubiquitination, control of centrosome duplication [7, 74, 75, 268], H2A/H2AX ubiquitination, epigenetic control [69, 76], and NPM stabilization via ubiquitination [77]. In the G2/M cell- cycle check-point, γ-tubulin is ubiquitinated at Lys48 as well as Lys344 residue [74] by BARD1-BRCA1 complex [74, 269, 270]. BRCA1-BARD1 complex is present as heterodimer, at the nuclear foci during the S-phage of the cell-cycle [271]. BRCA1-BARD1 E3 ubiquitin ligase activity also regulates increased expression levels of NPM (nucleophosmin or B23), which controls the centrosomal amplifications [7]. Protein-Protein interaction of BRCA1-BARD1 with MSH2-MSH6 complex is crucial to carry out mismatch repair and double stranded DNA damage repair [272]. BRCA1 is a well-established component of RNA polymerase holo-enzyme complex [72], and is a transcriptional co-activator of both P53 and p21WAF1/CIP1 [16, 273, 274].

BARD1-BRCA1 complex has unique property of ubiquitinating the target proteins at Lys6, position which is quite irregular. This complex also undergo auto-ubiquitination at Lys48 which enhances their E3 ubiquitin ligase activity by 20 folds [21, 69, 76, 275]. Reported cancer predisposing mutations of BRCA1 Cys61Gly and Cys64Gly, are identified in BRCA1 RING domain, which abolishes the E3 ubiquitin-ligase activity and tumor suppressor function [85, 237, 276] by interrupting the BRCA1–BARD1 interaction *in- vitro* [277].

Recent developments in cancer therapy encompass targeting vital cell survival pathways involved in maintenance of genomic integrity [1, 278]. In this approach these pathways are intentionally deregulated to induce genetic instability and subsequent cell death. The dysfunctioning of E3 ubiquitin ligase activity of BRCA1 can be a potential therapeutic target [119, 120]. Recent reports have suggested that platinum compounds like transplatin, cisplatin, oxaloplatin, and carboplatin inhibit tumor suppressor activity of BRCA1-BARD1 E3-ubiquitin ligase complex, thereby making tumor cells more sensitive to chemotherapeutic drugs [121]. The binding of anticancer drug cisplatin has been shown to affect the Apo-conformation of the BRCA1 RING domain. Mass spectrometric analysis has also revealed that cisplatin covalently binds to BRCA1 RING finger domain at His-117 position forming inter-molecular as well as intra-molecular adducts [121].

Hence, there is a possibility that platinum compounds or metal-based drugs, which target BRCA1 E3 ubiquitin ligase activity, can be significantly helpful in the combinatorial therapy and improving the efficacy of anti-breast cancer drugs. To understand the atomic association of platinum compounds with BRCA1-BARD1 complex and molecular mechanism of E3 ubiquitin ligase inhibition, BRCA1-inhibitor complex structure should be explored which will open new avenues to design the small molecule inhibitors for effective treatment.

5.2 Material and Methods

A comparative structural and binding analysis between transplatin and Cis-platin, and BRCA1 Ring domain (1-303aa) was carried out. RING domain of BRCA1 was sub-cloned in pET41a vector (Amersham Pharmacia). BARD1 Ring domain (16-119aa) was sub-cloned into the bacterial expression vector pGEX-kT (Amersham Pharmacia) using one step cloning method. Oligonucleotide primers complementary to the template sequence used in the PCR reactions are listed below.

5.2.1 Gene cloning

The different genes of interest were amplified using following PCR condition: 95°C denaturation (5 minutes), 95°C denaturation (45 seconds), and annealing 65°C for 35 seconds, extension 72°C at 0.5kb/min, final extension 72°C for 10 minutes, and 25 cycles. The PCR amplified product was digested with EcoR1 and BamH1 and ligated in the pGEX-kT and pET41a vector (a kind gift from John Ladias-HMS, Boston US). The ligation mixture was transformed into *E.coli* DH5α cells. Colony screening were performed for BARD1(16-119aa) and BRCA1(1-303aa), and positive clone were selected after the insert release from the plasmid using EcoR1 and BamH1 restriction digestion further confirmed by DNA sequencing. Positive clones with correct DNA sequence were used for protein expression and purification.

5.2.2 PCR Primers for BARD1 (16-119aa)

Forward primer: 5'-GTCGGATCCGAGAACCTGTACTTTCAGGGTGGTCCTCTTGTACTTATAGG- **3'**

Reverse primer 5'-GTCGAATTCCTATTAGCTGTCAAGAGGAAGC-**3'**

Primer details BRCA1 (1-303aa)

Forward primer: 5'-GTCGGATCCGAGAACCTGTACTTTCAGGGTATGGATTCTGCTCTTCGCG- **3'**

Reverse primer: 5'-GTCGAATTCGTCGACCTATTATTCAGCCTTTTCTACATT-**3'**

BARD1 (16-119aa) and BRCA1 (1-303aa) RING domains were expressed in *Rosetta 2(DE3)* bacterial strain. For protein expression, 50 ng/µl of plasmid construct was transformed into *Rosetta 2(DE3)* cells and grown on LB agar plate containing ampicillin (100 µg/ml). Detailed protocol is described below.

5.2.3 Protocol for protein expression:-

1. **Inoculation:** Picked a single freshly transformed colony from antibiotic resistant LB agar plate and inoculated it in 100 ml LB broth preinoculum containing 100 µg/ml of ampicillin. Grown the culture at 37^{o}C overnight in shaking condition.
2. **Dilution:** Added 10 ml of preinoculum to 1000 ml (1:100) of autoclaved LB broth containing 100 µg/ml of ampicillin. Grown medium on shaker incubator at 37^{o}C until it has reached mid-log phase i.e. A_{600} between 0.6-0.8.
3. **Induction:** Culture flasks were cooled and added 400 µl IPTG (0.4mM IPTG final concentration) to induce the culture, and incubated on shaker incubator at 24^{o}C for 16 hours.
4. **Harvesting:** The culture was centrifuged for 10 minutes at 6000 rpm at 4^{o}C. The pellet was resuspended in a small volume of supernatant and centrifuged for 15 minutes at 5000 rpm, 4^{o}C.
5. **Storage:** The bacterial pellets obtained were stored at -80^{o}C for further use.

Proteins were purified by affinity chromatography followed by FPLC.

5.2.4 Protocol for purification of BARD1 (16-119aa):-

Purification buffer*:* 50 mM HEPES (pH 7.5), 300 mM NaCl, 2mM $ZnCl_2$, 5% glycerol, pH 7.5 (buffer A)

FPLC buffer*:* 50 mM HEPES (pH 7.5), 300 mM NaCl, 1 mM $ZnCl_2$ (buffer B)

Re-suspension: Re-suspended the induced pellet of BARD1 (16-119aa) in 40 ml of (buffer A); supplemented with 200 mM PMSF and 20µl of protease inhibitor.

1. **Ultra sonication**: Transferred the resuspended cell pellet into centrifuge tube and sonicated at 70-pulse rate and 50 power with 1 minutes of duty cycle. Repeated the cycle 5 times with 1 min break.
2. **Centrifugation:** After sonication, the suspension was subjected to centrifugation at 18000 rpm for 45 minutes at 4^oC to obtain cleared lysate. Collected the supernatant and discarded the cell debris.
3. **Binding**: The soluble protein fractions obtained was brought at 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 bind specifically to the proteins having GST (Glutathione-S-transferase) tag. ~10-µl of resin was loaded onto SDS page to check the protein bound on the resin.
4. **Washing**: After binding, the protein bound column was washed with 5-column wash buffer to remove any non-specific protein.
5. **Cleavage**: Added 400 µl (20 units) of TEV protease, 40 µl of protease inhibitor cocktail and 100 µl of PMSF in 20 ml of buffer A and completed the cleavage step in 3 hours by passing

the solution through column at interval of 1 hour. ~10 µl of beads was loaded on to SDS-PAGE to detect the cleavage of protein.

6. **Elution**: After TEV cleavage, the protein was eluted with 30 ml of buffer A.

7. **Equilibration of Ni-NTA resin**: Given 2 column washes with double distilled water and then provided 5 to 6 column washes with wash Buffer.

8. **Metal Ion Chelate Affinity Chromatography**: After calibration of Ni-NTA resin, pass the eluted fractions to remove His-tagged TEV protease.

9. **Concentrating the protein:** Transferred the eluted protein in a 10 KDa pre-equilibrated centricon, and concentrate the protein up to 2 ml by centrifuging at 4500 rpm for 10 minutes at 4^{o}C. The concentration of protein was measured on Nanodrop spectrophotometer at λ=280 nm**.**

10. **Gel filtration:** Injected 2 ml of concentrated protein in AKTA- FPLC on Sperdex-75 column

11. **Fraction collection**: FPLC fractions were collected in 1.7 ml microfuge tube at its elution volume as per gel filtration spectra profile of the sample.

12. **Loading on SDS-PAGE 12% gel**: Loaded 20 µl of FPLC fractions on SDS-PAGE gel. The SDS-PAGE gel was stained with coomassie dye, and then destained it to visualize the protein of interest.

13. **Concentration of protein**: The fractions, which showed purified protein was concentrated as per the experiments requirement.

Protocol for purification of BRCA1 (1-303aa) and BARD1 (16-119aa) region:-

The BRCA1 (1-303aa) and BARD1 (16-119aa) were cloned and expressed in Rosetta 2(DE3) cells and purified using the similar protocol described in material and method section. Induction with IPTG (0.1 to 0.4 mM) was performed at 20^{o}C for 22 hrs. No expression of BRCA1 (1-303aa) soluble protein was observed at wide induction temperature range. The complete details of expression and purification have been provided in section 3 material and methods.

5.3 Results and discussion

5.3.1 Cloning, expression and purification of BARD1(16-119aa) and BRCA1(1-303aa) RING domains:- Positive clones were selected after confirmation of double digestion with the EcoR1 and BamH1 restriction enzymes, leading to insert release of appropriate size (**Figure:-5.3.1**). DNA Sequencing results confirmed the presence of ligated gene of interest in the vector with desired frame of codon sequence.

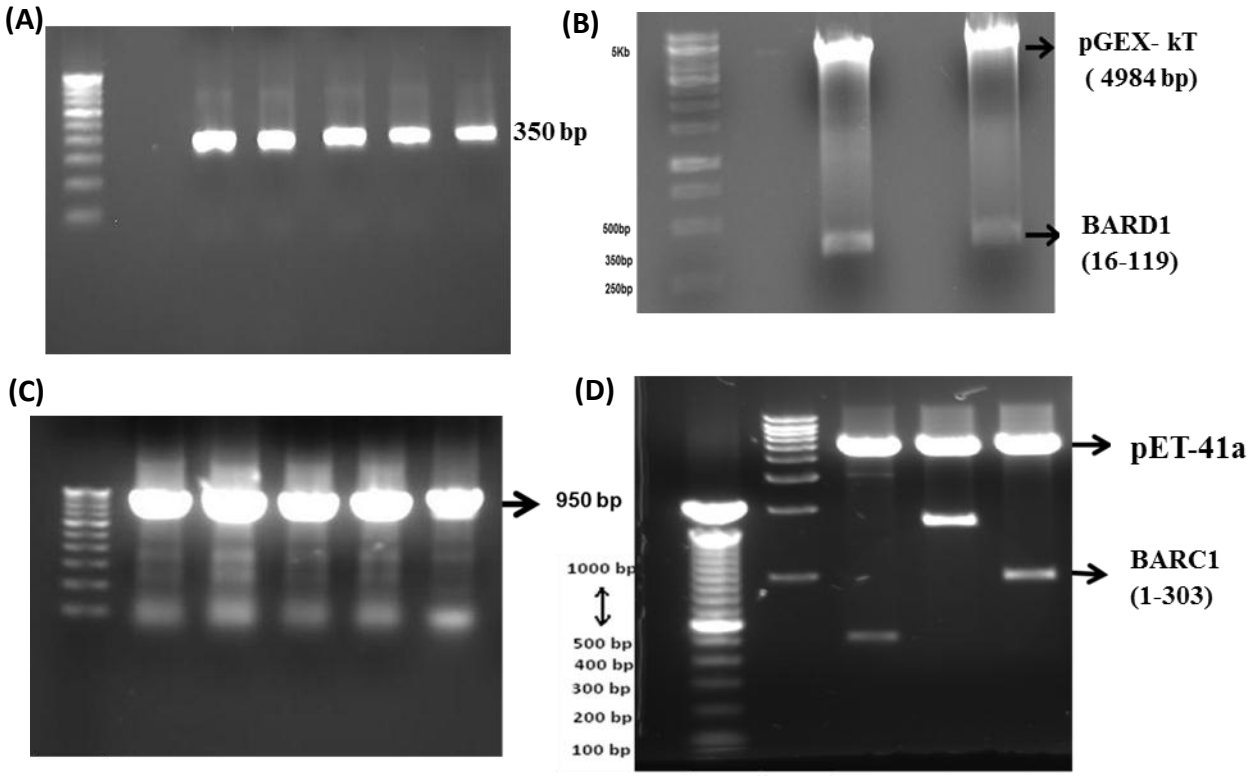


Figure 5.3.1: - Cloning of BRCA1 and BARD1 Ring domain region. **(A)** PCR profile of BARD1 Ring domain and **(B)** clones showing insert release. **(C)** PCR profile of BRCA1 Ring domain and **(D)** clones showing insert release.

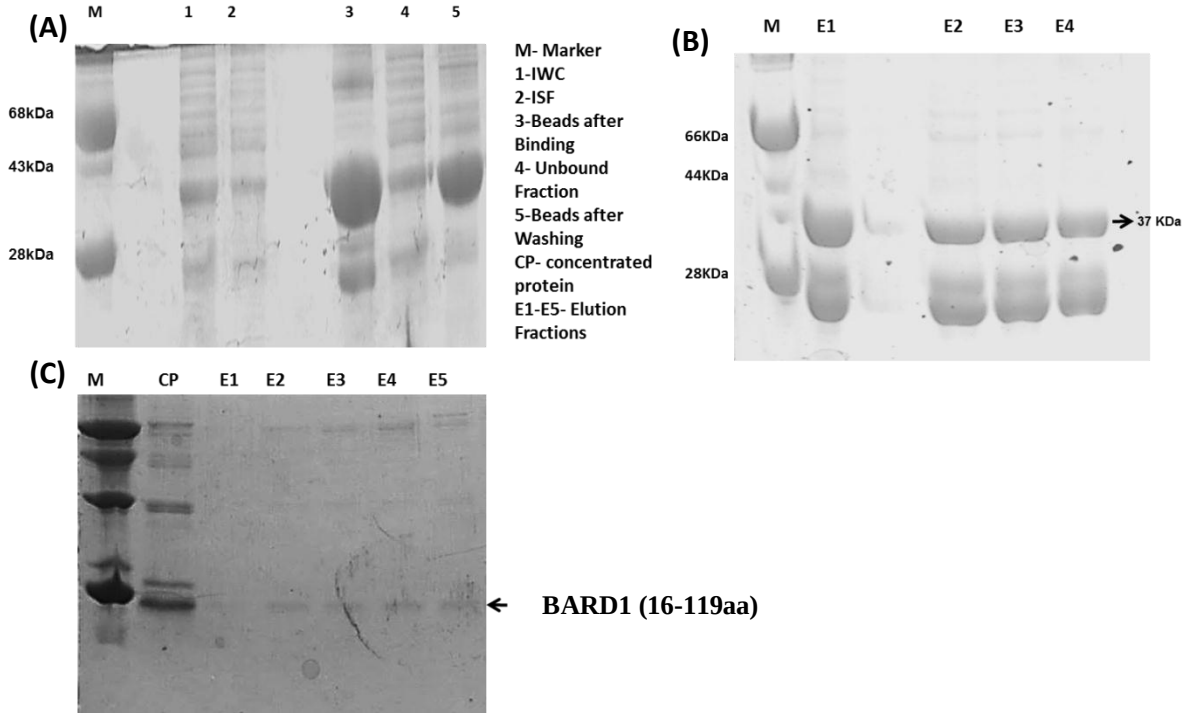


Figure 5.3.2: - Expression of BARD1 Ring domain region. **(A** and **B)** Expression profile of BARD1 Ring domain and **(C)** FPLC purification profile.

BARD1 (16-119aa) was expressed and purified as a GST fusion protein **(Figure: - 5.3.2)**. The Fusion protein was further treated with TEV protease to get rid of GST tag. The TEV was removed using Ni-NTA affinity chromatography as it contains 6His tag.

5.3.2 *In-silico* approach to unravel Structural insight into BRCA1-platinum drug complex:-

In order to analyze the binding of platinum drug to BRCA1 functional motif, RING domain (1-125aa) was modeled using Robetta server [37] using one ensemble of NMR solution structure (PDB ID: 1JM7) as the template [279].

Best model was selected on the basis of stereo chemistry and acceptable Ramachandran plot. The model was validated using server "**SAVES**" (Metaserver for analyzing and validating protein structures, (http://nihserver.mbi.ucla.edu/SAVES/). The server SAVES mainly comprises of five programs, Procheck, What_check, Errat, Verify_3D, and Prove. Platinum drugs were docked onto the structure of BRCA1(1-125aa) using discovery studio C-DOCK module [280]. Ligplot was used to analyze the weak intermolecular interactions [218].

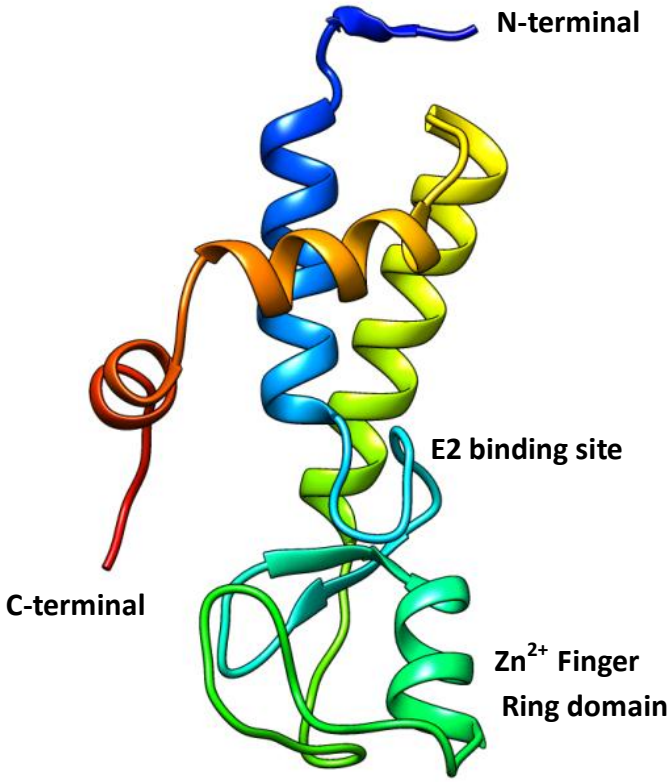


Figure 5.3.3: - Model structure BRCA1 Ring domain (1-125aa)

Modeled structure of BRCA1 shows a total of 5 α-helices and one double stranded β-sheet as shown in **(Figure:- 5.3.3)**. Two of these α-helices perturb in a shape of a finger from a ring shaped base with two Zinc binding sites within it.

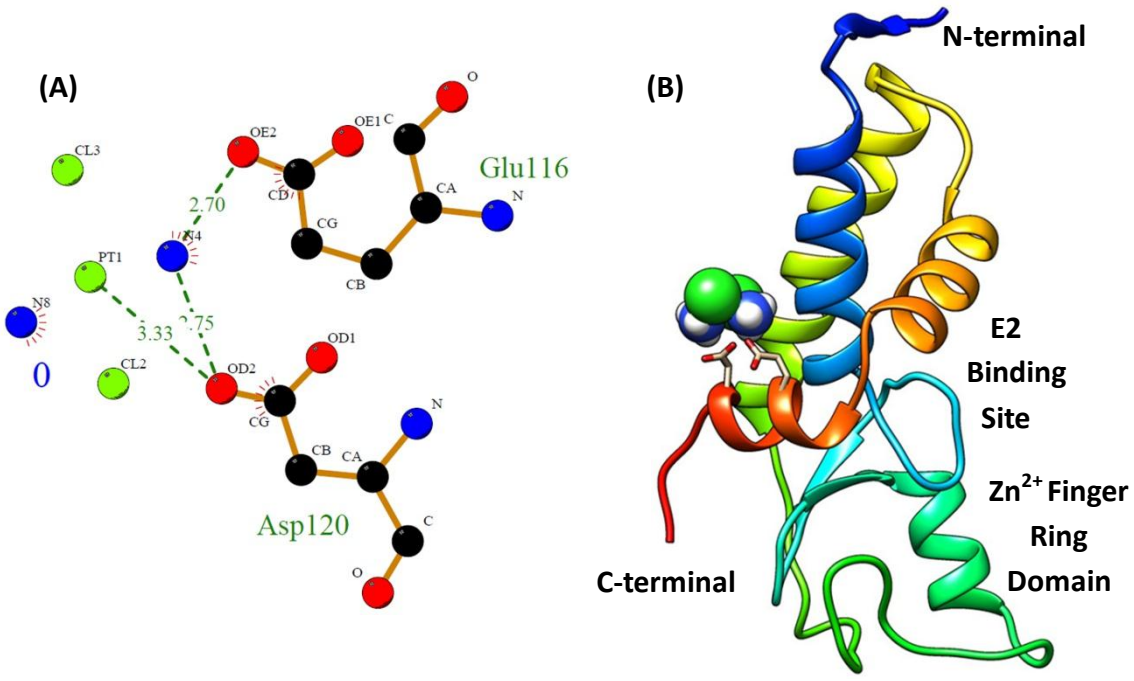


Figure 5.3.4:- **(A)** Hydrogen bonding and hydrophobic interactions between BRCA1 and cisplatin **(B)** complex model structure of cisplatin and BRCA1. The dotted lines indicate H-bond and brush lines indicate hydrophobic interactions.

Helix and strand intertwined themselves in such a way that they acquire a finger and ring like structure (**Figure: - 5.3.3)**. The 4-cysteine residues and one histidine residue in the loop region provide close interacting space for binding two Zn^{2+} ions. **Figure: - 5.3.4** and **5.3.5** shows cisplatin and transplatin molecules docked to the BRCA1 ring domain. It can be seen that these drug are binding in the small α-helical region of BRCA1, which is very close to the E2 binding site. Cisplatin shows hydrogen-bonding interactions with Glu116 and Asp120 **(Figure: - 5.3.4)**.

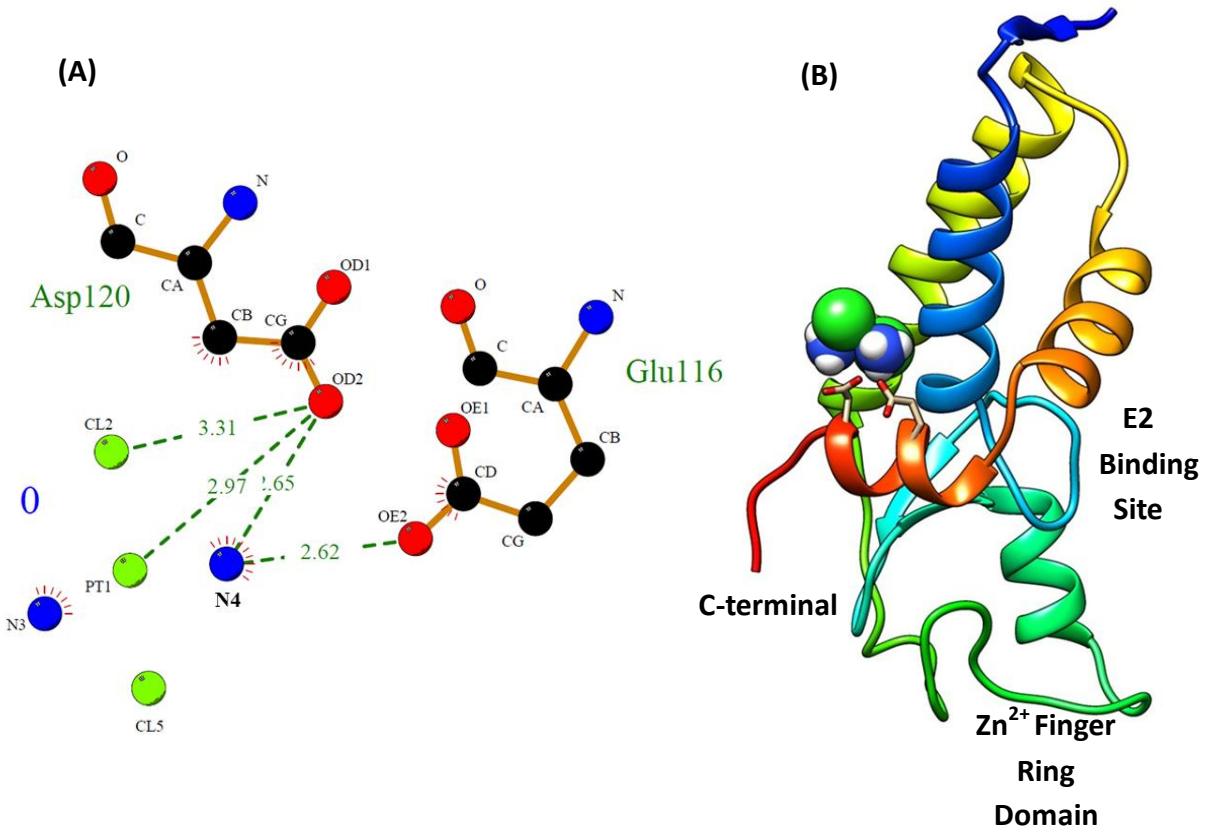


Figure 5.3.5: - **(A)** Inter-molecular interaction analysis and **(B)** complex structure of transplatin and BRCA1. The dotted lines indicate H-bond and brush lines indicate hydrophobic interactions.

The NH_3 of cisplatin at Cis position interacts with BRCA1 Glu116 and Asp120, and platinum shows interaction with BRCA1 Asp120. On the other hand, transplatin at binding site shows that the NH_3 group forms two hydrogen bonds with BRCA1 Glu116, Asp120 **(Figure:- 5.3.5)**. However, an extra hydrogen bond is formed in case of transplatin Cl^- and Asp 120 of BRCA1, which is not the case with the cisplatin molecule **(Figure:- 5.3.5)**. There is also an increase in number of weak intermolecular interactions between transplatin and BRCA1 ring domain. Therefore, Trans conformation provides the favorable positioning of NH3 group, which is the

basis of more favorable interactions and inhibition of E3 ubiquitin ligase activity of BRCA1 as compared to the cisplatin.

5.4 Conclusion:-

BRCA1 and BARD1 RING domain complex has E3 ubiquitin ligase activity, which is inhibited by platinum drugs like transplatin and cisplatin. A BRCA1-drug interactions study with modeled structure of BRCA1 RING domain and platinum molecules predicts that platinum drugs interact with BRCA1 through the RING domain. The Trans conformation of NH_3 group in transplatin molecule forms more number of hydrogen bonds as compared to cisplatin. Hence, transplatin inhibits BRCA1-BARD1 RING domain E3 ubiquitin ligase activity at lower concentration than cisplatin. This information will be very helpful in designing the small molecules inhibitor with higher affinity to BRCA1.

Chapter

6

Summary

And

Future

Perspective

6.1 BARD1-CstF50 complex

BARD1 and BRCA1 are well known for their tumor suppressor function. Complex of BRCA1 and BARD1 heterodimer with CstF50 has indispensable role in the pre-mRNA processing, and transcription coupled repair. However, BARD1 Glu564His mutation leads to the reduced binding to CstF50. Therefore, our aim was to delineate the structural details of the complex to understand the intermolecular interactions and effect of mutations on the protein-protein interactions. BARD1 ARD-BRCT *wild-type* and Glu564His mutant have been purified, and attempts were made to crystallize the proteins. It has been observed that the flexible linker region connecting ARD and BRCT domain keeps the ARD domain dynamic with respect to the BRCT domain. Furthermore, we have characterized the *wild-type* and mutant protein using different biophysical tools. We tried to purify the CstF50, which interacts at the linker region of the BARD1 ARD-BRCT domain. Insolubility of the CstF50, leads us to model BARD1 ARD-BARD1BRCT and CstF50 using Robetta server. Models obtained were validated and the extent of domain motion in ARD-BRCT domain was characterized using NMA and molecular dynamics simulation. We found similarity in results from NMA and molecular dynamics simulation studies. On the basis of these results, we conclude that in BARD1 ARD-BRCT domain, ARD domain has no specific orientation with respect to the BRCT domain. Molecular docking was performed to unravel the interacting residues between BARD1 ARD-BRCT and CstF50 and structural rigidity of the complex. The docked complex was simulated further to study the complex stability. RMSD, R_g and RMSF profile indicated that the binding has stabilized the complex. Further, protein-protein interactions revealed from docking studies indicate that BARD1 Glu564 plays an important role in binding to CstF50. Loss of this interaction with CstF50 due to Glu564His mutant

rationalizes abrogation of productive complex with cstF50. Hence, crystallographic studies with the complex will be helpful in understanding the precise mechanism of interaction, residual details and better understanding of the disease predisposition.

6.2 BARD1BRCT and Cancer predisposing mutations

BRCTs are conserved domains present in different proteins involved in the DNA damage repair process. BARD1BRCT has been found to be involved in the early recruitment of BRCA1 at the DNA damage site and involved in recruiting DNA damage repair proteins in a phosphodependent manner. Cancer predisposing mutations are reported in the BARD1BRCT domain in different populations across the world. In the present study, we have decided to characterize the deleterious effects of the mutations on the protein structure using *in-vitro*, biophysical techniques and computational tools like MD simulation. Comparative thermal and chemical denaturation studies with *wild-type* and mutant proteins indicated that besides BARD1 Cys645Arg mutant, other mutants like, Val695Leu, Ser761Asn, Arg658Cys and Ile738Val show loss of thermal stability. These results are further supported by less number of the H-bonds observed in MD simulation of the mutant proteins. Mutations Val695Leu and Ile738Val are present in the hydrophobic core of the BARD1BRCT and destabilize the core, which is reflected in the increase in the volume, and solvent accessible surface area, and trace of covariance matrix compared to the *wild-type* protein. According to the reports, BARD1 Cys645 and Val695 are involved in Poly-A ribose binding, and early recruitment of BRCA1 to the DNA damage site. Mutations such as Cys645Arg and Val695Leu lead to loss of Poly-A ribose binding, and consequently failed BRCA1 recruitment to DNA damage site. Therefore, structure determination of BARD1BRCT with Poly-A ribose will be helpful in understanding the molecular mechanism of Poly-A ribose binding to BARD1BRCT and will further explore

the possibility of structure based inhibitor design of pharmacological molecules for therapeutic application that can compensate the effect of such mutation.

6.3 BRCA1-BARD1 E3 ubiquitin ligase complex

Several pre-clinical, clinical and animal model studies have established that modulation of BRCA1 function has proven to be encouraging therapeutic answer for breast and ovarian cancer treatment. One such study indicates that platinum drugs like Transplatin and Cisplatin inhibit BRCA1-BARD1 E3 ubiquitin ligase activity. The mode of action of platinum drug binding to BRCA1 ring domain was poorly understood. Therefore, to unravel the mechanism of the drug binding, we have modeled the Ring domain of the BRCA1 and performed the binding with the platinum drugs. The validated models were docked with platinum drugs to understand the detailed mechanism of atomic interactions. In complex structure, transplatin shows more effective binding than the cisplatin molecule due to the Trans orientation of the NH_3 group. So, structural biology based approach will be more helpful in understanding the mode of inhibition more precisely and development of effective platinum drug derivatives.

6.4 Future perspective

The structural and functional analysis of BRCA1-BARD1-CstF50 complex was explored to understand its role in DNA damage repair mechanism. Non-homologous expression system like bacterial system often poses a great hurdle during protein expression and purification. CstF50 is a member of WD40 repeat domain protein, and was found insoluble in the bacterial system. So a better approach could be the expression of protein in eukaryotic expression system such as yeast or insect cells lines in order to get a well folded protein with all desired post-translation modification. Such a preparation will be more suitable for crystallization.

Inter-domain motion within a protein can be overcome by co-crystallizing it with its binding partner. *In-silico* approach such as docking can provide a possible binding interface, but stereochemical structure determination using x-ray crystallography will provide a rational approach of precisely understanding the mechanism of interaction. Often mutations in one residue in the same class (Ile to Val and Val to Leu) cannot have drastic effect on the secondary structure of the protein and probing such substitution with less sensitive techniques like CD spectroscopy are an alternative approach. Therefore, to delineate such changes techniques with high sensitivity like x-ray crystallography must be applied to unravel the small critical changes induced due to the act of mutation and disease predisposition.

References

References:-

1. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-674.
2. Alberg, A.J., A.P. Lam, and K.J. Helzlsouer, *Epidemiology, prevention, and early detection of breast cancer.* Current opinion in oncology, 1999. **11**(6): p. 435.
3. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.* Science, 1994. **266**(5182): p. 66-71.
4. Venkitaraman, A.R., *Cancer susceptibility and the functions of BRCA1 and BRCA2.* Cell, 2002. **108**(2): p. 171-82.
5. Lane, T.F., *BRCA1 and transcription.* Cancer Biol Ther, 2004. **3**(6): p. 528-33.
6. Kleiman, F.E. and J.L. Manley, *The BARD1-CstF-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression.* Cell, 2001. **104**(5): p. 743-53.
7. Sankaran, S., et al., *Centrosomal microtubule nucleation activity is inhibited by BRCA1-dependent ubiquitination.* Mol Cell Biol, 2005. **25**(19): p. 8656-68.
8. Abbott, D.W., et al., *BRCA1 expression restores radiation resistance in BRCA1-defective cancer cells through enhancement of transcription-coupled DNA repair.* J Biol Chem, 1999. **274**(26): p. 18808-12.
9. Hu, Y.F., Z.L. Hao, and R. Li, *Chromatin remodeling and activation of chromosomal DNA replication by an acidic transcriptional activation domain from BRCA1.* Genes Dev, 1999. **13**(6): p. 637-42.
10. Xia, Y., et al., *Enhancement of BRCA1 E3 ubiquitin ligase activity through direct interaction with the BARD1 protein.* Journal of Biological Chemistry, 2003. **278**(7): p. 5255-5263.
11. Eakin, C.M., et al., *Estrogen receptor α is a putative substrate for the BRCA1 ubiquitin ligase.* Proceedings of the National Academy of Sciences, 2007. **104**(14): p. 5794-5799.
12. Kleiman, F.E. and J.L. Manley, *Functional interaction of BRCA1-associated BARD1 with polyadenylation factor CstF-50.* Science, 1999. **285**(5433): p. 1576-9.
13. 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.
14. Cantor, S., et al., *The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations.* Proc Natl Acad Sci U S A, 2004. **101**(8): p. 2357-62.
15. Cantor, S.B., et al., *BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function.* Cell, 2001. **105**(1): p. 149-60.
16. 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.
17. 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.
18. Li, M. and X. Yu, *Function of BRCA1 in the DNA damage response is mediated by ADP-ribosylation.* Cancer Cell, 2013. **23**(5): p. 693-704.
19. 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.
20. Zhou, B.-B.S. and S.J. Elledge, *The DNA damage response: putting checkpoints in perspective.* Nature, 2000. **408**(6811): p. 433-439.
21. Wu-Baer, F., et al., *The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin.* J Biol Chem, 2003. **278**(37): p. 34743-6.

22. Atipairin, A., B. Canyuk, and A. Ratanaphan, *Cisplatin affects the conformation of apo form, not holo form, of BRCA1 RING finger domain and confers thermal stability.* Chemistry & biodiversity, 2010. **7**(8): p. 1949-1967.
23. Atipairin, A., B. Canyuk, and A. Ratanaphan, *The RING heterodimer BRCA1–BARD1 is a ubiquitin ligase inactivated by the platinum-based anticancer drugs.* Breast cancer research and treatment, 2011. **126**(1): p. 203-209.
24. Atipairin, A. and A. Ratanaphan, *In Vitro enhanced sensitivity to cisplatin in D67Y BRcA1 RInG Domain protein.* Breast cancer: basic and clinical research, 2011. **5**: p. 201.
25. Nazeer, F.I., et al., *p53 inhibits mRNA 3' processing through its interaction with the CstF/BARD1 complex.* Oncogene, 2011. **30**(27): p. 3073-83.
26. Starita, L.M., et al., *BRCA1/BARD1 ubiquitinate phosphorylated RNA polymerase II.* J Biol Chem, 2005. **280**(26): p. 24498-505.
27. 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.
28. Yu, X., et al., *The BRCT domain is a phospho-protein binding domain.* Science, 2003. **302**(5645): p. 639-642.
29. Glover, J.N., R.S. Williams, and M.S. Lee, *Interactions between BRCT repeats and phosphoproteins: tangled up in two.* Trends Biochem Sci, 2004. **29**(11): p. 579-85.
30. Manke, I.A., et al., *BRCT repeats as phosphopeptide-binding modules involved in protein targeting.* Science, 2003. **302**(5645): p. 636-9.
31. Yu, X., et al., *The BRCT domain is a phospho-protein binding domain.* Science, 2003. **302**(5645): p. 639-42.
32. De Brakeleer, S., et al., *Cancer predisposing missense and protein truncating BARD1 mutations in non-BRCA1 or BRCA2 breast cancer families.* Hum Mutat, 2010. **31**(3): p. E1175-85.
33. Ghimenti, C., et al., *Germline mutations of the BRCA1-associated ring domain (BARD1) gene in breast and breast/ovarian families negative for BRCA1 and BRCA2 alterations.* Genes Chromosomes Cancer, 2002. **33**(3): p. 235-42.
34. De Vries, S.J., M. van Dijk, and A.M. Bonvin, *The HADDOCK web server for data-driven biomolecular docking.* Nature protocols, 2010. **5**(5): p. 883-897.
35. Studio, D., *version 3.5.* Accelrys Inc.: San Diego, CA, USA, 2012.
36. Colgan, D.F. and J.L. Manley, *Mechanism and regulation of mRNA polyadenylation.* Genes Dev, 1997. **11**(21): p. 2755-66.
37. 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.
38. Birrane, G., et al., *Crystal structure of the BARD1 BRCT domains.* Biochemistry, 2007. **46**(26): p. 7706-12.
39. Fox, D., et al., *Crystal structure of the BARD1 ankyrin repeat domain and its functional consequences.* Journal of Biological Chemistry, 2008. **283**(30): p. 21179-21186.
40. Guex, N. and M.C. Peitsch, *SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling.* electrophoresis, 1997. **18**(15): p. 2714-2723.
41. Berendsen, H.J., D. van der Spoel, and R. van Drunen, *GROMACS: A message-passing parallel molecular dynamics implementation.* Computer Physics Communications, 1995. **91**(1): p. 43-56.
42. Hess, B., et al., *GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation.* Journal of chemical theory and computation, 2008. **4**(3): p. 435-447.
43. Van Der Spoel, D., et al., *GROMACS: fast, flexible, and free.* J Comput Chem, 2005. **26**(16): p. 1701-18.
44. Jackson, S.P., *The DNA-damage response: new molecular insights and new approaches to cancer therapy.* Biochem Soc Trans, 2009. **37**(Pt 3): p. 483-94.

45. Jackson, S.P. and J. Bartek, *The DNA-damage response in human biology and disease.* Nature, 2009. **461**(7267): p. 1071-8.
46. Slamon, D.J., et al., *Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer.* Science, 1989. **244**(4905): p. 707-712.
47. Aghib, D.F. and J.M. Bishop, *A 3' truncation of myc caused by chromosomal translocation in a human T-cell leukemia is tumorigenic when tested in established rat fibroblasts.* Oncogene, 1991. **6**(12): p. 2371-5.
48. Rowley, J.D., *Chromosome translocations: dangerous liaisons revisited.* Nat Rev Cancer, 2001. **1**(3): p. 245-50.
49. Vogelstein, B. and K.W. Kinzler, *Cancer genes and the pathways they control.* Nature medicine, 2004. **10**(8): p. 789-799.
50. 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.
51. Wooster, R., et al., *Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13.* Science, 1994. **265**(5181): p. 2088-2090.
52. Levy-Lahad, E. and E. Friedman, *Cancer risks among BRCA1 and BRCA2 mutation carriers.* Br J Cancer, 2007. **96**(1): p. 11-5.
53. Hall, J.M., et al., *Linkage of early-onset familial breast cancer to chromosome 17q21.* Science, 1990. **250**(4988): p. 1684-9.
54. Scully, R. and N. Puget, *BRCA1 and BRCA2 in hereditary breast cancer.* Biochimie, 2002. **84**(1): p. 95-102.
55. Welcsh, P.L. and M.C. King, *BRCA1 and BRCA2 and the genetics of breast and ovarian cancer.* Hum Mol Genet, 2001. **10**(7): p. 705-13.
56. Welcsh, P.L., K.N. Owens, and M.C. King, *Insights into the functions of BRCA1 and BRCA2.* Trends Genet, 2000. **16**(2): p. 69-74.
57. Yu, V., *Caretaker Brca1: keeping the genome in the straight and narrow.* Breast Cancer Res, 2000. **2**(2): p. 82-5.
58. Yu, X. and R. Baer, *Nuclear localization and cell cycle-specific expression of CtIP, a protein that associates with the BRCA1 tumor suppressor.* J Biol Chem, 2000. **275**(24): p. 18541-9.
59. 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.
60. Ayi, T.C., et al., *Conservation of function and primary structure in the BRCA1-associated RING domain (BARD1) protein.* Oncogene, 1998. **17**(16): p. 2143-8.
61. Gautier, F., et al., *Identification of an apoptotic cleavage product of BARD1 as an autoantigen: a potential factor in the antitumoral response mediated by apoptotic bodies.* Cancer Res, 2000. **60**(24): p. 6895-900.
62. 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.
63. Brzovic, P.S., et al., *BRCA1 RING domain cancer-predisposing mutations. Structural consequences and effects on protein-protein interactions.* J Biol Chem, 2001. **276**(44): p. 41399-406.
64. Brzovic, P.S., et al., *Structure of a BRCA1-BARD1 heterodimeric RING-RING complex.* Nat Struct Biol, 2001. **8**(10): p. 833-7.
65. Irminger-Finger, I. and C.E. Jefford, *Is there more to BARD1 than BRCA1?* Nat Rev Cancer, 2006. **6**(5): p. 382-91.
66. Wilkinson, K.D., *Roles of ubiquitinylation in proteolysis and cellular regulation.* Annu Rev Nutr, 1995. **15**: p. 161-89.
67. Hochstrasser, M., *Ubiquitin-dependent protein degradation.* Annual review of genetics, 1996. **30**(1): p. 405-439.

68. Laney, J.D. and M. Hochstrasser, *Substrate targeting in the ubiquitin system.* Cell, 1999. **97**(4): p. 427-430.
69. Mallery, D.L., C.J. Vandenberg, and K. Hiom, *Activation of the E3 ligase function of the BRCA1/BARD1 complex by polyubiquitin chains.* EMBO J, 2002. **21**(24): p. 6755-62.
70. 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.
71. Roy, R., J. Chun, and S.N. Powell, *BRCA1 and BRCA2: different roles in a common pathway of genome protection.* Nature Reviews Cancer, 2012. **12**(1): p. 68-78.
72. 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.
73. Jin, Y., et al., *Cell cycle-dependent colocalization of BARD1 and BRCA1 proteins in discrete nuclear domains.* Proc Natl Acad Sci U S A, 1997. **94**(22): p. 12075-80.
74. Starita, L.M., et al., *BRCA1-dependent ubiquitination of gamma-tubulin regulates centrosome number.* Mol Cell Biol, 2004. **24**(19): p. 8457-66.
75. Hsu, L.C., T.P. Doan, and R.L. White, *Identification of a gamma-tubulin-binding domain in BRCA1.* Cancer Res, 2001. **61**(21): p. 7713-8.
76. Chen, A., et al., *Autoubiquitination of the BRCA1*BARD1 RING ubiquitin ligase.* J Biol Chem, 2002. **277**(24): p. 22085-92.
77. Sato, K., et al., *Nucleophosmin/B23 is a candidate substrate for the BRCA1-BARD1 ubiquitin ligase.* J Biol Chem, 2004. **279**(30): p. 30919-22.
78. 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.
79. Zhang, J., et al., *Chk2 phosphorylation of BRCA1 regulates DNA double-strand break repair.* Mol Cell Biol, 2004. **24**(2): p. 708-18.
80. 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.
81. Uziel, T., et al., *Requirement of the MRN complex for ATM activation by DNA damage.* The EMBO journal, 2003. **22**(20): p. 5612-5621.
82. 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.
83. Krum, S.A., et al., *BRCA1 associates with processive RNA polymerase II.* J Biol Chem, 2003. **278**(52): p. 52012-20.
84. Nishikawa, H., et al., *BRCA1-associated protein 1 interferes with BRCA1/BARD1 RING heterodimer activity.* Cancer Res, 2009. **69**(1): p. 111-9.
85. Hashizume, R., et al., *The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation.* J Biol Chem, 2001. **276**(18): p. 14537-40.
86. 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.
87. 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.
88. 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.
89. Gundersen, G.G. and T.A. Cook, *Microtubules and signal transduction.* Current opinion in cell biology, 1999. **11**(1): p. 81-94.
90. Huen, M.S., et al., *RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly.* Cell, 2007. **131**(5): p. 901-914.

91. Wang, B., et al., *Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response.* Science, 2007. **316**(5828): p. 1194-1198.
92. Kim, H., J. Huang, and J. Chen, *CCDC98 is a BRCA1-BRCT domain–binding protein involved in the DNA damage response.* Nature structural & molecular biology, 2007. **14**(8): p. 710-715.
93. Chen, J.J., et al., *BRCA1, BRCA2, and Rad51 operate in a common DNA damage response pathway.* Cancer Res, 1999. **59**(7 Suppl): p. 1752s-1756s.
94. Westermark, U.K., et al., *BARD1 participates with BRCA1 in homology-directed repair of chromosome breaks.* Mol Cell Biol, 2003. **23**(21): p. 7926-36.
95. 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.
96. Feki, A., et al., *BARD1 induces apoptosis by catalysing phosphorylation of p53 by DNA-damage response kinase.* Oncogene, 2005. **24**(23): p. 3726-36.
97. Hayami, R., et al., *Down-regulation of BRCA1-BARD1 ubiquitin ligase by CDK2.* Cancer Res, 2005. **65**(1): p. 6-10.
98. Dechend, R., et al., *The Bcl-3 oncoprotein acts as a bridging factor between NF-kappaB/Rel and nuclear co-regulators.* Oncogene, 1999. **18**(22): p. 3316-23.
99. Palmer, S. and Y.H. Chen, *Bcl-3, a multifaceted modulator of NF-κB-mediated gene transcription.* Immunologic research, 2008. **42**(1-3): p. 210-218.
100. Morris, J.R., et al., *The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress.* Nature, 2009. **462**(7275): p. 886-90.
101. Ruffner, H., et al., *BRCA1 is phosphorylated at serine 1497 in vivo at a cyclin-dependent kinase 2 phosphorylation site.* Mol Cell Biol, 1999. **19**(7): p. 4843-54.
102. Cortez, D., et al., *Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks.* Science, 1999. **286**(5442): p. 1162-6.
103. Tibbetts, R.S., et al., *Functional interactions between BRCA1 and the checkpoint kinase ATR during genotoxic stress.* Genes Dev, 2000. **14**(23): p. 2989-3002.
104. O'Brien, K.A., et al., *Casein kinase 2 binds to and phosphorylates BRCA1.* Biochem Biophys Res Commun, 1999. **260**(3): p. 658-64.
105. Kim, H.S., et al., *DNA damage-induced BARD1 phosphorylation is critical for the inhibition of messenger RNA processing by BRCA1/BARD1 complex.* Cancer Res, 2006. **66**(9): p. 4561-5.
106. Wu-Baer, F., T. Ludwig, and R. Baer, *The UBXN1 protein associates with autoubiquitinated forms of the BRCA1 tumor suppressor and inhibits its enzymatic function.* Mol Cell Biol, 2010. **30**(11): p. 2787-98.
107. Jensen, D.E., et al., *BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression.* Oncogene, 1998. **16**(9): p. 1097-112.
108. Nijman, S.M., et al., *A genomic and functional inventory of deubiquitinating enzymes.* Cell, 2005. **123**(5): p. 773-86.
109. Nishikawa, H., et al., *BRCA1-associated protein 1 interferes with BRCA1/BARD1 RING heterodimer activity.* Cancer research, 2009. **69**(1): p. 111-119.
110. Koopmans, A.E., et al., *Clinical significance of immunohistochemistry for detection of BAP1 mutations in uveal melanoma.* Modern Pathology, 2014. **27**(10): p. 1321-1330.
111. Harbour, J.W., et al., *Frequent mutation of BAP1 in metastasizing uveal melanomas.* Science, 2010. **330**(6009): p. 1410-3.
112. Wilkinson, K.D., *Regulation of ubiquitin-dependent processes by deubiquitinating enzymes.* FASEB J, 1997. **11**(14): p. 1245-56.
113. Lu, Y., et al., *The F-box protein FBXO44 mediates BRCA1 ubiquitination and degradation.* Journal of Biological Chemistry, 2012. **287**(49): p. 41014-41022.

114. Wu, W., et al., *HERC2 is an E3 ligase that targets BRCA1 for degradation.* Cancer Res, 2010. **70**(15): p. 6384-92.
115. Puigvert, J.C., K. Sanjiv, and T. Helleday, *Targeting DNA repair, DNA metabolism and replication stress as anti-cancer strategies.* FEBS Journal, 2016. **283**(2): p. 232-245.
116. Tassone, P., et al., *Loss of BRCA1 function increases the antitumor activity of cisplatin against human breast cancer xenografts in vivo.* Cancer Biol Ther, 2009. **8**(7): p. 648-53.
117. Byrski, T., et al., *Response to neoadjuvant therapy with cisplatin in BRCA1-positive breast cancer patients.* Breast Cancer Res Treat, 2009. **115**(2): p. 359-63.
118. Price, M. and A.N. Monteiro, *Fine tuning chemotherapy to match BRCA1 status.* Biochem Pharmacol, 2010. **80**(5): p. 647-53.
119. Sun, Y., *E3 ubiquitin ligases as cancer targets and biomarkers.* Neoplasia, 2006. **8**(8): p. 645-54.
120. Hoeller, D. and I. Dikic, *Targeting the ubiquitin system in cancer therapy.* Nature, 2009. **458**(7237): p. 438-44.
121. Atipairin, A., B. Canyuk, and A. Ratanaphan, *Cisplatin affects the conformation of apo form, not holo form, of BRCA1 RING finger domain and confers thermal stability.* Chem Biodivers, 2010. **7**(8): p. 1949-67.
122. Anderson, J.T., *RNA turnover: unexpected consequences of being tailed.* Curr Biol, 2005. **15**(16): p. R635-8.
123. Mangus, D.A., M.C. Evans, and A. Jacobson, *Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression.* Genome Biol, 2003. **4**(7): p. 223.
124. Neugebauer, K.M., *On the importance of being co-transcriptional.* J Cell Sci, 2002. **115**(Pt 20): p. 3865-71.
125. Nunes, N.M., et al., *A functional human Poly(A) site requires only a potent DSE and an A-rich upstream sequence.* EMBO J, 2010. **29**(9): p. 1523-36.
126. Zhang, H., et al., *PolyA_DB: a database for mammalian mRNA polyadenylation.* Nucleic Acids Res, 2005. **33**(Database issue): p. D116-20.
127. Keller, W., et al., *Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA.* EMBO J, 1991. **10**(13): p. 4241-9.
128. Kolev, N.G., et al., *Conserved motifs in both CPSF73 and CPSF100 are required to assemble the active endonuclease for histone mRNA 3'-end maturation.* EMBO Rep, 2008. **9**(10): p. 1013-8.
129. Mandel, C.R., et al., *Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease.* Nature, 2006. **444**(7121): p. 953-6.
130. Hirose, Y. and J.L. Manley, *RNA polymerase II is an essential mRNA polyadenylation factor.* Nature, 1998. **395**(6697): p. 93-6.
131. McCracken, S., et al., *The C-terminal domain of RNA polymerase II couples mRNA processing to transcription.* Nature, 1997. **385**(6614): p. 357-61.
132. Mirkin, N., et al., *The 3' processing factor CstF functions in the DNA repair response.* Nucleic Acids Res, 2008. **36**(6): p. 1792-804.
133. Balajee, A.S., et al., *Efficient PCNA complex formation is dependent upon both transcription coupled repair and genome overall repair.* Mutat Res, 1998. **409**(3): p. 135-46.
134. Starita, L.M., et al., *BRCA1-dependent ubiquitination of γ-tubulin regulates centrosome number.* Molecular and cellular biology, 2004. **24**(19): p. 8457-8466.
135. Li, D. and R. Roberts, *Human Genome and Diseases:¶ WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases.* Cellular and Molecular Life Sciences CMLS, 2001. **58**(14): p. 2085-2097.
136. Gaudet, R., A. Bohm, and P.B. Sigler, *Crystal structure at 2.4 angstroms resolution of the complex of transducin betagamma and its regulator, phosducin.* Cell, 1996. **87**(3): p. 577-88.

137. Wall, M.A., et al., *The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2.* Cell, 1995. **83**(6): p. 1047-58.
138. Singh, P., et al., *Global changes in processing of mRNA 3' untranslated regions characterize clinically distinct cancer subtypes.* Cancer Res, 2009. **69**(24): p. 9422-30.
139. Cevher, M.A. and F.E. Kleiman, *Connections between 3'-end processing and DNA damage response.* Wiley Interdiscip Rev RNA, 2010. **1**(1): p. 193-9.
140. Futreal, P.A., et al., *BRCA1 mutations in primary breast and ovarian carcinomas.* Science, 1994. **266**(5182): p. 120-122.
141. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.* Science, 1994. **266**(5182): p. 66-71.
142. Bork, P., et al., *A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins.* FASEB J, 1997. **11**(1): p. 68-76.
143. Callebaut, I. and J.P. Mornon, *From BRCA1 to RAP1: a widespread BRCT module closely associated with DNA repair.* FEBS Lett, 1997. **400**(1): p. 25-30.
144. 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.
145. Liu, Z., J. Wu, and X. Yu, *CCDC98 targets BRCA1 to DNA damage sites.* Nat Struct Mol Biol, 2007. **14**(8): p. 716-20.
146. Lee, M.S., et al., *Structure of the BRCT repeat domain of MDC1 and its specificity for the free COOH-terminal end of the γ-H2AX histone tail.* Journal of Biological Chemistry, 2005. **280**(37): p. 32053-32056.
147. Edwards, R.A., et al., *The BARD1 C-terminal domain structure and interactions with polyadenylation factor CstF-50.* Biochemistry, 2008. **47**(44): p. 11446-56.
148. 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.
149. Zhang, X., et al., *Structure of an XRCC1 BRCT domain: a new protein–protein interaction module.* The EMBO journal, 1998. **17**(21): p. 6404-6411.
150. Leung, C.C.Y. and J.M. Glover, *BRCT domains: easy as one, two, three.* Cell Cycle, 2011. **10**(15): p. 2461-2470.
151. Rodriguez, M.C. and Z. Songyang, *BRCT domains: phosphopeptide binding and signaling modules.* Frontiers in bioscience: a journal and virtual library, 2007. **13**: p. 5905-5915.
152. 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.
153. Thanassoulas, A., et al., *Thermodynamic study of the BRCT domain of BARD1 and its interaction with the -pSER-X-X-Phe- motif-containing BRIP1 peptide.* Biochim Biophys Acta, 2010. **1804**(9): p. 1908-16.
154. Rodriguez, M., et al., *Phosphopeptide binding specificities of BRCA1 COOH-terminal (BRCT) domains.* J Biol Chem, 2003. **278**(52): p. 52914-8.
155. Ray, H., et al., *ACCA phosphopeptide recognition by the BRCT repeats of BRCA1.* Journal of molecular biology, 2006. **359**(4): p. 973-982.
156. 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.
157. Lee, S.B., et al., *Destabilization of CHK2 by a missense mutation associated with Li-Fraumeni Syndrome.* Cancer Res, 2001. **61**(22): p. 8062-7.
158. Yarden, R.I., et al., *BRCA1 regulates the G2/M checkpoint by activating Chk1 kinase upon DNA damage.* Nature genetics, 2002. **30**(3): p. 285-289.

159. Xu, X., et al., *Centrosome amplification and a defective G 2–M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform–deficient cells.* Molecular cell, 1999. **3**(3): p. 389-395.
160. Rodríguez-Pinilla, S.M., et al., *Prognostic significance of basal-like phenotype and fascin expression in node-negative invasive breast carcinomas.* Clinical cancer research, 2006. **12**(5): p. 1533-1539.
161. Koonin, E.V., S.F. Altschul, and P. Bork, *BRCA1 protein products ... Functional motifs.* Nat Genet, 1996. **13**(3): p. 266-8.
162. Shakya, R., et al., *BRCA1 tumor suppression depends on BRCT phosphoprotein binding, but not its E3 ligase activity.* Science, 2011. **334**(6055): p. 525-528.
163. Untergasser, A., et al., *Primer3Plus, an enhanced web interface to Primer3.* Nucleic acids research, 2007. **35**(suppl 2): p. W71-W74.
164. Zhu, H. and R.A. Dean, *A novel method for increasing the transformation efficiency of Escherichia coli-application forbacterial artificial chromosome library construction.* Nucleic Acids Res, 1999. **27**(3): p. 910-1.
165. Pace, C.N. and K.L. Shaw, *Linear extrapolation method of analyzing solvent denaturation curves.* Proteins, 2000. **Suppl 4**: p. 1-7.
166. Pace, C.N., *Determination and analysis of urea and guanidine hydrochloride denaturation curves.* Methods Enzymol, 1986. **131**: p. 266-80.
167. 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.
168. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data.* Electrophoresis, 1999. **20**(18): p. 3551-67.
169. Jaroszewski, L., *Protein structure prediction based on sequence similarity*, in *Biomedical Informatics*2009, Springer. p. 129-156.
170. Kopp, J. and T. Schwede, *Automated protein structure homology modeling: a progress report.* Pharmacogenomics, 2004. **5**(4): p. 405-416.
171. Henderson, R., et al., *Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy*. Journal of molecular biology, 1990. **213**(4): p. 899-929.
172. Rost, B. and S. O'Donoghue, *Sisyphus and protein structure prediction*, 1997, CABIOS.
173. Hilbert, M., G. Böhm, and R. Jaenicke, *Structural relationships of homologous proteins as a fundamental principle in homology modeling.* Proteins: Structure, Function, and Bioinformatics, 1993. **17**(2): p. 138-151.
174. Bonneau, R. and D. Baker, *Ab initio protein structure prediction: progress and prospects.* Annual review of biophysics and biomolecular structure, 2001. **30**(1): p. 173-189.
175. Sticht, H., et al., *Homology modeling of adenylosuccinate synthetase from Saccharomyces cerevisiae reveals a possible binding region for single-stranded ARS sequences.* Journal of Biomolecular Structure and Dynamics, 1997. **14**(6): p. 667-675.
176. Xiang, Z., *Homology-based modeling of protein structure*, in *Computational Methods for Protein Structure Prediction and Modeling*2007, Springer. p. 319-357.
177. Vogt, G., S. Woell, and P. Argos, *Protein thermal stability, hydrogen bonds, and ion pairs.* Journal of molecular biology, 1997. **269**(4): p. 631-643.
178. Bowie, J.U. and D. Eisenberg, *An evolutionary approach to folding small alpha-helical proteins that uses sequence information and an empirical guiding fitness function.* Proceedings of the National Academy of Sciences, 1994. **91**(10): p. 4436-4440.
179. Simons, K.T., et al., *Ab initio protein structure prediction of CASP III targets using ROSETTA.* Proteins: Structure, Function, and Bioinformatics, 1999. **37**(S3): p. 171-176.

180. Simons, K.T., et al., *Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions.* Journal of molecular biology, 1997. **268**(1): p. 209-225.
181. Das, R., et al., *Structure prediction for CASP7 targets using extensive all-atom refinement with Rosetta@ home.* Proteins: Structure, Function, and Bioinformatics, 2007. **69**(S8): p. 118-128.
182. Bradley, P., et al., *Free modeling with Rosetta in CASP6.* Proteins: Structure, Function, and Bioinformatics, 2005. **61**(S7): p. 128-134.
183. Li, Z. and H.A. Scheraga, *Monte Carlo-minimization approach to the multiple-minima problem in protein folding.* Proceedings of the National Academy of Sciences, 1987. **84**(19): p. 6611-6615.
184. Fujitsuka, Y., G. Chikenji, and S. Takada, *SimFold energy function for de novo protein structure prediction: consensus with Rosetta.* Proteins: Structure, Function, and Bioinformatics, 2006. **62**(2): p. 381-398.
185. Lee, J., et al., *Prediction of protein tertiary structure using PROFESY, a novel method based on fragment assembly and conformational space annealing.* Proteins: Structure, Function, and Bioinformatics, 2004. **56**(4): p. 704-714.
186. Kuntz, I.D., et al., *A geometric approach to macromolecule-ligand interactions.* Journal of molecular biology, 1982. **161**(2): p. 269-288.
187. SchuÈttelkopf, 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.
188. Jones, G., et al., *Development and validation of a genetic algorithm for flexible docking.* Journal of molecular biology, 1997. **267**(3): p. 727-748.
189. Wang, J., P.A. Kollman, and I.D. Kuntz, *Flexible ligand docking: a multistep strategy approach.* Proteins: Structure, Function, and Bioinformatics, 1999. **36**(1): p. 1-19.
190. Rarey, M., B. Kramer, and T. Lengauer, *Multiple automatic base selection: protein-ligand docking based on incremental construction without manual intervention.* J Comput Aided Mol Des, 1997. **11**(4): p. 369-84.
191. Kramer, B., M. Rarey, and T. Lengauer, *CASP2 experiences with docking flexible ligands using FlexX.* Proteins, 1997. **Suppl 1**: p. 221-5.
192. Sousa, S.F., P.A. Fernandes, and M.J. Ramos, *Protein–ligand docking: current status and future challenges.* Proteins: Structure, Function, and Bioinformatics, 2006. **65**(1): p. 15-26.
193. Laurie, A.T. and R.M. Jackson, *Q-SiteFinder: an energy-based method for the prediction of protein–ligand binding sites.* Bioinformatics, 2005. **21**(9): p. 1908-1916.
194. Eisen, M.B., et al., *HOOK: a program for finding novel molecular architectures that satisfy the chemical and steric requirements of a macromolecule binding site.* Proteins: Structure, Function, and Bioinformatics, 1994. **19**(3): p. 199-221.
195. De Vries, S.J., et al., *HADDOCK versus HADDOCK: new features and performance of HADDOCK2. 0 on the CAPRI targets.* Proteins: Structure, Function, and Bioinformatics, 2007. **69**(4): p. 726-733.
196. de Vries, S.J., M. van Dijk, and A.M. Bonvin, *The HADDOCK web server for data-driven biomolecular docking.* Nat Protoc, 2010. **5**(5): p. 883-97.
197. Dominguez, C., R. Boelens, and A.M. Bonvin, *HADDOCK: a protein-protein docking approach based on biochemical or biophysical information.* J Am Chem Soc, 2003. **125**(7): p. 1731-7.
198. van Dijk, A.D., D. Fushman, and A.M. Bonvin, *Various strategies of using residual dipolar couplings in NMR-driven protein docking: Application to Lys48-linked di-ubiquitin and validation against 15N-relaxation data.* Proteins: Structure, Function, and Bioinformatics, 2005. **60**(3): p. 367-381.
199. van Dijk, A.D., et al., *Combining NMR relaxation with chemical shift perturbation data to drive protein–protein docking.* Journal of biomolecular NMR, 2006. **34**(4): p. 237-244.

200. Schmitz, C. and A.M. Bonvin, *Protein–protein HADDocking using exclusively pseudocontact shifts.* Journal of biomolecular NMR, 2011. **50**(3): p. 263-266.
201. Verlet, L., *Computer" experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules.* Physical review, 1967. **159**(1): p. 98.
202. Katchalski-Katzir, E., et al., *Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques.* Proceedings of the National Academy of Sciences, 1992. **89**(6): p. 2195-2199.
203. Gabb, H.A., R.M. Jackson, and M.J. Sternberg, *Modelling protein docking using shape complementarity, electrostatics and biochemical information.* Journal of molecular biology, 1997. **272**(1): p. 106-120.
204. Meyer, M., P. Wilson, and D. Schomburg, *Hydrogen bonding and molecular surface shape complementarity as a basis for protein docking.* Journal of molecular biology, 1996. **264**(1): p. 199-210.
205. Moont, G., H.A. Gabb, and M.J. Sternberg, *Use of pair potentials across protein interfaces in screening predicted docked complexes.* Proteins: Structure, Function, and Bioinformatics, 1999. **35**(3): p. 364-373.
206. Fernández-Recio, J., R. Abagyan, and M. Totrov, *Improving CAPRI predictions: Optimized desolvation for rigid-body docking.* Proteins: Structure, Function, and Bioinformatics, 2005. **60**(2): p. 308-313.
207. Kozakov, D., et al., *Optimal clustering for detecting near-native conformations in protein docking.* Biophysical Journal, 2005. **89**(2): p. 867-875.
208. Murphy, J., et al., *Combination of scoring functions improves discrimination in protein–protein docking.* Proteins: Structure, Function, and Bioinformatics, 2003. **53**(4): p. 840-854.
209. Li, C.H., et al., *Complex-type-dependent scoring functions in protein–protein docking.* Biophysical chemistry, 2007. **129**(1): p. 1-10.
210. Martin, O. and D. Schomburg, *Efficient comprehensive scoring of docked protein complexes using probabilistic support vector machines.* Proteins: Structure, Function, and Bioinformatics, 2008. **70**(4): p. 1367-1378.
211. Dominguez, C., R. Boelens, and A.M. Bonvin, *HADDOCK: a protein-protein docking approach based on biochemical or biophysical information.* Journal of the American Chemical Society, 2003. **125**(7): p. 1731-1737.
212. Paterlini, M.G. and D.M. Ferguson, *Constant temperature simulations using the Langevin equation with velocity Verlet integration.* Chemical Physics, 1998. **236**(1): p. 243-252.
213. Kresse, G. and J. Hafner, *Ab initio molecular dynamics for liquid metals.* Physical Review B, 1993. **47**(1): p. 558.
214. Holm, L. and P. Rosenström, *Dali server: conservation mapping in 3D.* Nucleic acids research, 2010. **38**(suppl 2): p. W545-W549.
215. 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.
216. Bakan, A., L.M. Meireles, and I. Bahar, *ProDy: protein dynamics inferred from theory and experiments.* Bioinformatics, 2011. **27**(11): p. 1575-1577.
217. Grant, B.J., et al., *Bio3d: an R package for the comparative analysis of protein structures.* Bioinformatics, 2006. **22**(21): p. 2695-2696.
218. 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.
219. Amadei, A., A. Linssen, and H. Berendsen, *Essential dynamics of proteins.* Proteins Structure Function and Bioinformatics, 1994(17): p. 412-25.

220. Amadei, A., A.B. Linssen, and H.J. Berendsen, *Essential dynamics of proteins.* Proteins, 1993. **17**(4): p. 412-25.
221. Alexandrov, V., et al., *Normal modes for predicting protein motions: a comprehensive database assessment and associated Web tool.* Protein Science, 2005. **14**(3): p. 633-643.
222. Levitt, M., C. Sander, and P.S. Stern, *Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme.* Journal of molecular biology, 1985. **181**(3): p. 423-447.
223. Hayward, S., A. Kitao, and H.J. Berendsen, *Model-free methods of analyzing domain motions in proteins from simulation: a comparison of normal mode analysis and molecular dynamics simulation of lysozyme.* Proteins: Structure, Function, and Bioinformatics, 1997. **27**(3): p. 425-437.
224. 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.
225. Wang, X., et al., *Structural NMR of protein oligomers using hybrid methods.* J Struct Biol, 2011. **173**(3): p. 515-29.
226. Cizas, P., et al., *Size-dependent neurotoxicity of β-amyloid oligomers.* Archives of biochemistry and biophysics, 2010. **496**(2): p. 84-92.
227. Cizas, P., et al., *Size-dependent neurotoxicity of beta-amyloid oligomers.* Arch Biochem Biophys, 2010. **496**(2): p. 84-92.
228. 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.
229. Feeney, R.E., *Chemical modification of proteins: comments and perspectives.* Int J Pept Protein Res, 1987. **29**(2): p. 145-61.
230. Molgaard, A. and S. Larsen, *Crystal packing in two pH-dependent crystal forms of rhamnogalacturonan acetylesterase.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 3): p. 472-8.
231. D'arcy, A., et al., *A novel approach to crystallising proteins under oil.* Journal of crystal growth, 1996. **168**(1): p. 175-180.
232. 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.
233. 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.
234. Easton, D.F., et al., *Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium.* Am J Hum Genet, 1993. **52**(4): p. 678-701.
235. Ford, D., D.F. Easton, and J. Peto, *Estimates of the gene frequency of BRCA1 and its contribution to breast and ovarian cancer incidence.* American journal of human genetics, 1995. **57**(6): p. 1457.
236. Wu, L.C., et al., *Identification of a RING protein that can interact in vivo with the BRCA1 gene product.* Nature genetics, 1996. **14**(4): p. 430-440.
237. 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.
238. 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.
239. 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.
240. Takagaki, Y., et al., *A multisubunit factor, CstF, is required for polyadenylation of mammalian pre-mRNAs.* Genes Dev, 1990. **4**(12A): p. 2112-20.

241. Takagaki, Y. and J.L. Manley, *Levels of polyadenylation factor CstF-64 control IgM heavy chain mRNA accumulation and other events associated with B cell differentiation.* Mol Cell, 1998. **2**(6): p. 761-71.
242. Le Page, F., et al., *BRCA1 and BRCA2 are necessary for the transcription-coupled repair of the oxidative 8-oxoguanine lesion in human cells.* Cancer Res, 2000. **60**(19): p. 5548-52.
243. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures.* Journal of applied crystallography, 1993. **26**(2): p. 283-291.
244. Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallographica Section D: Biological Crystallography, 2009. **66**(1): p. 12-21.
245. Davis, I.W., et al., *MolProbity: all-atom contacts and structure validation for proteins and nucleic acids.* Nucleic acids research, 2007. **35**(suppl 2): p. W375-W383.
246. Fiser, A. and A. Sali, *ModLoop: automated modeling of loops in protein structures.* Bioinformatics, 2003. **19**(18): p. 2500-2501.
247. Ramachandran, G., C. Ramakrishnan, and V. Sasisekharan, *Stereochemistry of polypeptide chain configurations.* Journal of molecular biology, 1963. **7**(1): p. 95-99.
248. Colovos, C. and T. Yeates, *ERRAT: an empirical atom-based method for validating protein structures.* Protein Science, 1993. **2**: p. 1511-1519.
249. Liithy, R., J. Bowie, and D. Eisenberg, *Assessment of protein models with three-dimensional profiles.* Nature, 1992. **356**(6364): p. 83-85.
250. Laskowski, R.A., et al., *AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR.* Journal of biomolecular NMR, 1996. **8**(4): p. 477-486.
251. Rost, B. and C. Sander, *Combining evolutionary information and neural networks to predict protein secondary structure.* Proteins: Structure, Function, and Bioinformatics, 1994. **19**(1): p. 55-72.
252. Frishman, D. and P. Argos, *Knowledge-based protein secondary structure assignment.* Proteins, 1995. **23**(4): p. 566-79.
253. Heinig, M. and D. Frishman, *STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins.* Nucleic Acids Res, 2004. **32**(Web Server issue): p. W500-2.
254. Arnold, G.E. and R.L. Ornstein, *Molecular dynamics study of time-correlated protein domain motions and molecular flexibility: cytochrome P450BM-3.* Biophys J, 1997. **73**(3): p. 1147-59.
255. Arnold, G.E. and R.L. Ornstein, *Protein hinge bending as seen in molecular dynamics simulations of native and M61 mutant T4 lysozymes.* Biopolymers, 1997. **41**(5): p. 533-44.
256. Ichiye, T. and M. Karplus, *Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations.* Proteins, 1991. **11**(3): p. 205-17.
257. Laskowski, R.A., *PDBsum: summaries and analyses of PDB structures.* Nucleic acids research, 2001. **29**(1): p. 221-222.
258. Kleiman, F.E. and J.L. Manley, *The BARD1-CstF-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression.* Cell, 2001. **104**(5): p. 743-753.
259. Collins, N., et al., *Consistent loss of the wild type allele in breast cancers from a family linked to the BRCA2 gene on chromosome 13q12-13.* Oncogene, 1995. **10**(8): p. 1673-5.
260. Wooster, R., et al., *Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13.* Science, 1994. **265**(5181): p. 2088-90.
261. Huber, L.J., et al., *Impaired DNA damage response in cells expressing an exon 11-deleted murine Brca1 variant that localizes to nuclear foci.* Mol Cell Biol, 2001. **21**(12): p. 4005-15.
262. Mohammad, D.H. and M.B. Yaffe, *14-3-3 proteins, FHA domains and BRCT domains in the DNA damage response.* DNA Repair (Amst), 2009. **8**(9): p. 1009-17.

263. Li, M., et al., *The FHA and BRCT domains recognize ADP-ribosylation during DNA damage response.* Genes Dev, 2013. **27**(16): p. 1752-68.
264. Alshatwi, A.A., et al., *Identification of functional SNPs in BARD1 gene and in silico analysis of damaging SNPs: based on data procured from dbSNP database.* PLoS One, 2012. **7**(10): p. e43939.
265. Antoniou, A.C., et al., *Common breast cancer susceptibility alleles and the risk of breast cancer for BRCA1 and BRCA2 mutation carriers: implications for risk prediction.* Cancer Res, 2010. **70**(23): p. 9742-54.
266. Karppinen, S.-M., et al., *Mutation screening of the BARD1 gene: evidence for involvement of the Cys557Ser allele in hereditary susceptibility to breast cancer.* Journal of medical genetics, 2004. **41**(9): p. e114-e114.
267. Yoshikawa, K., et al., *Abnormal expression of BRCA1 and BRCA1-interactive DNA-repair proteins in breast carcinomas.* Int J Cancer, 2000. **88**(1): p. 28-36.
268. Parvin, J.D. and S. Sankaran, *The BRCA1 E3 ubiquitin ligase controls centrosome dynamics.* Cell Cycle, 2006. **5**(17): p. 1946-50.
269. 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.
270. 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.
271. Scully, R., et al., *Association of BRCA1 with Rad51 in mitotic and meiotic cells.* Cell, 1997. **88**(2): p. 265-75.
272. Wang, Q., et al., *Adenosine nucleotide modulates the physical interaction between hMSH2 and BRCA1.* Oncogene, 2001. **20**(34): p. 4640-9.
273. Zhang, H., et al., *BRCA1 physically associates with p53 and stimulates its transcriptional activity.* Oncogene, 1998. **16**(13): p. 1713-21.
274. Somasundaram, K., et al., *Arrest of the cell cycle by the tumour-suppressor BRCA1 requires the CDK-inhibitor p21WAF1/CiP1.* Nature, 1997. **389**(6647): p. 187-90.
275. Xia, Y., et al., *Enhancement of BRCA1 E3 ubiquitin ligase activity through direct interaction with the BARD1 protein.* J Biol Chem, 2003. **278**(7): p. 5255-63.
276. Ruffner, H., et al., *Cancer-predisposing mutations within the RING domain of BRCA1: loss of ubiquitin protein ligase activity and protection from radiation hypersensitivity.* Proc Natl Acad Sci U S A, 2001. **98**(9): p. 5134-9.
277. Wu, W., et al., *The ubiquitin E3 ligase activity of BRCA1 and its biological functions.* Cell Div, 2008. **3**: p. 1.
278. Lord, C.J., M.D. Garrett, and A. Ashworth, *Targeting the double-strand DNA break repair pathway as a therapeutic strategy.* Clinical cancer research, 2006. **12**(15): p. 4463-4468.
279. Arnold, K., et al., *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling.* Bioinformatics, 2006. **22**(2): p. 195-201.
280. Studio, D., *version 2.1.* Accelrys: San Diego, CA, 2008.