

SCIENTIFIC REPORTS



OPEN

Structural basis to stabilize the domain motion of BARD1-ARD BRCT by CstF50

Rajan Kumar Choudhary^{1,2}, Mohd Quadir Siddiqui^{1,2}, Pankaj S. Thapa^{1,2}, Nikhil Gadewal^{1,2}, Senthil Kumar Nachimuthu³ & Ashok K. Varma^{1,2}

BRCA1 associated ring domain protein 1 (BARD1) is a tumor suppressor protein having a wide role in cellular processes like cell-cycle checkpoint, DNA damage repair and maintenance of genomic integrity. Germ-line mutation Gln 564 His discovered in linker region of BARD1 leads to loss of binding to Cleavage stimulating factor (CstF50), which in turn instigates the premature mRNA transcript formation and apoptosis. We have studied the dynamics of ARD domain present in the BARD1 *wild-type* and mutant protein in association with CstF50 using biophysical, biochemical and molecular dynamics simulations. It has been observed that the ARD domain is relatively more flexible than the BRCT domain of BARD1. Further relative orientations of both the ARD and BRCT domains varies due to the highly flexible nature of the connecting linker region present between the domains. It has been observed that mutant ARD domain is more dynamic in nature compared to *wild-type* protein. Molecular docking studies between BARD1 Gln 564 His mutant and CstF50 shows the loss of interactions. Furthermore, domain motion of ARD present in BARD1 was stabilized when complexed with CstF50.

BRCA1 associated ring domain 1 (BARD1) is a binding partner of breast cancer associated protein (BRCA1)^{1,2}. BARD1 comprises different domains including N-terminus ring domain, an ankyrin repeat domain (ARD) and two-tandem BRCT domains³. BARD1-BRCA1, RING-RING domain complex is an E3 ubiquitin ligase complex which is important for many cellular processes like DNA damage repair, cell cycle checkpoint and genomic integrity^{1,4-8}. Mutations associated to breast and ovarian tumors present in the ring domain of BRCA1 impairs the protein-protein interactions and complex formation with BARD1^{6,9,10}. ARD and BRCT domains are connected via 14 residue linker region and provide binding site for CstF50^{11,12}. The BARD1 ARD-BRCT domain is essentially required for homology-directed DNA (HDR) repair, as cells transfected with truncated BARD1 are unable to carry out HDR¹³. BRCA1-BARD1 complex with Cleavage stimulating factor (CstF50) has pivotal role in transcription-coupled DNA-damage repair (TCR)^{11,12}. CstF50 has 7 Tryptophan-Aspartate (WD-40) repeats that interacts to RNA polymerase II (RNAP II) through its N-terminal region^{14,15} whereas BARD1 interacts through its 7th WD-40 repeat^{16,17}. BARD1/BRCA1-CstF50 complex, after UV exposure leads to RNAP II CTD degradation¹⁷⁻¹⁹ and inhibition of 3'end- mRNA processing^{16,17,20}. Missense mutations in the BARD1 have been reported in sporadic cases of ovarian, uterine, and breast carcinoma as well as in patients with familial breast or ovarian cancer²¹⁻²⁵. Germ line mutation Gln 564 His reported in BARD1 disrupts the interactions with CstF50 and p53. BARD1 Gln 564 His mutation affects BARD1 mediated p53-dependent apoptosis and 3'- mRNA processing^{12,21,26}.

Multidisciplinary approaches were explored to study the effect of Gln 564 His mutation on the stability of BARD1 ARD-BRCT domain and interactions with CstF50. Furthermore, protein-protein docking and molecular dynamics simulation approach were employed to unravel the effect of mutation on domain dynamics of BARD1 ARD-BRCT, Gln 564 His mutant and BARD1 ARD-BRCT-CstF50 complex.

Results and Discussion

Oligomeric behavior of BARD1 ARD-BRCT. To delineate the effect of mutation on the oligomeric behavior of BARD1 ARD-BRCT, the purified BARD1 ARD-BRCT *wild-type* and mutant proteins (Figure 1A,B

¹Advanced Centre for Treatment, Research and Education in Cancer, Kharghar, Navi Mumbai, Maharashtra, 410 210, India. ²Homi Bhabha National Institute, Training School Complex, Anushaktinagar, Mumbai, 400 094, India. ³Department of Biotechnology, Mizoram University (A Central University) Aizawl, 796 004, Mizoram, India. Correspondence and requests for materials should be addressed to A.K.V. (email: avarma@actrec.gov.in)

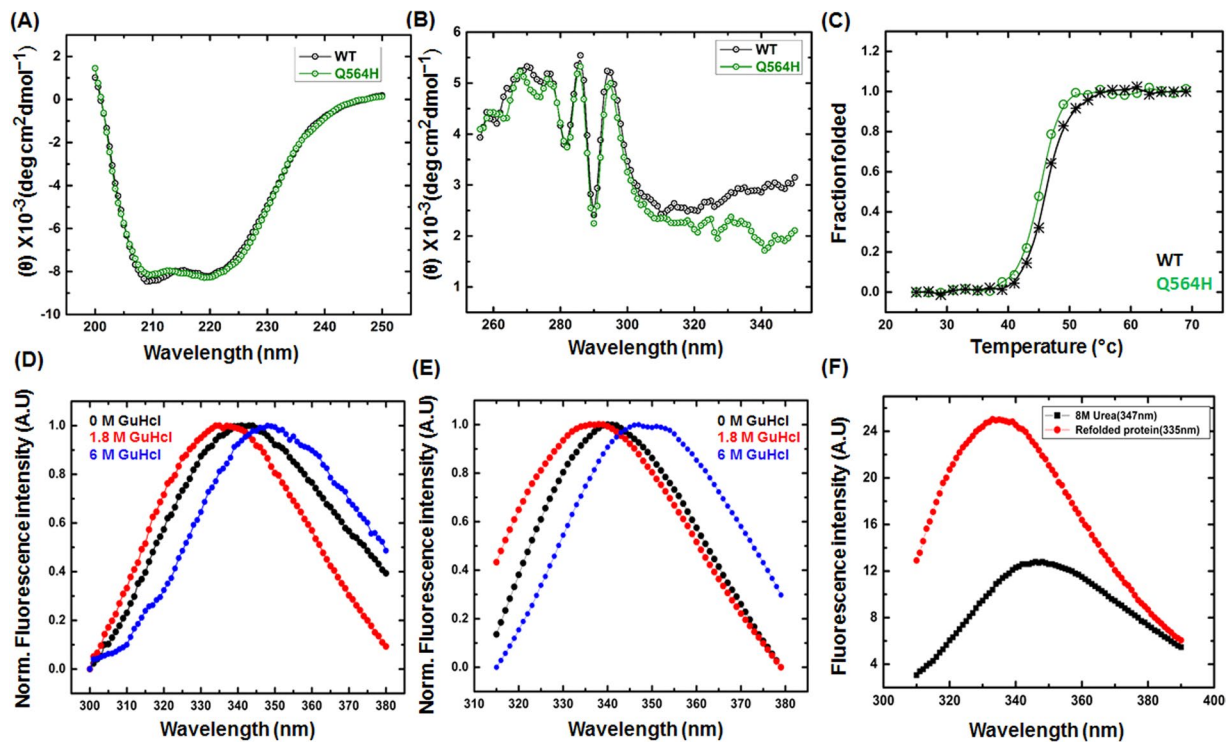


Figure 1. (A and B) Comparative Far -UV and Near -UV, CD spectroscopy profile of BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant respectively, (C) Comparative thermal denaturation profile of *wild-type* and Gln 564 His mutant protein, (D and E) Comparative chemical denaturation profile of profile of *wild-type* and Gln 564 His mutant protein respectively and (F) Refolding profile of CstF50.

Supplementary) were subjected to size exclusion chromatography and dynamic light scattering (DLS). The hydrodynamic radii of BARD1 ARD-BRCT *wild-type* and mutant protein from DLS experiment were 3.3 ± 0.45 nm and 3.38 ± 0.15 nm respectively which suggest that BARD1 Gln 564 His mutation does not affect the overall packing within the hydrophobic core of the protein (Figure 1E Supplementary). Both BARD1 ARD-BRCT *wild-type* and mutant proteins were eluted in the monomeric elution volume, suggesting their monomeric nature (Figure 2 Supplementary). Glutaraldehyde cross-linking assay for *wild-type* and mutant proteins do not show higher molecular weight conjugates on the 12% SDS-PAGE gel which further support the monomeric nature of proteins (Figure 3A and B Supplementary).

ARD and BRCT domains of BARD1 are compact and folds independently. CD Spectroscopy (CD) was performed to explore the effect of mutation on the secondary structure of BARD1 ARD-BRCT. The CD spectra for *wild-type* and mutant protein shows that secondary structure of the protein is majorly stabilized by α -helices because of prominent ellipticity minima at $\lambda = 208$ nm and 222 nm (Fig. 1A). However, the CD spectra do not show change in the secondary structure of mutant protein which indicates that mutation has no effect on the secondary structure of the *wild-type* BARD1 ARD-BRCT protein. However, the absence of characteristic random coils and β -sheets are due to the dominant contribution of CD signal from the α -helix portion of the structure. Additionally, this masking of random coil and β -sheets can also be due to the disulfide chromophores present in ARD-BRCT region which strongly contributes to the CD signal in the range of $\lambda = 215$ –235 nm^{27, 28}. CD spectroscopy was further performed at near-UV range to investigate the effect of mutation on the tertiary structure of the protein. Near-UV spectra of *wild-type* and mutant proteins show peaks at $\lambda_{\max} = 295$ nm and 285 nm respectively, which indicates the hydrophobic microenvironments of intrinsic fluorophores substantiating the compact 3D structure of the proteins²⁹ (Fig. 1B). Furthermore, limited proteolysis profile of *wild-type* and mutant protein indicates that ARD and BRCT domain of BARD1 are two compact and independently folding domains (Figure 3C and D Supplementary).

Thermal stability of BARD1 ARD-BRCT. Thermal denaturation was performed to delineate the effect of mutation on the thermal stability of the BARD1 ARD-BRCT *wild-type*. Thermal induced denaturation was monitored from 25 °C–65 °C using CD spectroscopy. The loss in secondary structure was observed at 65 °C for both the *wild-type* and mutant proteins (Figure 4A and B Supplementary). The unfolded fraction was calculated by using change in millidegree ellipticity at $\lambda = 222$ nm, which suggest that *wild-type* and mutant protein unfold via a two state pathway in a co-operative manner. The T_m calculated for *wild-type* and mutant were 45.6 ± 0.60 °C and 45.2 ± 0.42 °C respectively, which further suggest that mutation does not drastically affect the thermal stability of the BARD1 ARD-BRCT *wild-type* (Fig. 1C). Chemical denaturation induced by GuHCl was monitored using fluorescence spectroscopy. Fluorescence emission maximum at $\lambda_{\max} = 340$ nm for *wild-type* and at $\lambda_{\max} = 341$ nm

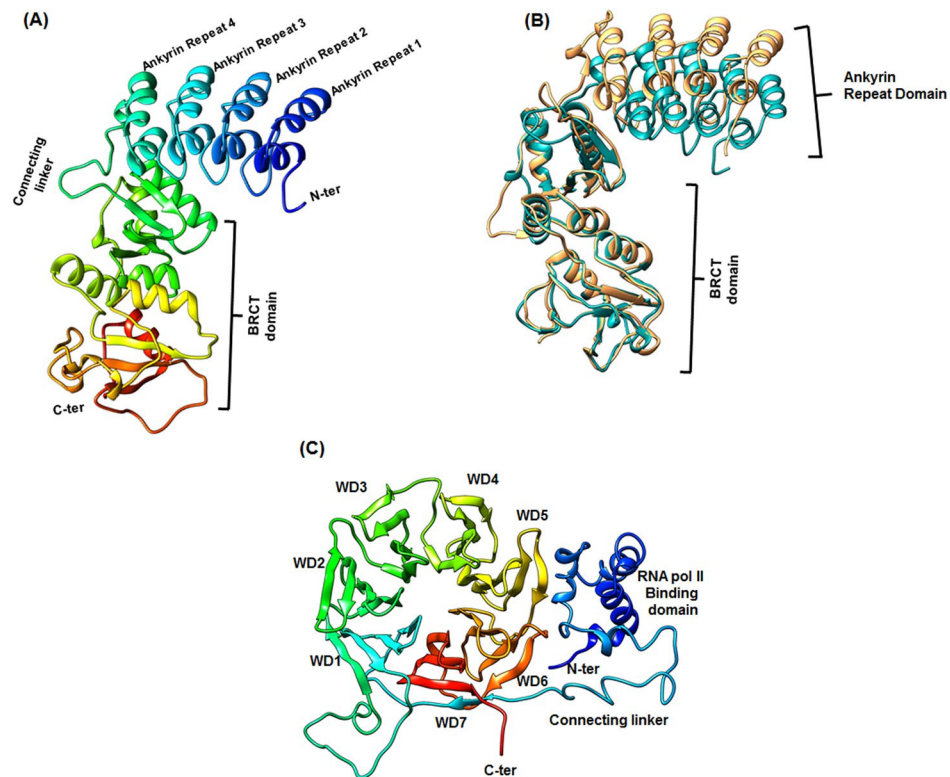


Figure 2. (A) Representative molecular model of BARD1 ARD-BRCT *wild-type*, (B) Comparative structural alignment profile of BARD1 ARD-BRCT *wild-type* (Navy Blue) and Gln 564 His Mutant (Yellow) and (C) Molecular Model structure of CstF50.

for mutant indicate that proteins are in folded conformation. However, at 1.8 M GuHCl concentration significant blue-shifts in the emission maxima was observed for *wild-type* at $\lambda_{\max} = 335$ nm and mutant at $\lambda = 336$ nm respectively. *Wild-type* and mutant completely unfold at 6 M GuHCl as indicated by the emission maximum at $\lambda_{\max} = 348$ nm. These observations suggest that *wild-type* and mutant unfold via an intermediate molten globule formation (Fig. 1D and E), and (Figure 4C and D Supplementary).

Furthermore, to investigate the molecular interactions between CstF50 and BARD1, we have expressed His-tagged CstF50 in bacterial system that expressed mostly in inclusion bodies (Figure 1C and D Supplementary). Hence 6XHis tagged fused CstF50 protein was extracted using 8 M urea, and then further refolded in the buffer D (50 mM Tris, pH 6.2, 500 mM NaCl, 0.1% triton, 10% glycerol pH 7.2, 2 mM EDTA, 700 mM arginine, 100 mM KCl, 50 mM MgCl₂, 100 mM CaCl₂ and 5 mM DTT) containing Arginine. Refolding pattern of CstF50 protein was further confirmed by fluorescence spectroscopy. The fluorescence spectroscopy profile for denatured CstF50 in 8 M urea shows an emission maximum at $\lambda_{\max} = 347$ nm, however, folded protein shows a blue-shifted emission maximum at $\lambda_{\max} = 337$ nm, which indicates complete burial of tryptophans in the hydrophobic core of the protein (Fig. 1F). Furthermore, good quality CD spectra could not be achieved because arginine hydrochloride present in buffer D interfered with the measurement^{30,31}. An attempt to dialyze out arginine hydrochloride resulted in precipitation of CstF50 protein.

CstF50 stabilizes the ARD domain of BARD1: *In-silico* approach. To understand the function associated to protein-protein interactions for BARD1-CstF50 complex, BARD1 ARD-BRCT region and CstF50 were modeled using Robetta server^{32,33}. This Model of BARD1 ARD-BRCT and CstF50 were further validated by MolProbity server³⁴ (Table 1 and 2 Supplementary). The 3-D structure of BARD1 ARD-BRCT which comprises 4, N-terminal ankyrin repeats containing α -helices, a linker, two BRCT repeats containing 3- α helices and three β -strands (Fig. 2A). BARD1 ARD-BRCT Gln 564 His mutation was generated using SPDB viewer³⁵. The BARD1 ARD-BRCT and Gln 564 His mutant structures were superimposed for direct comparison between *wild-type* and Mutant models (Fig. 2B). Modeled structure of CstF50 shows N-terminal helical region and C-terminal 7 WD-40 repeat comprised of beta sheets that fold into a propeller like structure (Fig. 2C).

Normal Mode Analysis (NMA) was performed on the BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant structures to investigate the effect of mutation on protein dynamics. First few modes known as a low frequency modes describe the large domain motion within the protein³⁶. Hence, to understand the structural dynamics between ARD and BRCT domain, we have selected the 1st lowest frequency mode for BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant protein structures (Fig. 3A–D). A large inter-domain motion was observed between the ARD and BRCT domain of BARD1. The inter-domain motion is often observed in multidomain proteins having hinge bending³⁷. The residues at ARD and BRCT domain show high positive correlated motion at

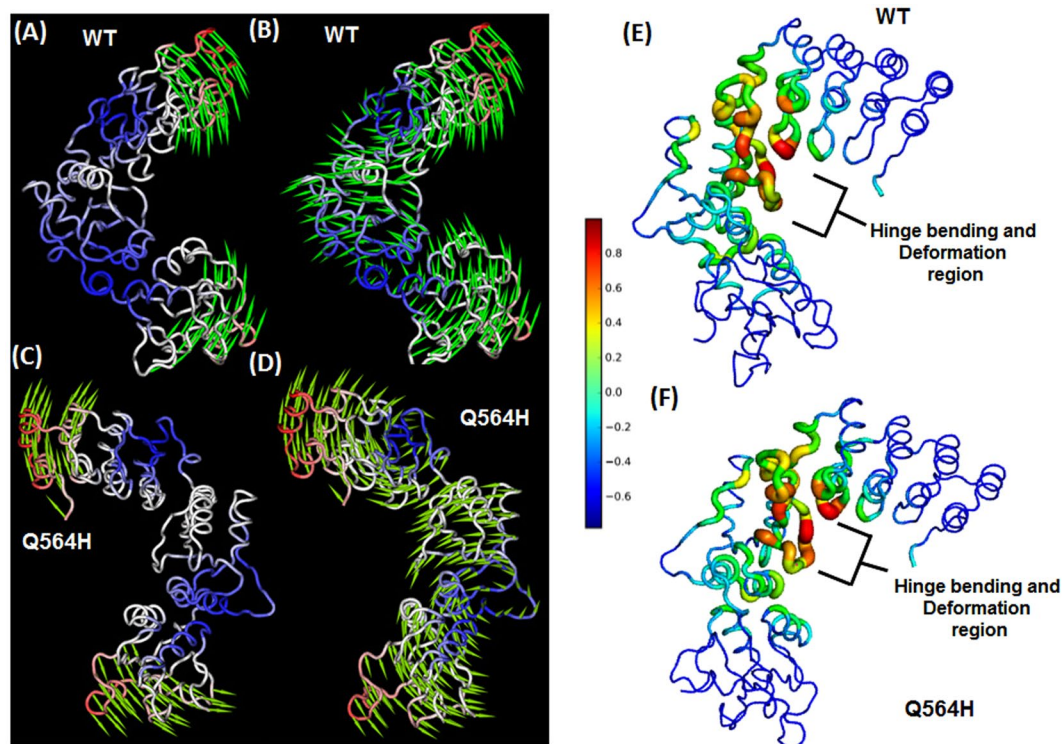


Figure 3. (A and B) Comparative Normal Mode Analysis and representation of mode 1 profile of BARD1 ARD-BRCT *wild-type*, (C and D) Gln 564 His mutant protein. (E and F) Comparative deformation analysis profile of BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant protein.

the terminal regions. Interestingly, the linker region residues show high anti-correlated motion for the ARD and BRCT domain residues in *wild-type* and mutant protein structures. The region of hinge bending and local flexibility were investigated using deformation analysis (Fig. 3E and F). The linker region was found to possess highest flexibility in both *wild-type* and mutant (thickness and color intensity) which shows that the linker between inter-domain region is responsible for large dynamics between ARD domain and BRCT domains.

BARD1- ARD domain motion and CstF50 complex. To understand the function associated to protein-protein interactions of BARD1-CstF50 complex, and inter-domain motion between ARD and BRCT domains, the model structure of BARD1 ARD-BRCT, BARD1 Gln 564 His mutant, CstF50 and BARD1 ARD-BRCT-CstF50 complex were simulated for 100 ns under periodic boundary conditions. The RMSD (Root Mean Square Deviation) profile for BARD1 shows that mutant protein structure exhibits higher RMSD than the *wild-type* (Fig. 4A). During the first 5 ns of simulation, RMSD values of *wild-type* and mutant have increased up to a value of 5 Å. However, after 10 ns, RMSD of *wild-type* decreases to 3.5 Å and mutant increases up to 6.5 Å. At 16 ns, RMSD value for *wild-type* show sharp increase and reaches up to 8 Å. Hence, the mutant structure RMSD shows only a gradual rise to reach the level of *wild-type* RMSD (Fig. 4A). Beyond 16 ns, the decrease in the RMSD up to 4 Å for *wild-type* and mutant structure is indicative of structural stabilization. However, the RMSD of mutant protein structure remained high as compared to *wild-type*. After 35 ns of simulation, *wild-type* and mutant structures are stabilized for next 50 ns. Differences in RMSD were again observed at 98 ns and 96 ns of simulation for mutant and *wild-type* structures respectively. To investigate change in the RMSD, 0 ns input structures were superimposed onto the structure of 16 ns and 98 ns time frame for *wild-type* and mutant respectively. The two superimposed structures (Fig. 5A and B) show large ARD domain movement with respect to BRCT domain for *wild-type* and mutant protein structures. Superimposed structure of *wild-type* with input structure shows that ARD domain rotates about 70° relative to BRCT domain contributing to the RMSD value of 8.55 Å. However, the mutant shows that rotation of ARD domain by 84.4° leading to a RMSD of 9.27 Å with respect to the structure of 0 ns (Fig. 5A and B). The RMSD attained during the entire simulation by ARD domain of *wild-type* mutant shows that ARD domain is dynamic in nature comparative to the BRCT domain, which is stable during entire simulation (Fig. 4C). The RMSD change for ARD domain in mutant protein shows gradual increase but *wild-type* ARD domain is stable. Furthermore, 10 superimposed structures extracted for *wild-type* and mutant protein at different time points of trajectory show highly dynamic behavior of ARD domain, whereas the BRCT domain is stable in nature. The ARD domain in mutant shows conformational changes compared to *wild-type*, signifying that ARD domain in mutant protein is more dynamic than in *wild-type* (Fig. 4F and G). However, RMSD profile for CstF50 structure indicates that CstF50 is compact protein and does not show any changes in the model structure (Fig. 4B).

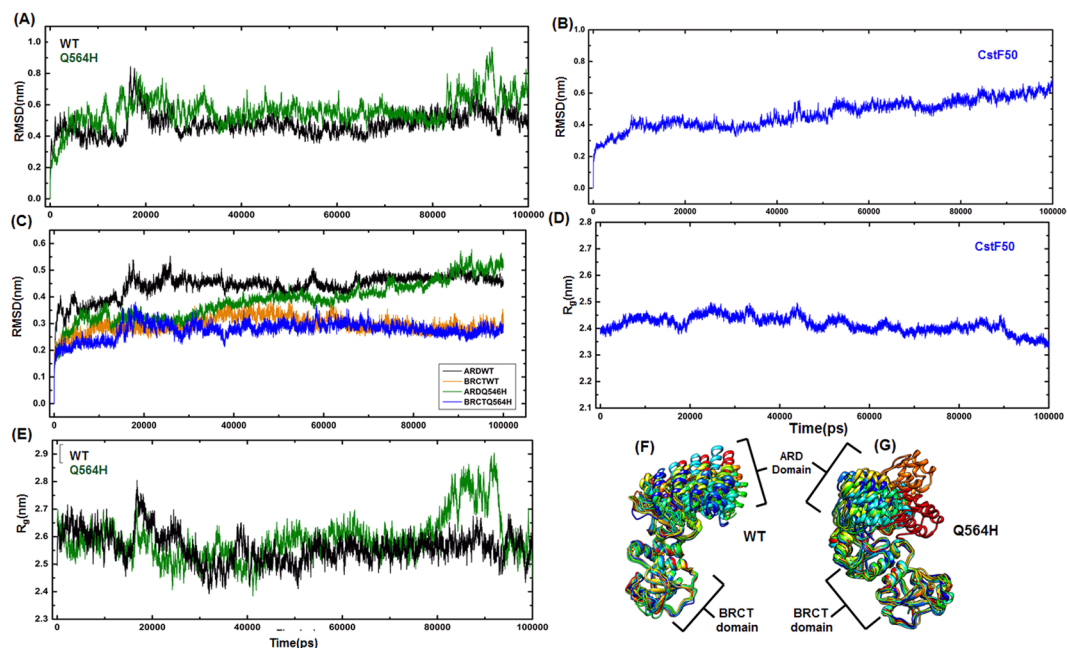


Figure 4. (A and C) Comparative RMSD and R_g profile of *wild-type* and Gln 564 His mutant protein respectively (B and D) RMSD and R_g profile of CstF50 respectively. (E) Individual RMSD profile of *wild-type* and Gln 564 His mutant protein ARD and BRCT domain respectively. (F and G) Comparative superimposition of 10 frames extracted from trajectory of *wild-type* and Gln 564 His mutant protein respectively.

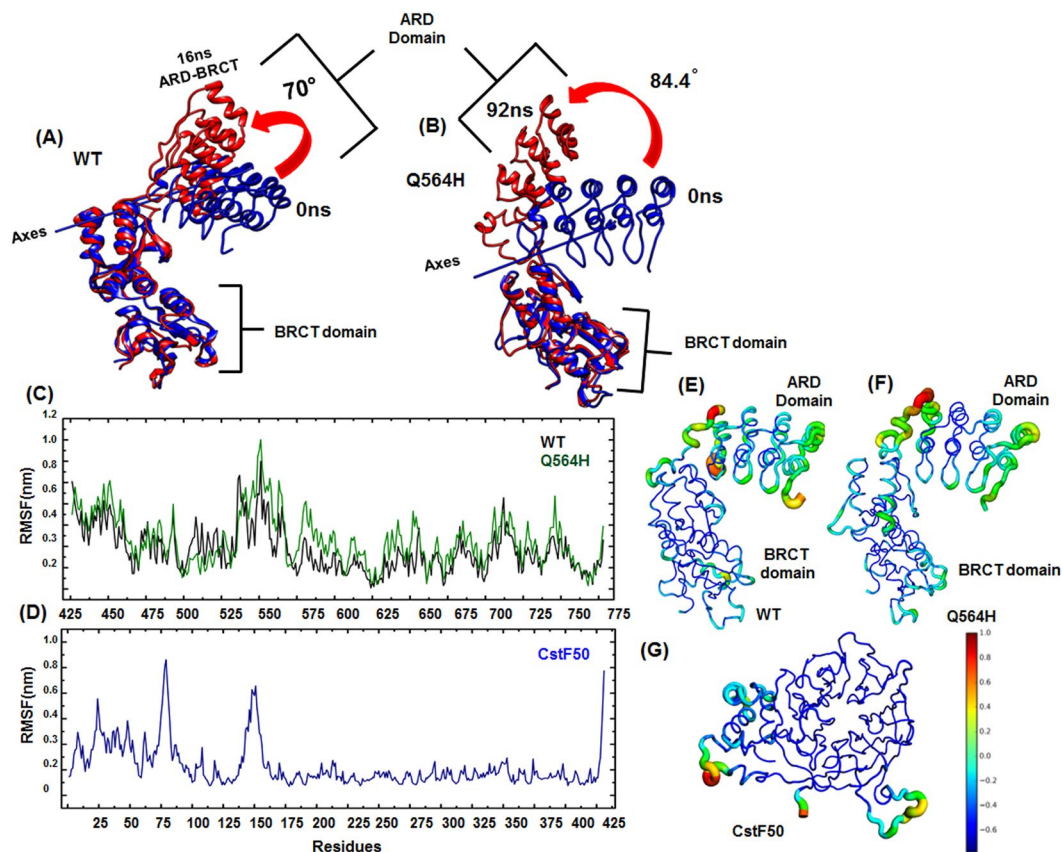


Figure 5. (A and B) Comparative superimposition of lowest RMSD and highest RMSD attained structure for *wild-type* and Gln 564 His mutant protein respectively (C) RMSF of *wild-type* and Gln 564 His mutant protein respectively. (D) RMSF profile of CstF50. (E, F and G) RMSF structure extracted from trajectory of *wild-type*, Gln 564 His mutant and CstF50 respectively.

Furthermore, changes in the radius of gyration (R_{gyr}) in a time dependent manner during the course of simulation determines the protein structure compactness³⁸. During the early phase of simulation *wild-type* and mutant did not show any sharp changes in the R_{gyr} but steep rise in the R_{gyr} was observed at 16 ns in *wild-type* and at 98 ns in mutant structure of BARD1. In mutant structure, R_{gyr} change is significantly higher than that of *wild-type*, substantiating the inference that mutation has made the mutant protein dynamic than the *wild-type* structure (Fig. 4E). Change in R_{gyr} is largely due to change in the position of ARD domain relative to the BRCT domain of BARD1. However, R_{gyr} profile for CstF50 structure indicates that CstF50 has overall compact structure and does not show any changes in the conformation (Fig. 4D).

Comparative residual RMSF (Root Mean Square Fluctuation) values attained by *wild-type* and mutant suggest that the N-terminal ARD domain of BARD1 is highly flexible. The linker region (550–568) amino acids in BARD1 ARD-BRCT shows highest RMSF values in *wild-type* and Gln 564 His mutant structures (Fig. 5C). Hence, to explore the residual flexibility of BARD1 ARD-BRCT, RMSF structures showing maximum fluctuation from the trajectory were extracted (Fig. 5D and E). The RMSF structures for *wild-type* and mutant indicate that ARD domain is highly flexible, and high value of RMSF is observed for the residues in N-terminal, linker and loop regions of second BRCT repeat (Fig. 5C, E and F). However, in the CstF50 structure, RMSF indicated two flexible regions, first at the connecting loops between N-terminal and WD-40 domain, and second at the loop region between WD1 and WD2 repeat which also corroborates with the residual RMSF plotted for the CstF50 (Fig. 5D and G).

Furthermore, domain movement of the protein structures have been studied in relation to possible involvement with protein function³⁹. To probe the collective motion during the simulation, the principal component analysis (PCA) was performed over the BARD1-ARD-BRCT, Gln 564 His mutant, CstF50 and BARD1-ARD-BRCT-CstF50 complex trajectories. The Eigenvalues and eigenvectors were calculated after the diagonalization of the covariance matrix. The area of trace of covariance matrix calculated for BARD1-ARD-BRCT, Gln 564 His mutant and CstF50 were 487.990 nm², 714.963 nm² and 540.810 nm² respectively, which indicates that mutant shows more dynamic tertiary structural conformations than the *wild-type* protein. Hence, the trajectories for domain motion of BARD1 ARD-BRCT *wild-type*, Gln 564 His mutant and CstF50 were projected on the first three eigenvectors. Projection on eigenvector 2 and 1 for *wild-type* and mutant protein indicates that mutant protein is more dynamic in nature than the *wild-type*. Trajectory Projection on eigenvectors combination 3 and 1, 2 and 1 and 3 and 2 for *wild-type* (Fig. 6A, B and C) and mutant (Figure 6D, E and F) indicates that both the structures can attain three major tertiary structure conformations whereas mutant structure shows higher periodic jump. Further, projection of eigenvectors 3 and 1, 2 and 1 for CstF50 shows that two major groups of tertiary structural conformation can be attained by the CstF50. The large transitions and periodic fluctuation between these conformers throughout the trajectory have been observed (Fig. 6G, H and I).

Comparative residual displacement along the Eigenvectors. To investigate the residual fluctuation, first two eigenvector extracted from trajectories of BARD1 ARD-BRCT *wild-type*, Gln 564 His mutant and CstF50 were projected on the residues. The plot of projection on eigenvector 1 and 2 of residues for *wild-type* (Figure 5A, B Supplementary) and mutant (Figure 5C, D Supplementary) demonstrates concerted residual fluctuation, which shows high flexibility in ARD domain, linker region and loop regions of BARD1 BRCT domain. Furthermore, eigenvector 1 and 2 were also projected on residue for CstF50 (Figure 5E, F Supplementary). The projection of eigenvector 1 and 2 show different profile of residual fluctuation and do not shows any similarity. The projection of eigenvector 1 on residue shows flexible loop regions connecting N-ter, WD-40 domain. On the other hand projection of eigenvector 2 on residues shows stable WD-40 domain but fluctuations in the N-terminal domain that may be due to the connecting loop regions and untethered N-terminal (Figure 5E, F Supplementary).

Positive correlation motion of BARD1 ARD domain. Cross correlation for PCA of BARD1 *wild-type*, Gln 564 His mutant protein, CstF50 and complex was plotted to investigate the correlated motion within the proteins. Cross-correlation of PCA indicates high positive correlation within the ARD domain. The residues within the ARD domain show strong negative correlation with the linker region which is in consistent with the NMA results (Fig. 7A, B). Comparatively, BARD1 *wild-type* shows higher positive correlation motion within the ARD domain and linker region residues as compared to the mutant protein. Furthermore, mutation has also affected the positive correlation within the BRCT repeat as *wild-type* shows higher positive correlation within the N-terminal BRCT and C-terminal repeat residues. CstF50 cross-correlation for PCA indicates that the N-terminal residues show little positive correlation within themselves as compared to the WD-40 repeat domain, which shows very high positive correlated motion within their domain (Fig. 7C). However, BARD1 ARD-BRCT-CstF50 complex cross-correlation for PCA indicates increased positive correlated motion within the N-terminal residues and WD-40 repeat for CstF50 in complex as compared to native condition. Furthermore, higher positive correlation is also observed in the ARD domain as well as BRCT domain in complex compared to the BARD1 ARD-BRCT alone (Fig. 7D).

Protein-Protein Interactions between BARD1 ARD-BRCT and CstF50. The minimum free energy model structures of BARD1 ARD-BRCT, Gln 564 His mutant and CstF50 were selected for protein-protein docking studies (Figure 6A, B and C Supplementary). It is well established that 7th WD 40 (395–431) repeat domain of CstF50 is required to establish BARD1 complex²⁶. The Gln at 564 to His mutation reduces the binding affinity of CstF50 to BARD1¹⁷. Therefore, we considered BARD1 Gln 564 as an important residue for interaction study with CstF50. In-silico docking was performed to understand the residual interactions between BARD1 ARD-BRCT, Gln 564 His mutant with CstF50. In the intermolecular interface of BARD1-CstF50 complex, glutamine 564 of BARD1 ARD-BRCT forms hydrogen bond with Asn 409 and Tyr 425 of CstF50, and also participates in

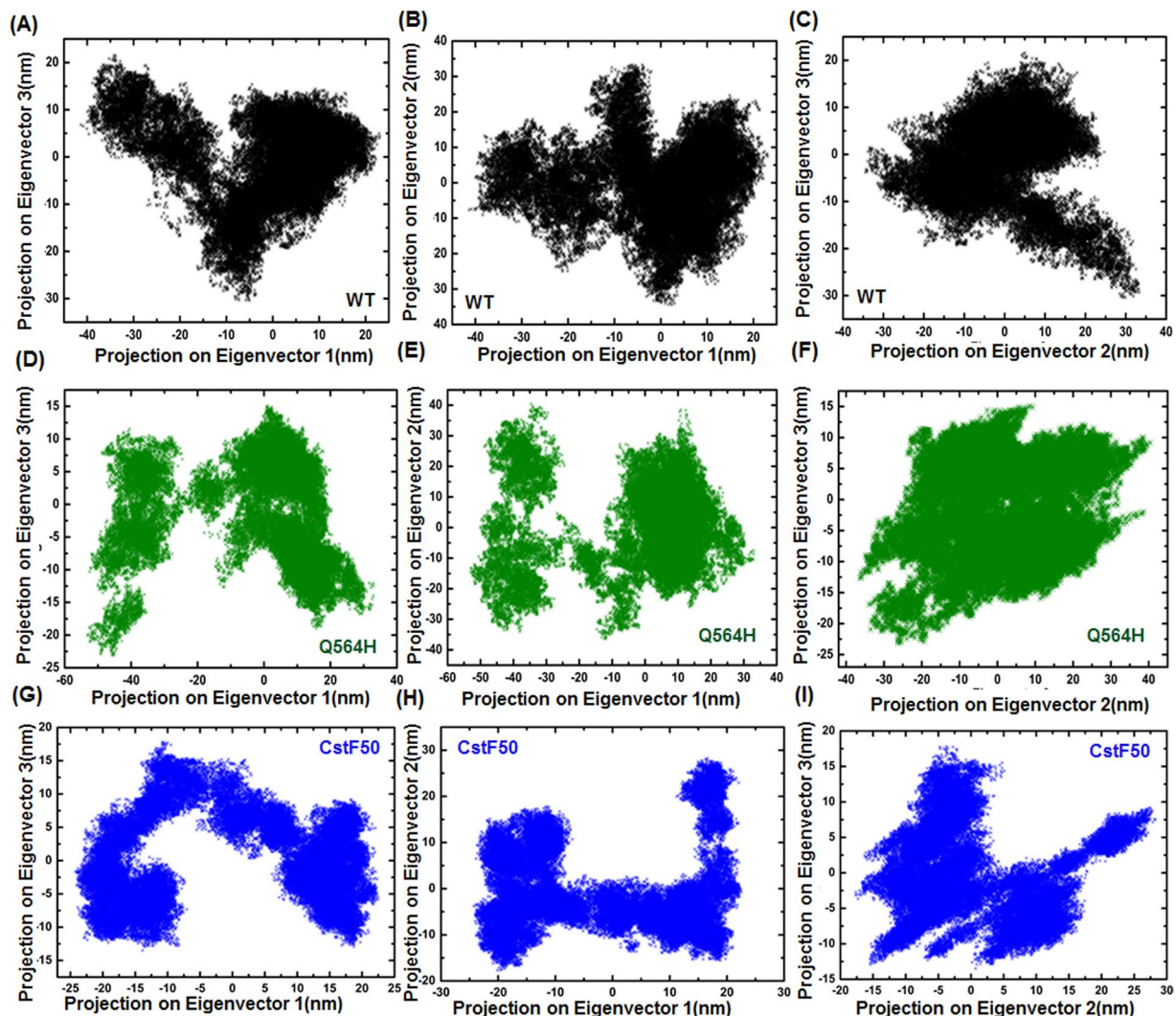


Figure 6. (A,B and C) Comparative Eigenvector projection profile of BARD1 ARD-BRCT *wild-type* (D,E and F) Gln 564 His mutant protein respectively (G,H and I) CstF50 respectively.

non-hydrogen bonding interactions with Tyr 425, Asn 409, Pro 410, and Thr 408 of CstF50. The other BARD1 residues Thr 562 and Gly 563 also form hydrogen bonds with Thr 430, Tyr 425 and Ser 428 of CstF50 respectively (Figure 7A and B Supplementary). However, in the BARD1 Gln 564 His mutant-CstF50 complex, BARD1 His 564 lost hydrogen bonding interactions but forms non-covalent interactions with Gly 394 and Leu 395 of the CstF50 (Figure 7C and D Supplementary). The BARD1 Thr 562 was found to be involved in the non-covalent interactions with CstF50 Glu 100. Loss of weak intermolecular interactions in case of mutant rationalizes the loss of binding affinity between mutant BARD1 and CstF50 (Figure 7D Supplementary).

Binding Interface of BARD1 ARD-BRCT and CstF50 complex. To determine the stability of BARD1 ARD-BRCT-CstF50 complex molecule we have simulated the structure for 100 ns to analyze the comparative RMSD, RMSE, R_{gyr} and hydrogen bonding (Fig. 8). Comparative RMSD and R_{gyr} indicate that binding of CstF50 to the BARD1 ARD-BRCT restricts the ARD domain fluctuation relative to the BRCT domain (Fig. 8A and C). Hydrogen bonding analysis between the linker region and 7th WD-40 domain was performed to analyze the domain stability of complex (Fig. 8B). It was found that BARD1 ARD-BRCT *wild-type* linker and 7th WD-40 repeat form stable hydrogen bonds throughout the simulation, thus enabling formation of a stable complex (Fig. 8D). Further, binding energy and other energy component calculated for BARD1-CstF50 complex for entire simulation indicated that interaction between BARD1 and CstF50 is stable (Table 4 Supplementary). Interestingly, interactions between CstF50 and the BARD1 linker rigidifies the linker flexibility and thereby controls the architectural placement of the BARD1 C-terminal domain relative to the N-terminal ARD domain (Fig. 8E). This phenomenon thereby reduce the relative inter domain flexibility of the ankyrin repeats and BRCT domains. The short length of the linker and germ-line cancer-predisposing mutations suggest that the relative positioning of ARD and BRCT domains within BARD1 is critical in the binding to proteins and functionality of BARD1 (Fig. 8E and F).

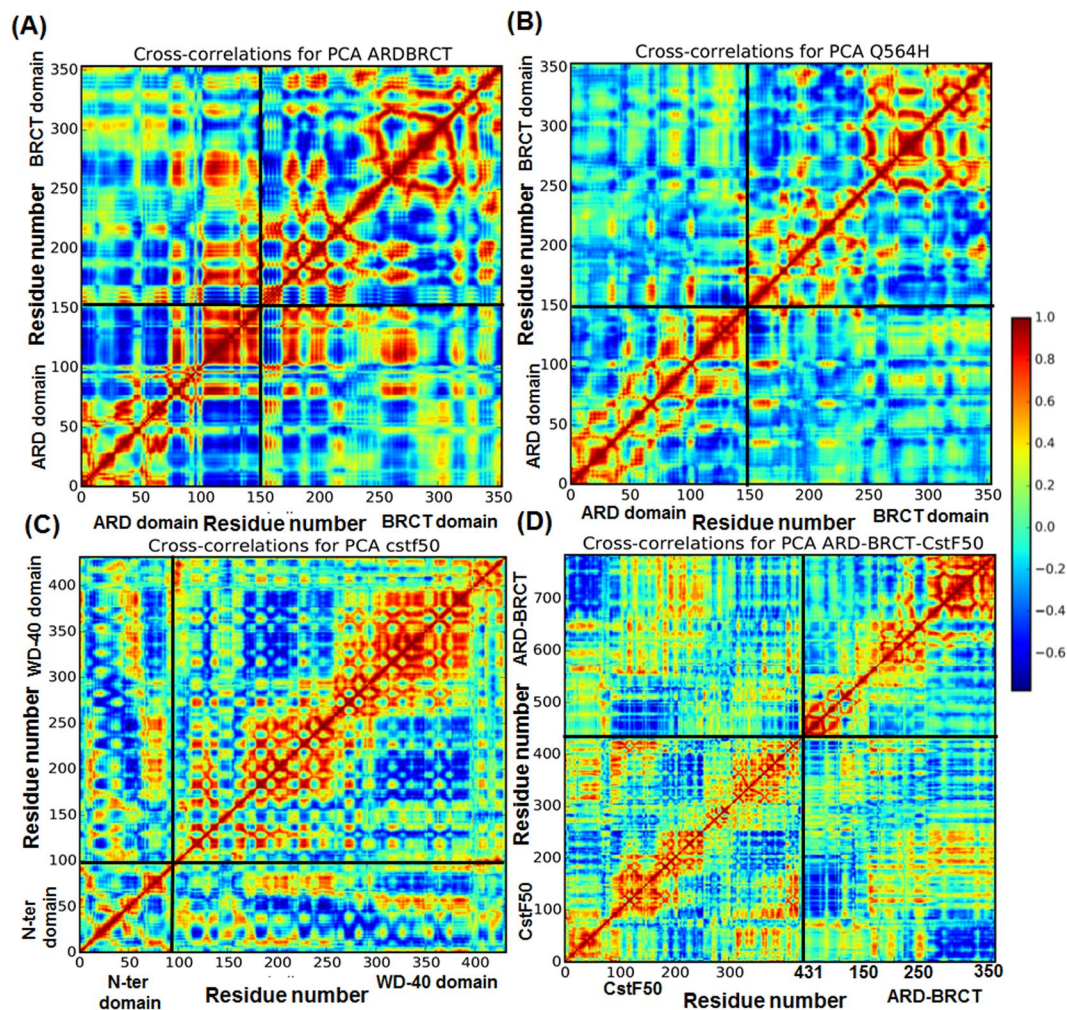


Figure 7. Comparative residual cross-correlation for PCA of (A) BARD1 ARD-BRCT *wild-type*, (B) Gln 564 His mutant, (C) CstF50 respectively and (D) BARD1 ARD-BRCT *wild-type* - CstF50 complex respectively. Red and yellow regions show highly positively correlated regions while dark blue represent anti-correlated movement of residues.

Conclusion

Multidisciplinary approaches were applied to characterize the domain motion and binding affinity between BARD1 ARD-BRCT, Gln 564 His mutant and CstF50 proteins. It has been observed that the BARD1 ARD-BRCT Gln 564 His mutant is not affecting the monomeric property, secondary structure, hydrodynamic radii and thermal stability of the *wild-type* protein. The limited proteolysis of *wild-type* and mutant confirms that ARD and BRCT domain are stable and independent folded domains of BARD1 connected by a short stretch of a linker region. These results from limited proteolysis are consistent with earlier NMR studies which conclude that in the absence of any interacting protein, the ARD and tandem BRCT region behave as an independent domain in solution^{40, 41}. The thermal unfolding study indicates that BARD1 ARD-BRCT *wild-type* and mutant unfold via a two state pathway. However, chemical unfolding of *wild-type* and mutant suggest that both proteins unfold via a molten globule intermediate.

The results from NMA and MD simulation substantiate well with the earlier reported short angle X-ray scattering (SAXS) findings. In solution ARD and BRCT domains sample a wide range of relative orientations with respect to one another⁴². The Small Angle X-ray Scattering (SAXS) for BARD1 ARD-BRCT domain shows two additional features in the $p(r)$ function: a shoulder at $r \sim 40$ Å and tailing-off of the at $r > 90$ Å⁴². SAXS for ARD-BARD1BRCT suggests that D_{max} of the domain is around 90 Å⁴². The RMSD values calculated for *wild-type* and mutant are very close (87 Å) to the experimental values ($r > 90$ Å)⁴². The distance calculated for *wild-type* N-terminal ARD domain and C-terminal BRCT domain shows a maximum change of 40 Å, which shows dynamic behavior of the protein due to the flexible linker region. The NMA and deformation analysis indicate that the linker connecting ARD and BRCT domains makes the *wild-type* and mutant protein flexible. MD simulation studies show that ARD domain is highly dynamic than the BRCT domain due to flexible linker which acts as the connecting bridge between the ARD and BRCT domain of BARD1. Flexibility in linker segments of the BARD1 ARD-BRCT backbone allow corresponding domain motions to occur with minor structural

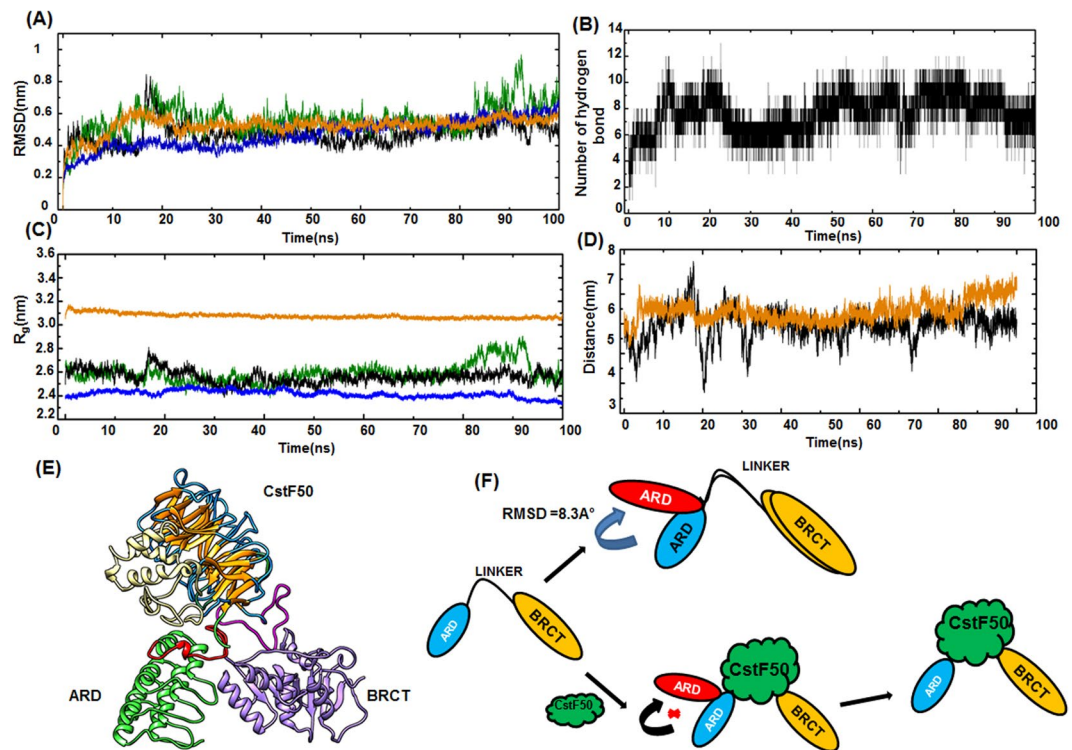


Figure 8. (A and C) RMSD and R_g profile of BARD1 ARD-BRCT *wild-type*, Gln 564 His mutant, CstF50 and *wild-type*-CstF50 complex respectively (B) H bonding profile of BARD1 ARD-BRCT *wild-type*-CstF50 complex (D) distance fluctuation profile of ARD domain in BARD1 ARD-BRCT *wild-type* and complex (E) Structure of *wild-type*-CstF50 complex and (F) Model for mechanism of *wild-type*-CstF50 complex stabilization.

perturbations. The RMSD, R_{gyr} and RMSF profiles show that BARD1 Gln 564 His mutant protein structure is more flexible than the *wild-type*. The relative orientation of ARD domain is significantly affected due to the Gln 564 His mutation. In BARD1 ARD-BRCT domain, flexible linker has different orientation that allows free twist and rotation of the ARD domain through space which can facilitates the recruitment of CstF50. The mutation Gln 564 His is present in the disordered flexible linker region of BARD1 ARD-BARD1BRCT domain. Lack of specific secondary structure provides the accessibility of different orientations to the linker region. The increase in the relative flexibility may be due to twist in the linker region by histidine. The linker region acts as a binding region for two large proteins CstF50 and p53. The difference in hydrogen bond donor group in glutamine and histidine may be the other reason in the loss of BARD1-CstF50 interactions. The results from *wild-type*, mutant and CstF50 docking study indicates that glutamine 564 to histidine mutation shows loss in hydrogen bonding interactions between *wild-type* and CstF50. To our conclusion, BARD1 Gln at 564 position plays a pivotal role in the BARD1 ARD-BRCT-CstF50 complex formation. The study further shows CstF50 binding provides rigidity to the ARD domain of BARD1. Therefore, flexibility in the linker region is immensely required to adopt such unique orientation to accommodate CstF50 protein.

Material and Methods

All the chemicals used in this study were purchased from Sigma-Aldrich, unless otherwise specified. The buffers were prepared in double distilled, 0.44 μM filtered Milli-Q water (Millipore, USA). Filtered 0.44 μM (Millipore, USA) protein and buffer solutions were degassed prior to use.

Protein expression and purification. BARD1 ARD-BRCT (425–777) cloned into a modified pET-6H vector (generous gift by Dr. Richard Baer, Institute for Cancer Genetics, Columbia University, USA). BARD1 ARD-BRCT *wild-type* domain was further PCR amplified and cloned into pGEX-KT vector (Amersham) in such a way that TEV protease site was incorporated between the GST and BARD1 ARD-BRCT. Site directed mutagenesis was performed to engineer the Gln 564 His mutation in BARD1 ARD-BRCT, and further sequenced for confirmation. BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant were expressed in *Escherichia coli* BL21 (DE3). The GST-BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant fusion protein were purified using affinity chromatography (GST-sepharose 4B column) in 50 mM Borate buffer, pH 9.0, 300 mM NaCl (Buffer A). The affinity purified proteins were treated with TEV to remove the fusion tag. Molecular exclusion chromatography was further performed to achieve high purity and homogenous monomeric population of BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant proteins. Superdex-75 column was used for the purification of the BARD1 ARD-BRCT *wild-type* protein. The Mutant protein was purified using the superdex-200 analytical column.

Human-CstF50 cDNA was bought from DNASU cDNA repository (Arizona state university) was PCR amplified and cloned in to pET-28a vector (Invitrogen). CstF50 cloned in pET28a vector was expressed in *Escherichia coli* BL21 (DE3) cells grown at 37 °C until absorbance at $\lambda = 600$ nm reached a value in between 0.6–0.8, followed by induction with 0.1 mM IPTG at 18 °C for 18 hrs. Most of the CstF50 was found as insoluble aggregates at 18 °C when induced with 0.1 mM IPTG. The insoluble aggregate of CstF50 was washed with buffer composition 50 mM, 500 mM NaCl, 5% glycerol, 1% sarkosyl, 1% SDS, pH 8 (Buffer B) by resuspending the pellet and again centrifuging at 18000 r.p.m. After washing, pellet was resuspended in 50 mM, 500 mM NaCl, 8 M urea, 5% glycerol (Buffer C), pH 8 for 5 hours to solubilize completely. Then 8M-urea-solubilized-CstF50 was refolded using buffer 50 mM Tris, 500 mM NaCl, 0.1% triton, 10% glycerol pH 7.2, 2 mM EDTA, 700 mM arginine, 100 mM KCl, 50 mM MgCl₂, 100 mM CaCl₂, and 5 mM DTT, pH 6.2 (Buffer D). In the refolding process we have added 10% glycerol and 0.5 M arginine hydrochloride to improve the refolding process^{43, 44}. To overcome the aggregate formation, high concentration of arginine was used to suppress aggregation of folding intermediates.

Circular - Dichroism spectroscopy. Circular dichroism spectra for BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant were collected using 0.1 cm path length sealed quartz cuvette on a JASCO J-715 spectropolarimeter (Jasco, Easton, MD), equipped with a JASCO PTC 348 WI temperature controller. BARD1 ARD-BRCT and Gln 564 His mutant at a concentration of 15 μ M were scanned in the far-UV ($\lambda = 180$ –260 nm) and near-UV range ($\lambda = 350$ –260 nm). Seven spectra were collected with 20 nm scan speed at a resolution of 1 nm and, response time of 1 s was averaged for each experiment. Blank spectra using buffer E (50 mM Borate buffer, pH 9.0, 150 mM NaCl) have been taken under the similar conditions and subtracted from the raw data. The spectra was recorded at the 25 °C. The results have been expressed terms of molar ellipticity $[\theta]$ (deg cm² dmol⁻¹).

Limited proteolysis. BARD1 ARD-BRCT *wild-type* and mutant at a concentration of 2 mg/ml were incubated with the 10 μ g/ μ l trypsin in different time intervals of 0, 5, 10, 30, 60 and 120 mins at 37 °C. After incubation, reaction was terminated by adding 1 mM PMSF (sigma Aldrich). Samples collected at different time intervals were analyzed over SDS-PAGE gel. BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant protein untreated with trypsin were taken as control.

Glutaraldehyde crosslinking. BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant at a concentration of 0.5 mg/ml in buffer A (50 mM Borate buffer, pH 9.0, 300 mM NaCl) was incubated with freshly prepared solution of glutaraldehyde (final concentration 0.1%) for 0, 1, 2, 3, 5, 10, 15 and 30 minutes at 37 °C. Crosslinking reaction was terminated by adding 5 μ l of 1 M Tris-HCl, pH 8.0, and the samples were analyzed on 12% SDS-PAGE gel.

Dynamic light scattering. Molecular size measurement and oligomeric behavior of BARD1 ARD-BRCT *wild-type* was performed using Malvern Zetasizer (Malvern). BARD1 ARD-BRCT *wild-type*, Gln 564 His mutant protein, and buffer were filtered (0.22 μ m), degassed thoroughly prior to every measurement. BARD1 ARD-BRCT *wild-type* protein and mutant protein at a concentration of 1 mg/ml were scanned at an interval of 5 minutes for 15 minutes. The scanning were performed at the 25 °C. The DLS experiments were repeated in three independent sets.

Thermal and chemical denaturation. Thermal denaturation was monitored by circular dichroism spectroscopy. BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant were allowed to unfold over the temperature ranges from 20 °C to 70 °C. Further fraction unfolded was calculated at each temperature by using millidegree ellipticity at λ_{222} and data was fitted in to a two state unfolding pathway. BARD1 ARD-BRCT *wild-type* and mutant Gln 564 His were allowed to unfold by incubating 2 μ M protein in different concentrations of GuHCl at 10 °C for 18 hrs. Chemical denaturation of BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant were performed at 10 °C and the fluorescence from frequently used fluorophores was monitored using fluorescence spectrophotometer (Horiba, USA) at excitation wavelength of $\lambda = 295$ nm. BARD1 ARD-BRCT *wild-type* and Gln 564 His mutant at a concentration of 2 μ M was mixed with the GuHCl. The concentration of GuHCl varied from 0 M to 6 M, while the protein concentration was fixed at 2 μ M. Fluorescence emission spectra were recorded over the range of wavelength $\lambda = 310$ –400 nm in order of increasing GuHCl concentration, and blank subtraction was done to increase signal to noise ratio.

Molecular dynamics simulation. The validated Model structures of BARD1 ARD-BRCT *wild-type*, mutant Gln 564 His, CstF50 and BARD1-CstF50 complex were subjected to MD simulations using GROMACS 4.5.5^{45–47} software. Protein were solvated with explicit solvent Single Point Charge water, in a cubic box which left 2 Å space nearby the solute. The counter ions in the form of Na⁺ and Cl⁻ ions were added to make the system electrically neutral. Further energy minimization by the steepest descents method was used, and system was coupled to an external bath using Berendsen's method. The reference temperature for all simulation was fixed at 300 K. LINC method was used to constrain all bonds^{48, 49}. OPLS-AA force field in GROMACS was used and grid type neighbor searching was done, and long range electrostatics was controlled using PME^{50, 51}. During simulations, the respective energy minimized structures were subjected to a position restrained MD Simulation for 100 ns with a time step of 2 fs. The trajectories were saved at every 2 ps to form respective trajectories. GROMACS built in tools were used to calculate RMSD (Root Mean Square Deviation), RMSF (Root Mean Square Fluctuations) and R_{gyr} (radius of gyration).

Normal mode, Deformation and fluctuation analysis. Normal mode, deformation and fluctuation analysis were performed to study the domain motion in BARD1 ARD-BRCT *wild-type* and mutant protein using R 3.2 bio3D package⁵². First five modes were calculated by building the hessian matrix, and further extracted to explain the relative fluctuation in ARD domain of *wild-type* and mutant protein.

Principal component analysis (PCA). PCA reduces the dimensionality of the MD trajectory data and provides a concise way to visualize, analyze and compare large-scale concerted motions observed over the course of the simulation. PCA was performed using GROMACS in-built tool *g_covar*, *g_anaeig* and covariance matrix was built after removing net translational or rotational motion of the system by fitting the coordinate data to a reference structure. Further, diagonalization of the symmetric $3N \times 3N$ covariance matrix was performed via eigenvector decomposition method and an orthogonal set of eigenvectors which are also called as functional “modes” were calculated with eigenvalues.

Molecular docking. BARD1 ARD-BRCT *wild-type*, Gln 564 His and CstF50 docking was performed using HADDOCK server 2.0 which uses data-driven method of docking, with support from an extensive variety of experimental data obtained from diverse biophysical and biochemical experimental method⁵³.

Binding energy calculation. In common term binding free of complex in solvent can be calculated by $\Delta G_{\text{Binding}} = G_{\text{complex}} - (G_{\text{protein}} + G_{\text{Ligand}})$. Free energy and other energy parameters of the complex were calculated by using molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method in GROMACS 4.5.5 package. *g_mmpbsa* tool was used to integrate high-throughput molecular dynamics simulation with the estimation of free energy of interaction⁵⁴.

References

- Wu, L. C. *et al.* Identification of a RING protein that can interact *in vivo* with the BRCA1 gene product. *Nat Genet* **14**, 430–440, doi:10.1038/ng1296-430 (1996).
- Yu, X. & Baer, R. Nuclear localization and cell cycle-specific expression of CtIP, a protein that associates with the BRCA1 tumor suppressor. *J Biol Chem* **275**, 18541–18549, doi:10.1074/jbc.M909494199M909494199 (2000).
- Li, M. & Yu, X. Function of BRCA1 in the DNA damage response is mediated by ADP-ribosylation. *Cancer Cell* **23**, 693–704 (2013).
- Jin, Y. *et al.* Cell cycle-dependent colocalization of BARD1 and BRCA1 proteins in discrete nuclear domains. *Proc Natl Acad Sci USA* **94**, 12075–12080 (1997).
- Yarden, R. I., Pardo-Reoyo, S., Sgagias, M., Cowan, K. H. & Brody, L. C. BRCA1 regulates the G2/M checkpoint by activating Chk1 kinase upon DNA damage. *Nat Genet* **30**, 285–289, doi:10.1038/ng837ng837 (2002).
- Baer, R. & Ludwig, T. The BRCA1/BARD1 heterodimer, a tumor suppressor complex with ubiquitin E3 ligase activity. *Curr Opin Genet Dev* **12**, 86–91, doi:S0959437X01002696 (2002).
- Khanna, K. K. & Jackson, S. P. DNA double-strand breaks: signaling, repair and the cancer connection. *Nat Genet* **27**, 247–254, doi:10.1038/85798 (2001).
- Wu-Baer, F., Lagazon, K., Yuan, W. & Baer, R. The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin. *J Biol Chem* **278**, 34743–34746, doi:10.1074/jbc.C300249200C300249200 (2003).
- Ruffner, H., Joazeiro, C. A., Hemmati, D., Hunter, T. & Verma, I. M. Cancer-predisposing mutations within the RING domain of BRCA1: loss of ubiquitin protein ligase activity and protection from radiation hypersensitivity. *Proc Natl Acad Sci USA* **98**, 5134–5139, doi:10.1073/pnas.08106839898/9/5134 (2001).
- Hashizume, R. *et al.* The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation. *J Biol Chem* **276**, 14537–14540, doi:10.1074/jbc.C000881200C000881200 (2001).
- Nazeer, F. I. *et al.* p53 inhibits mRNA 3' processing through its interaction with the CstF/BARD1 complex. *Oncogene* **30**, 3073–3083, doi:10.1038/ncr.2011.29onc201129 (2011).
- Irminger-Finger, I. *et al.* Identification of BARD1 as mediator between proapoptotic stress and p53-dependent apoptosis. *Mol Cell* **8**, 1255–1266, doi:S1097-2765(01)00406-3 (2001).
- Laufer, M. *et al.* Structural requirements for the BARD1 tumor suppressor in chromosomal stability and homology-directed DNA repair. *J Biol Chem* **282**, 34325–34333, doi:M70519820010.1074/jbc.M705198200 (2007).
- Fong, N., Bird, G., Vigneron, M. & Bentley, D. L. A 10 residue motif at the C-terminus of the RNA pol II CTD is required for transcription, splicing and 3' end processing. *EMBO J* **22**, 4274–4282, doi:10.1093/emboj/cdg396 (2003).
- Fong, N. & Bentley, D. L. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes Dev* **15**, 1783–1795, doi:10.1101/gad.889101 (2001).
- Cevher, M. A. *et al.* Nuclear deadenylation/polyadenylation factors regulate 3' processing in response to DNA damage. *EMBO J* **29**, 1674–1687, doi:emboj20105910.1038/emboj.2010.59 (2010).
- Kleiman, F. E. & Manley, J. L. Functional interaction of BRCA1-associated BARD1 with polyadenylation factor CstF-50. *Science* **285**, 1576–1579, doi:7797 (1999).
- Kleiman, F. E. *et al.* BRCA1/BARD1 inhibition of mRNA 3' processing involves targeted degradation of RNA polymerase II. *Genes Dev* **19**, 1227–1237, doi:19/10/122710.1101/gad.1309505 (2005).
- 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* **66**, 4561–4565, doi:66/9/456110.1158/0008-5472.CAN-05-3629 (2006).
- Li, L., Chen, R. & Weng, Z. RDOCK: refinement of rigid-body protein docking predictions. *Proteins* **53**, 693–707, doi:10.1002/prot.10460 (2003).
- Thai, T. H. *et al.* Mutations in the BRCA1-associated RING domain (BARD1) gene in primary breast, ovarian and uterine cancers. *Hum Mol Genet* **7**, 195–202, doi:ddb030 (1998).
- Sauer, M. K. & Andrulis, I. L. Identification and characterization of missense alterations in the BRCA1 associated RING domain (BARD1) gene in breast and ovarian cancer. *J Med Genet* **42**, 633–638, doi:42/8/63310.1136/jmg.2004.030049 (2005).
- Ghimanti, 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* **33**, 235–242 (2002).
- Lee, C. *et al.* Functional Analysis of BARD1 Missense Variants in Homology-Directed Repair of DNA Double Strand Breaks. *Human mutation* **36**, 1205–1214 (2015).
- Klonowska, K. *et al.* Analysis of large mutations in BARD1 in patients with breast and/or ovarian cancer: the Polish population as an example. *Scientific reports* **5** (2015).
- Kleiman, F. E. & Manley, J. L. The BARD1-CstF-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression. *Cell* **104**, 743–753, doi:S0092-8674(01)00270-7 (2001).
- Woody, R. W. Contributions of tryptophan side chains to the far-ultraviolet circular dichroism of proteins. *Eur Biophys J* **23**, 253–262 (1994).
- Myers, J. K., Morris, D. P., Greenleaf, A. L. & Oas, T. G. Phosphorylation of RNA polymerase II CTD fragments results in tight binding to the WW domain from the yeast prolyl isomerase Ess1. *Biochemistry* **40**, 8479–8486 (2001).
- Lakowicz, J. R. *Principles of fluorescence spectroscopy*. (Springer Science & Business Media, 2007).
- Ishibashi, M. *et al.* Is arginine a protein-denaturant? *Protein Expr Purif* **42**, 1–6, doi:10.1016/j.pep.2005.03.028 (2005).

31. Srinivas, V., Raman, B., Rao, K. S., Ramakrishna, T. & Rao Ch, M. Structural perturbation and enhancement of the chaperone-like activity of alpha-crystallin by arginine hydrochloride. *Protein Sci* **12**, 1262–1270, doi:10.1110/ps.0302003 (2003).
32. Fiser, A. & Sali, A. ModLoop: automated modeling of loops in protein structures. *Bioinformatics* **19**, 2500–2501 (2003).
33. Fiser, A., Do, R. K. & Sali, A. Modeling of loops in protein structures. *Protein Sci* **9**, 1753–1773, doi:10.1110/ps.9.9.1753 (2000).
34. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography* **66**, 12–21 (2009).
35. Kaplan, W. & Littlejohn, T. G. Swiss-PDB Viewer (Deep View). *Brief Bioinform* **2**, 195–197 (2001).
36. Yang, L., Song, G. & Jernigan, R. L. How well can we understand large-scale protein motions using normal modes of elastic network models? *Biophysical Journal* **93**, 920–929 (2007).
37. Brooks, B. & Karplus, M. Normal modes for specific motions of macromolecules: application to the hinge-bending mode of lysozyme. *Proceedings of the National Academy of Sciences* **82**, 4995–4999 (1985).
38. Paladino, A., Costantini, S., Colonna, G. & Facchiano, A. M. Molecular modelling of miraculin: Structural analyses and functional hypotheses. *Biochem Biophys Res Commun* **367**, 26–32, doi:10.1016/j.bbrc.2007.12.102 (2008).
39. Kurplus, M. & McCammon, J. Dynamics of proteins: elements and function. *Annual review of biochemistry* **52**, 263–300 (1983).
40. Irminger-Finger, I. & Jefford, C. E. Is there more to BARD1 than BRCA1? *Nat Rev Cancer* **6**, 382–391, doi:nrc187810.1038/nrc1878 (2006).
41. Fox, D. *et al.* Crystal structure of the BARD1 ankyrin repeat domain and its functional consequences. *Journal of Biological Chemistry* **283**, 21179–21186 (2008).
42. Edwards, R. A. *et al.* The BARD1 C-terminal domain structure and interactions with polyadenylation factor CstF-50. *Biochemistry* **47**, 11446–11456, doi:10.1021/bi801115g (2008).
43. Arora, D. & Khanna, N. Method for increasing the yield of properly folded recombinant human gamma interferon from inclusion bodies. *J Biotechnol* **52**, 127–133 (1996).
44. Buchner, J. & Rudolph, R. Renaturation, purification and characterization of recombinant Fab-fragments produced in *Escherichia coli*. *Biotechnology (NY)* **9**, 157–162 (1991).
45. Berendsen, H. J., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **91**, 43–56 (1995).
46. Hess, B., Kutzner, C., Van Der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of chemical theory and computation* **4**, 435–447 (2008).
47. Van Der Spoel, D. *et al.* GROMACS: fast, flexible, and free. *J Comput Chem* **26**, 1701–1718, doi:10.1002/jcc.20291 (2005).
48. Hess, B., Bekker, H., Berendsen, H. J. & Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *Journal of computational chemistry* **18**, 1463–1472 (1997).
49. Miyamoto, S. & Kollman, P. A. SETTLE: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of computational chemistry* **13**, 952–962 (1992).
50. Essmann, U. *et al.* A smooth particle mesh Ewald method. *The Journal of chemical physics* **103**, 8577–8593 (1995).
51. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *The Journal of chemical physics* **98**, 10089–10092 (1993).
52. Grant, B. J., Rodrigues, A. P., ElSawy, K. M., McCammon, J. A. & Caves, L. S. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* **22**, 2695–2696 (2006).
53. 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* **69**, 726–733 (2007).
54. Kumari, R., Kumar, R. & Lynn, A. g_mmpbsa A GROMACS Tool for High-Throughput MM-PBSA Calculations. *Journal of chemical information and modeling* **54**, 1951–1962 (2014).

Acknowledgements

Funding for this study was supported by Seed in Air grant from TMC. We thank DBT-BTIS facility at ACTREC and Mizoram University for providing necessary software to this study. We thanks M.V. hosur for critically reading the manuscript.

Author Contributions

R.K.C., M.Q.S., P.S.T. and N.G. performed the calculations, R.K.C., S.K.N. and A.K.V. designed the experiments and wrote the paper.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-03816-4

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017