




Cite this: *RSC Adv.*, 2018, 8, 34056

Biophysical evaluation to categorize pathogenicity of cancer-predisposing mutations identified in the BARD1 BRCT domain†

Rajan Kumar Choudhary,^{ae} M. Quadir Siddiqui,^{af} Nikhil Gadewal,^a Nachimuthu Senthil Kumar,^c Ekaterina S. Kuligina^d and Ashok K. Varma ^{*ab}

The BRCT domain of BARD1 (BARD1 BRCT) is involved in many cellular processes such as DNA damage repair (DDR) and cell-cycle checkpoint regulation. BARD1 BRCT performs tumor suppressor function by recruiting BRCA1 at DNA damage site *via* interactions with other DNA damage repair (DDR) proteins. Considering the importance of the BRCT domain in genomic integrity, we decided to evaluate reported mutations of BARD1 BRCT Cys645Arg, Val695Leu, and Ser761Asn for their pathogenicity. To explore the effect of the mutation on the structure and function, BARD1 BRCT wild-type proteins and the mutant proteins were studied using different biochemical, biophysical and *in silico* techniques. Comparative fluorescence, circular dichroism (CD) spectroscopy and limited proteolysis studies demonstrate the well-folded structural conformation of wild-type and mutant proteins. However, thermal and chemical denaturation studies revealed similarity in the folding pattern of BARD1 BRCT wild-type and Cys645Arg mutant proteins, whereas there was a significant loss in the thermodynamic stability of Val695Leu and Ser761Asn mutants. Molecular dynamics (MD) simulation studies on wild-type and mutant protein structures indicate the loss in structural integrity of mutants compared with the wild-type protein.

Received 2nd August 2018
Accepted 17th September 2018

DOI: 10.1039/c8ra06524a

rsc.li/rsc-advances

1 Introduction

BARD1 (BRCA1-associated Ring Domain protein) has been identified as an N-terminal binding partner of tumor suppressor protein BRCA1 (breast cancer-associated gene product 1).^{1–3} BRCA1/BARD1 N-terminal RING–RING domain heterodimer is an E3 ubiquitin ligase complex and has an indispensable role in tumor suppression.^{3–8} Mutations in the RING domain of BRCA1 that abolish the interactions with BARD1 have been found to be associated with breast and ovarian cancers, thus establishing BARD1 as a tumor suppressor.^{5,9} Up-regulated level of BARD1 after DNA damage,

hormone signaling, hypoxic condition,¹⁰ and induction of apoptosis by activation of caspase 3 in a p53-mediated fashion demonstrated the involvement of BARD1 in tumor suppression.^{11,12} In several breast, ovarian and uterine tumor cases, BARD1 was found to be either mutated or truncated.^{13–15} The BARD1 protein comprises 777 amino acids and different functional domains including one N-terminal ring domain region, one ankyrin repeat domain¹⁶ and two tandem BRCT repeat domains structurally similar to the BRCA1 BRCT domain.¹⁷

BRCT repeat domains are present in different DNA repair proteins, such as MDC1, BRCA1, BARD1, XRCC1, and 53BP1, and dynamically participate in DNA damage repair mechanism.^{18–24} BARD1 BRCT also has two conserved binding pockets, P₁ (hydrophilic pocket) and P₂ (hydrophobic pocket), for interaction with DNA damage repair protein in a phospho-dependent manner.²⁵ The role of BARD1 BRCT in DNA damage repair is crucial for early recruitment of BRCA1 to the DNA damage site.²⁶ Three cancer-predisposing mutations Cys645Arg, Val695Leu, and Ser761Asn are reported within the BARD1 BRCT domain^{15,27} in the breast, ovarian and uterine cancers.^{15,28,29} Recent studies have established that BARD1 BRCT is required for early recruitment of BRCA1 at the DNA damage site *via* directly interacting with PAR through the residues Cys645 and Val695.²⁶ In the crystal structure of BARD1 BRCT,³⁰ Ser761Asn mutation is located in very close proximity to the conserved P₂-binding pocket that may interfere with phosphopeptide binding at P₂ pocket.³⁰ A set of *in silico*

^aAdvanced Centre for Treatment, Research and Education in Cancer, Kharghar, Navi Mumbai, Maharashtra 410 210, India. E-mail: avarma@actrec.gov.in; Fax: +91-22-2740 5085; Tel: +91-22-2740 5112

^bHomi Bhabha National Institute, Training School Complex, Anushaktinagar, Mumbai - 400 094, India

^cDepartment of Biotechnology, Mizoram University (A Central University), Aizawl - 796 004, Mizoram, India

^dLaboratory of Molecular Oncology, Department of Tumor Growth Biology, N.N. Petrov Institute of Oncology, RU-197758, Pesochny-2, St.-Petersburg, Russia

^eDepartment of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, Edmond J. Safra Campus, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

^fUniversity of Nebraska Medical Centre, Omaha, NE, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8ra06524a



pathogenicity prediction tools such as Align-GVGD³¹ (Align-Grantham variation (GV) and Grantham deviation (GD)), I-mutant 3.0 (ref. 32) (prediction of protein stability changes upon single point mutation), and SNP&GO³³ (single amino acid polymorphisms using GO terms) have been used for the prediction of pathogenicity of mutations. Furthermore, to understand the abrogative effects of mutations, a combinatorial approach of biochemical, biophysical and *in silico* techniques were employed. This study will also help in understanding the role of BARD1 BRCT in DNA damage repair and future development of small molecule inhibitors, which can modulate or neutralize the pathogenic effects of mutations.

2 Results

2.1 Pathogenicity characterization of disease-associated mutants

The BARD1 BRCT Cys645Arg, Val695Leu, and Ser761Asn mutations are identified as cancer predisposing in nature (Fig. 1A, ESI†), and significantly affect the tumor-suppressor functions of the BARD1.¹⁵ The $\Delta\Delta G^\circ$ value predicted by I-mutant 3.0 (ref. 32) is tabulated in Table 1 (ESI†). It is observed that $\Delta\Delta G^\circ$ value predicted for BARD1 BRCT Cys645Arg mutant protein is $-0.29 \text{ kcal mol}^{-1}$, which is smaller than the set threshold of $-0.5 \text{ kcal mol}^{-1}$. Therefore, Cys645Arg mutation is classified as neutral without any pathogenicity. However, $\Delta\Delta G^\circ$ values for Val695Leu and Ser761Asn mutants are $-1.82 \text{ kcal mol}^{-1}$ and $-0.54 \text{ kcal mol}^{-1}$, respectively. Therefore, it can be predicted that Val695Leu and Ser761Asn mutations in the BARD1 BRCT domain may be pathogenic. The prediction is also supported by SNP&GO,³³ which indicates that except Cys645Arg substitution, Val695Leu and Ser761Asn substitutions have chances of disease predisposition (Table 1, ESI†). We also performed pathogenicity analysis Align-GVGD³¹ to characterize these mutants. Align-GVGD classifies BARD1 BRCT Cys645Arg mutant protein as most likely to be pathogenic and scores this protein as class 65, while Ser761Asn is scored as class 45 and Val695Leu is scored as class 25, classifying as less pathogenic (Table 1).

2.2 Oligomeric behavior of BARD1 BRCT mutants

To characterize the effect of the mutation on BARD1 BRCT, wild-type and mutant proteins were purified using Ni-NTA affinity chromatography (Fig. 1B, ESI†). The size exclusion chromatography profiles of BARD1 BRCT wild-type, and Cys645Arg, Val695Leu, and Ser761Asn mutants are shown in Fig. 1C (ESI†). As reported earlier, wild-type BARD1 BRCT is

monomeric in nature,³⁴ and comparative gel filtration elution profile of BARD1 BRCT wild-type and mutant proteins indicated that all the proteins elute at same elution volume, suggesting that mutations do not affect the monomeric property of the BARD1 BRCT protein (Fig. 1C, ESI†). DLS studies revealed that BARD1 BRCT wild-type and the mutants Cys645Arg, Val695Leu, and Ser761Asn had hydrodynamic radii in the range of 2.2–2.7 nm. These radii are comparable with the Stoke's radii of 2.4 nm obtained from the gel filtration chromatography calibrated against the known standards. The experimental hydrodynamic diameter obtained from DLS for wild-type was $4.2 \pm 0.19 \text{ nm}$ and for mutants, the hydrodynamic diameters were $4.88 \pm 0.32 \text{ nm}$ for Cys645Arg, $5.12 \text{ nm} \pm 0.67 \text{ nm}$ for Val695Leu and $5.35 \pm 0.43 \text{ nm}$ for Ser761Asn, which conclude that BARD1 BRCT and the mutants are monomeric in nature (Fig. 1A). This inference is also supported by SDS-PAGE glutaraldehyde cross-linking profiles of wild-type and mutant proteins. As there was no sign of any oligomeric species observed for wild-type and mutant proteins, mutations did not change the monomeric property of BARD1 BRCT (Fig. 2, ESI†).^{25,30}

2.3 Mutants have structures similar to wild-type

Far- and near-UV range CD (circular dichroism spectroscopy) spectra for BARD1 BRCT wild-type and Cys645Arg, Val695Leu, and Ser761Asn mutants are shown in Fig. 1B and C. The ellipticity minima at $\lambda = 222 \text{ nm}$ and $\lambda = 208 \text{ nm}$ indicate the presence of predominantly α -helical structures. The ellipticity maximum rise at $\lambda = 218 \text{ nm}$ indicates the presence of β -strands, and both these features are in good agreement with the crystal structure of BARD1 BRCT (PDB ID: 2NTE, Fig. 1A, ESI†). The far-UV spectra of wild-type and mutant proteins do not show any large change in the ellipticity pattern, indicating that mutations are not drastically changing the secondary structure of the protein. Near-UV spectra of BARD1 BRCT wild-type and mutant proteins are shown in Fig. 1C. All the proteins including wild-type and mutant proteins show a characteristic peak at $\lambda = 295 \text{ nm}$ and a plateau from $\lambda = 260\text{--}280$ (cumulative signal for Trp, Phe, Tyr amino acids), which is an indication of the hydrophobic micro-environment around the aromatic residues in the structure (Fig. 1C). Tryptophan fluorescence was also measured in the presence of GuHCl to probe the compactness of tertiary structural differences between BARD1 BRCT wild-type and mutant proteins. Emission maxima recorded for native BARD1 BRCT wild-type, and Cys645Arg, Val695Leu, and Ser761Asn mutant proteins were at $\lambda = 332 \text{ nm}$, 333 nm , 333 nm and 335 nm , respectively. However, a shift in the emission maxima is highest for the Ser761Asn mutant, which is very close

Table 1 *In silico* prediction of pathogenicity of BARD1 BRCT mutants. Comparative prediction output from different prediction servers indicates that mutation may have structural and functional effects on the BARD1 BRCT domain

Mutation	I-mutant prediction	$-\Delta\Delta G^\circ$	SNP&GO prediction	Align-GVGD
Cys645Arg	Decreased stability	$-0.29 \text{ kcal mol}^{-1}$	Neutral	Deleterious (class C65)
Val695Leu	Decreased stability	$-1.82 \text{ kcal mol}^{-1}$	Disease	Neutral (class C35)
Ser761Asn	Decreased stability	$-0.54 \text{ kcal mol}^{-1}$	Disease	Deleterious (class C65)



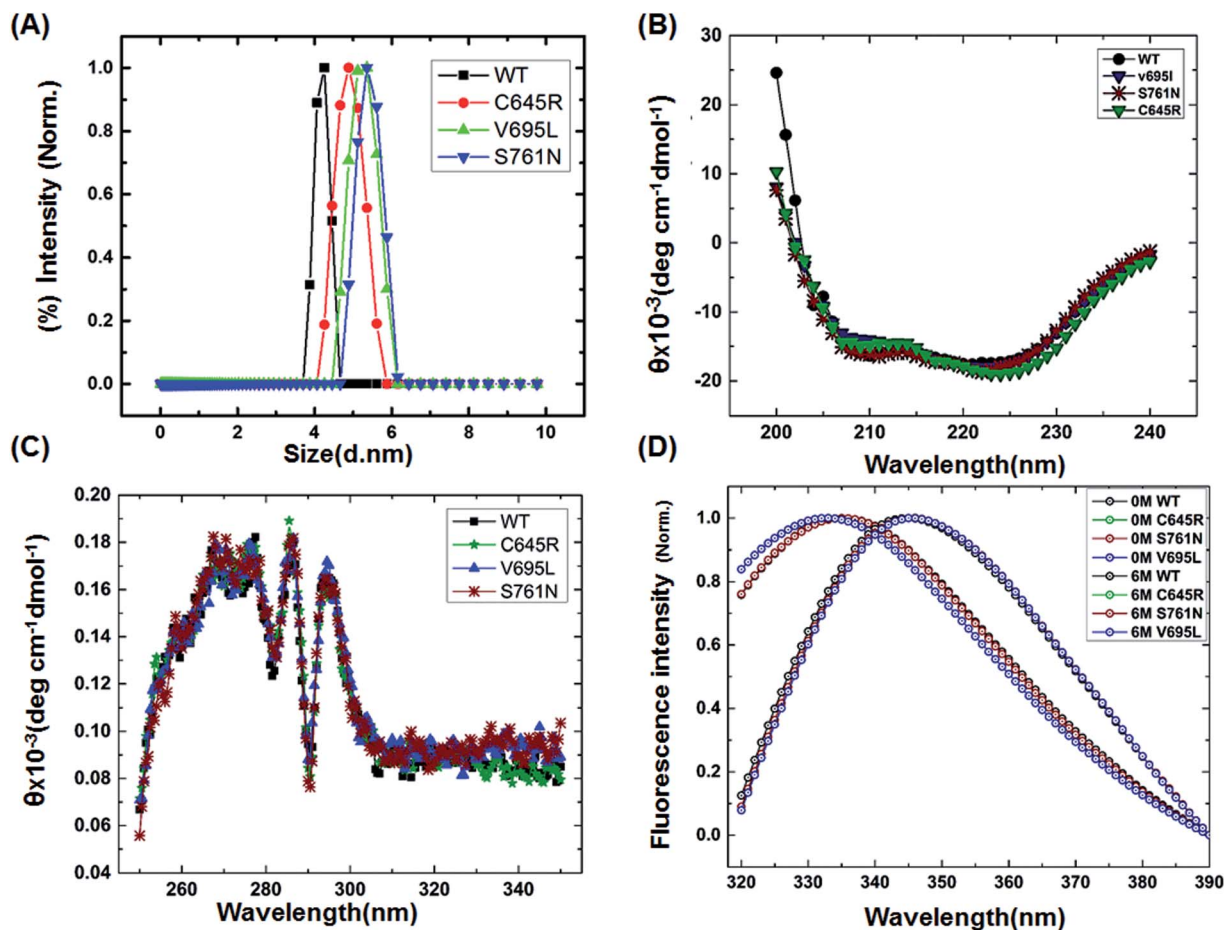


Fig. 1 Structural characterization of BARD1 BRCT wild-type and mutant proteins. (A) Dynamic light-scattering profile of the BARD1 BRCT wild-type and mutant proteins. (B) Secondary structure analysis of BARD1 BRCT wild-type and mutant proteins using far-UV, CD spectroscopy. (C) Tertiary structure characterization of BARD1 BRCT wild-type and mutant proteins using near-UV, CD spectroscopy. (D) Chemical denaturation of BARD1 BRCT wild-type and mutant proteins using GuHCl.

to that of the Trp762 fluorophore. BARD1 BRCT wild-type and mutant proteins show emission maxima at $\lambda = 348$ nm for the unfolded state at 6 M GuHCl (Fig. 1D). Furthermore, to explore the effect of the mutation on BARD1 BRCT tertiary structure, limited proteolysis was performed, and the results are shown in Fig. 2. Limited proteolysis profile of BARD1 BRCT wild-type and mutant proteins show similar resistivity profiles for trypsin (Fig. 2A–D) and chymotrypsin digestion, unravelling the facts that BARD1 BRCT wild-type is a stable domain and the mutations brought neither large tertiary structural changes nor overall packing of mutant proteins (Fig. 2E–H).

2.4 Thermodynamics and unfolding

It has been reported that BRCT domains unfold *via* intermediate molten globule formation.^{25,34–37} Thermal denaturation profiles of BARD1 BRCT wild-type and Cys645Arg, Val695Leu and Ser761Asn mutants are shown in Fig. 3A. It is observed that BARD1 BRCT wild-type and Cys645Arg have higher thermal stability compared with Val695Leu and Ser761Asn mutants. BARD1 BRCT wild-type and mutant proteins unfold *via* a two-state transition. However, Ser761Asn mutant protein lost the

cooperativity and T_m significantly (Fig. 3A). The T_m 's values calculated for BARD1 BRCT wild-type, Cys645Arg, Val695Leu and Ser761Asn assuming a two-state transition model^{34,38,39} were 47.7 ± 0.85 °C, 47.6 ± 0.65 °C, 42.3 ± 0.78 °C, and 41.1 ± 0.42 °C, respectively. ΔG_{N-D} calculated for BARD1 BRCT wild-type, Cys645Arg, Val695Leu, and Ser761Asn mutant proteins were 9.8 ± 0.39 kcal mol⁻¹, 9.6 ± 0.11 kcal mol⁻¹, 7.4 ± 0.69 kcal mol⁻¹ and 7.1 ± 0.12 kcal mol⁻¹, respectively (Fig. 3A). The chemical denaturation profiles of BARD1 BRCT wild-type and mutant proteins at 10 °C are shown in Fig. 3B. The chemical denaturation profile shows that BARD1 BRCT wild-type and mutant proteins completely unfold at a concentration of 6 M GuHCl, following a three-state unfolding pathway with an intermediate formation at 2.2 M GuHCl (Fig. 3B). The ΔG_{N-I-D} values calculated for BARD1 BRCT wild-type, Cys645Arg, Val695Leu, and Ser761Asn mutant proteins were 6.94 ± 0.23 kcal mol⁻¹, 6.28 ± 0.82 kcal mol⁻¹, 6.2 ± 0.45 kcal mol⁻¹, and 6.11 ± 0.18 kcal mol⁻¹, respectively. The near-UV spectra for wild-type and mutant proteins are shown in Fig. 4A–D. At 10 °C, BARD1 BRCT wild-type and mutant proteins are present in fully folded native conformation. However, with the increase in temperature to 45 °C, BARD1 BRCT Val695Leu and



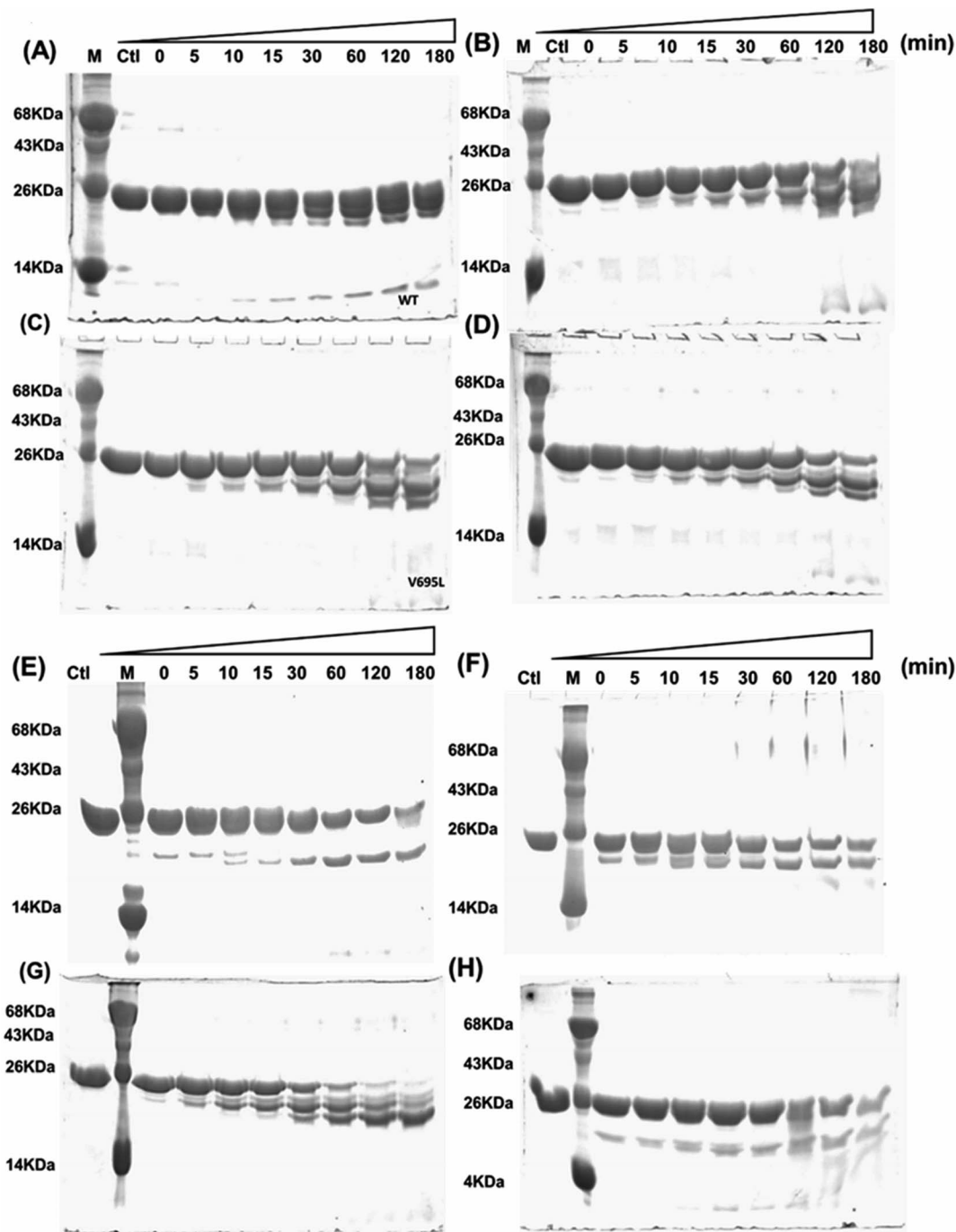


Fig. 2 Limited proteolysis of BARD1 BRCT wild-type and mutant proteins. (A–D) are trypsin digestion profiles of wild-type, Cys645Arg, Val695Leu and Ser761Asn mutant, respectively. (E–H) are chymotrypsin digestion profiles of wild-type, Cys645Arg, Val695Leu, and Ser761Asn mutant, respectively. Control (Ctl) shows untreated trypsin and chymotrypsin sample of wild-type and mutant proteins.

Ser761Asn mutant proteins show reduced CD signal compared with the wild-type protein (Fig. 4C and D), which is an indication of significant loss in the tertiary structure of mutant proteins (Fig. 4A–D). However, BARD1 BRCT wild-type and the BARD1 BRCT Cys645Arg mutant partially lose their tertiary structure at 45 °C (Fig. 4A and B), providing an indication that Ser761Asn and Val695Leu mutant proteins have lost their thermal stability compared with the BARD1 BRCT wild-type and Cys645Arg mutant protein.

2.5 Variable temperature fluorescence spectroscopy

Fluorescence spectra were recorded from 15 °C to 60 °C at 3 °C intervals, for BARD1 BRCT wild-type and the mutant proteins (Fig. 3, ESI†). The high fluorescence emission intensity recorded at 15 °C indicates complete burial of tryptophan in the hydrophobic core of the protein.⁴⁰ However, as the temperature increases up to 30 °C, the loss in fluorescence intensity for wild-type and mutant proteins suggests the restricted opening of the compact structure due to partial exposure of tryptophan to the



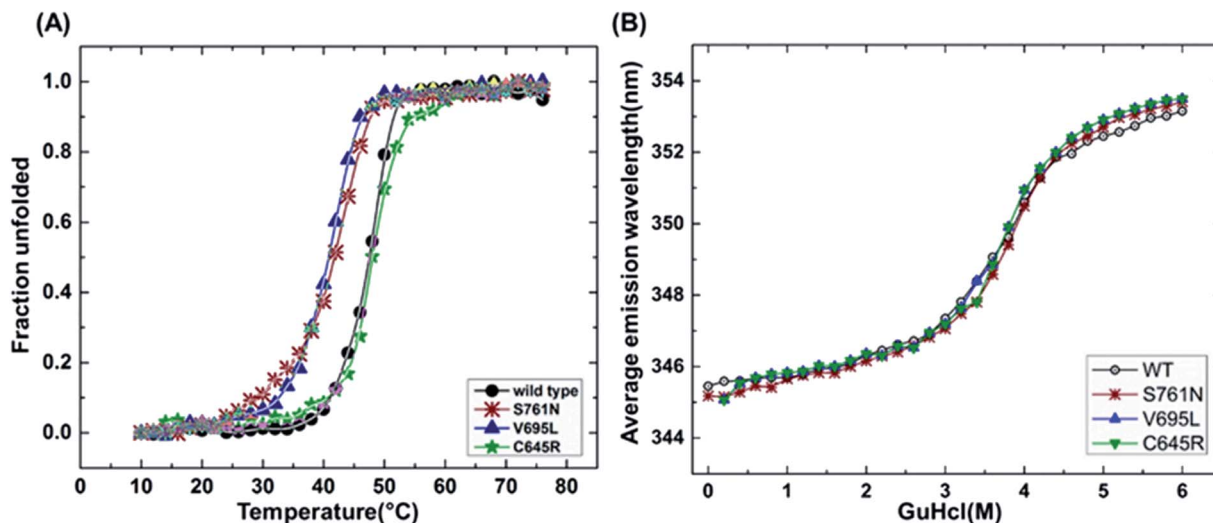


Fig. 3 Thermal and chemical denaturation profiles of BARD1 BRCT wild-type and mutant proteins. (A) Comparative overlay of fraction unfolded of BARD1 BRCT wild-type and Cys645Arg, Val695Leu and Ser761Asn mutants. (B) Chemical denaturation profiles of BARD1 BRCT wild-type and Cys645Arg, Val695Leu and Ser761Asn mutants.

polar environment. Furthermore, no significant decrease in the fluorescence intensity was observed at 45 °C for wild-type and Cys645Arg mutant proteins. However, at 45 °C, Ser761Asn and Val695Leu mutants show the increase in fluorescence intensity and blue-shifting in emission maxima as compared with wild-type and Cys645Arg mutant proteins. At 50 °C, the

fluorescence intensity was maximum for wild-type and mutant proteins with blue-shift in emission maxima, indicating burial of tryptophan fluorophores in the hydrophobic core of the protein for the formation of a molten globule structure. The higher fluorescence intensity and blue-shift at 45 °C for Val695Leu and Ser761Asn mutant indicate molten globule

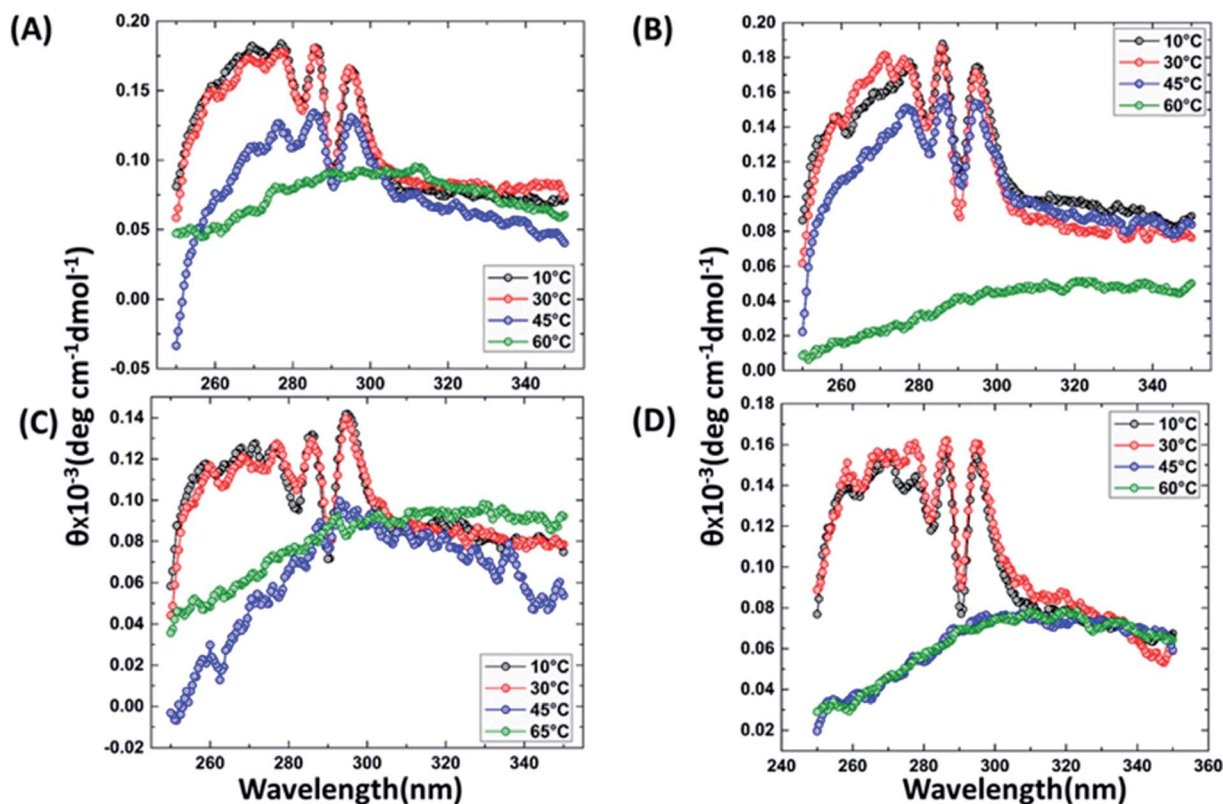


Fig. 4 Comparative thermal denaturation profile of BARD1 BRCT wild-type and mutant proteins in near-UV range. Thermal denaturation profile of BARD1 BRCT (A) wild-type, (B) Cys645Arg, (C) Ser761Asn and (D) Val695Leu.



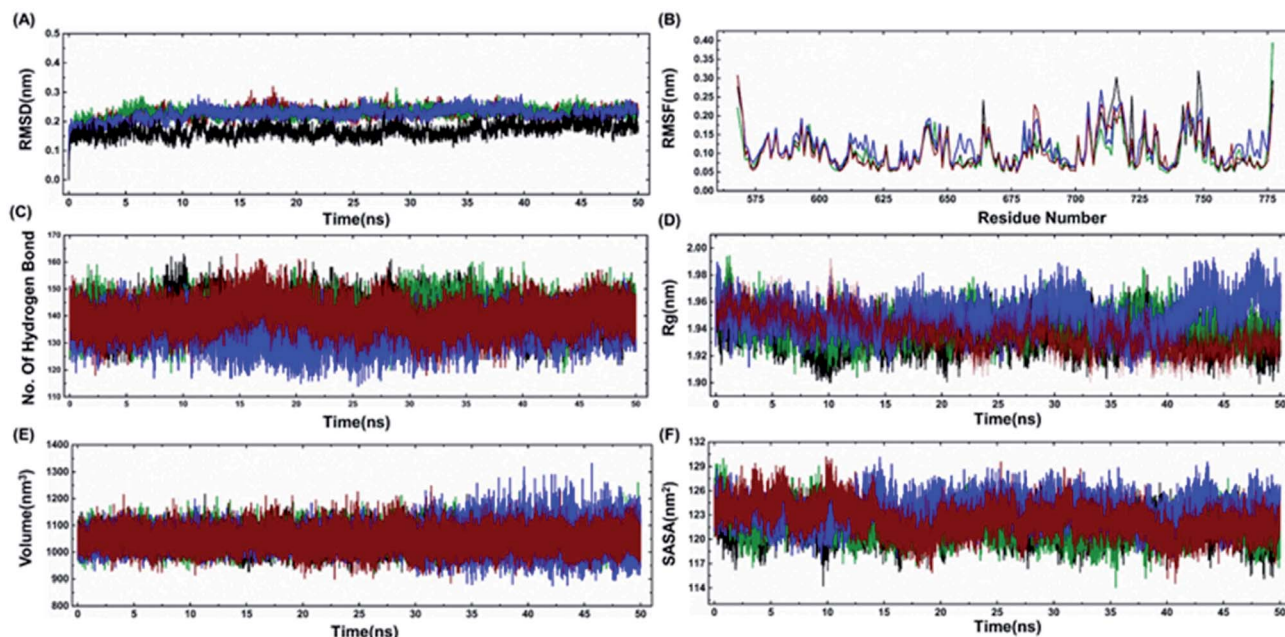


Fig. 5 Comparative representation of MDS trajectory for wild-type and mutant proteins. Comparative profile of (A) RMSD, (B) RMSF, (C) number of hydrogen bonds, (D) radius of gyration, (E) volume (nm^3), and (F) SASA (nm^2) of wild-type, Cys645Arg, Val695Leu and Ser761Asn (black color represents wild-type, green color indicates Cys645Arg profile, blue color indicates Val695Leu and red color line represents Ser761Asn mutant protein-structure profile).

formation at lower temperature compared with Cys645Arg and wild-type protein (Fig. 3, ESI[†]). The changes in the emission maximum of wild-type and mutant proteins are tabulated in Table 2 (ESI[†]).

2.6 Molecular dynamics simulation analysis of BARD1 BRCT and mutants

The model structures of all three cancer predisposing mutants Cys645Arg, Val695Leu and Ser761Asn reported in the BARD1 BRCT region were prepared *via* introducing mutations in the amino acid sequences of the wild-type protein using a Swiss PDB viewer.⁴¹ These structures were subjected to molecular dynamics simulation (MD) under periodic boundary conditions. After the equilibration period, the comparative RMSD (root-mean-square deviation) profile of BARD1 BRCT wild-type and mutant proteins indicated that the wild-type structure shows stable RMSD and trajectory compared with mutants throughout the simulation period. BARD1 BRCT Val695Leu mutant shows RMSD fluctuation in profiles similar to Cys645Arg and Ser761Asn mutant structures (Fig. 5A). The comparative RMSF (root-mean-square fluctuation) analysis indicates that there is a local structural change in mutant proteins compared with the BARD1 BRCT wild-type protein (Fig. 5B). The weak intramolecular hydrogen bonding profile indicates that except Cys645Arg mutation, BARD1 BRCT Val695Leu and Ser761Asn mutations have less number of hydrogen bonds as compared with the wild-type protein (Fig. 5C and 4, ESI[†]). Fluctuation in the radius of gyration (R_g) provides a measure of molecular compactness of the dynamic system (Fig. 5D). The R_g fluctuation was found to be higher in case of

BARD1 BRCT Val695Leu mutant structure, which has a positive correlation with the reduced number of intra-molecular hydrogen bonds (Fig. 5C and 4, ESI[†]). Comparative volume change observed for BARD1 BRCT wild-type and mutant protein structures clearly indicate that BARD1 BRCT Val695Leu substitution has destabilized the hydrophobic core in the mutant protein structure (Fig. 5E). The solvent accessible surface area (SASA) calculated after 50 ns of MD production run indicates that BARD1 BRCT Val695Leu mutant protein shows more exposed solvent accessible surface area compared with other mutants and wild-type protein (Fig. 5F) (Table 2).

2.7 Identifying the conformational fluctuation by principal component analysis (PCA)

Principal component analysis was performed for BARD1 BRCT wild-type and mutant protein trajectories after diagonalization of the covariance matrix to obtain eigenvalues and eigenvectors (Fig. 6). The trace of covariance matrix were 64.97 nm^2 , 58.49 nm^2 , 57.22 nm^2 , and 53.89 nm^2 for Val695Leu, Ser761Asn, Cys645Arg and wild-type proteins, respectively. Val695Leu

Table 2 Results obtained from fluorescence spectroscopy thermal denaturation showing change in emission maximum at different temperatures

	WT	Cys645Arg	Val695Leu	Ser761Asn
15 °C	332 nm	333 nm	333 nm	335 nm
30 °C	330 nm	333 nm	333 nm	333 nm
45 °C	328 nm	331 nm	331 nm	331 nm
50 °C	329 nm	330 nm	329 nm	329 nm



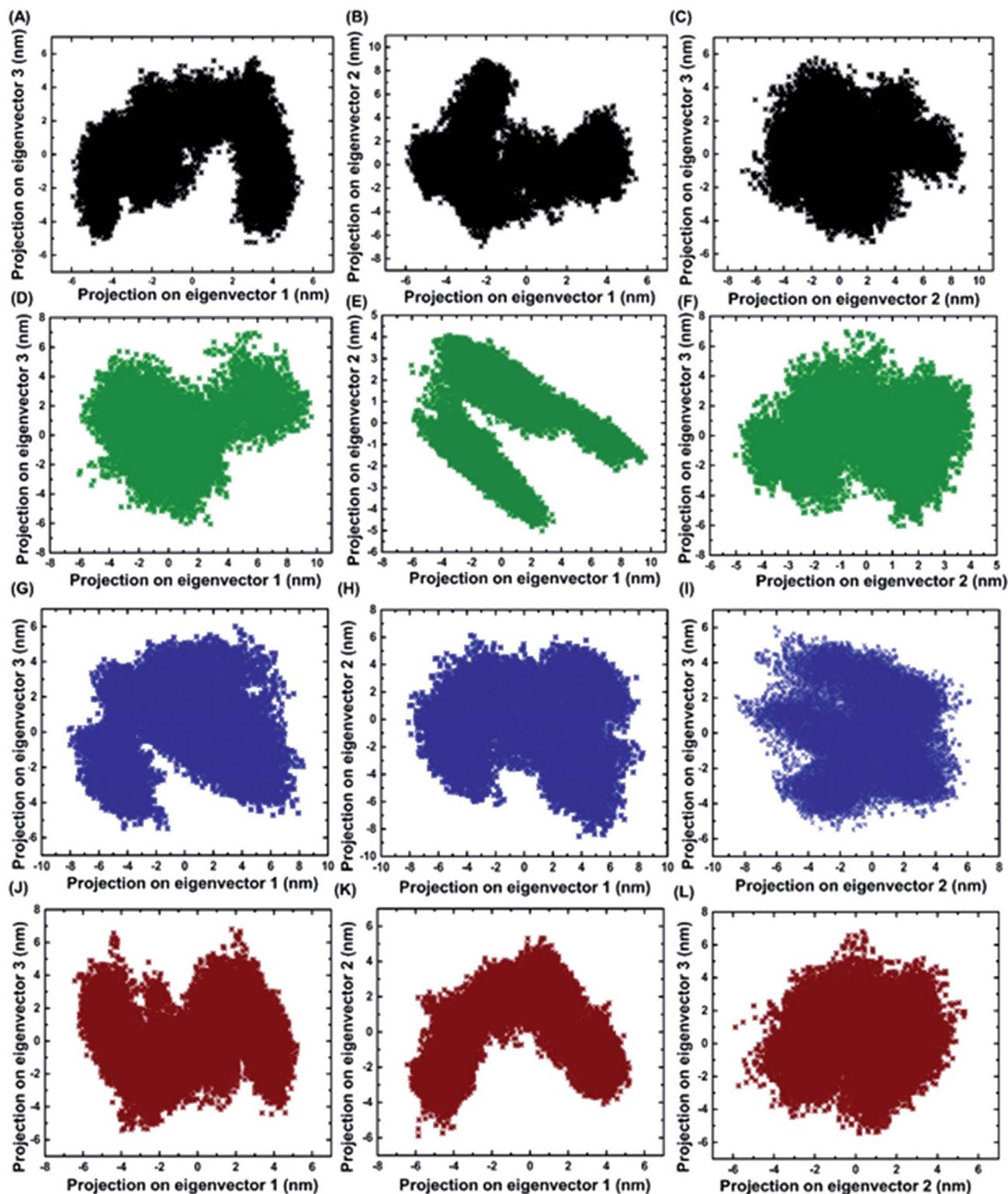


Fig. 6 Projections of first three eigenvectors for wild-type and mutant proteins. (A) Projection of eigenvector 3 on 1, (B) 2 on 1 and (C) 3 on 2 for wild-type. (D) Projection of eigenvector 3 on 1, (E) 2 on 1 and (F) 3 on 2 for Cys645Arg. (G) Projection of eigenvector 3 on 1, (H) 2 on 1 and (I) 3 on 2 for Val695Leu. (J) Projection of eigenvector 3 on 1, (K) 2 on 1 and (L) 3 on 2 for Ser761Asn.

shows the highest trace of covariance matrix, implying greater structural flexibility compared with wild-type, Ser761Asn and Cys645Arg mutant proteins (Table 3, ESI^{\dagger}). The first three eigenvectors having the highest corresponding eigenvalues

cover 90% of atomic motion attained by wild-type and mutant proteins. Fig. 5 (ESI^{\dagger}) shows that eigenvalue change for Val695Leu mutant for the first two eigenvectors is the highest as compared with wild-type and mutant proteins. Fig. 6A–L show



Table 3 Trace of covariance matrix obtained for BARD1 BRCT wild-type and mutant proteins after PCA

BARD1 BRCT and mutants	Trace of covariance matrix
WT	53.89 nm ²
Cys645Arg	57.22 nm ²
Val695Leu	64.97 nm ²
Ser761Asn	58.49 nm ²

the projection of trajectory on to first three eigenvectors for wild-type and mutant proteins. BARD1 Val695Leu protein structure shows the highest periodic transition in tertiary structure when projected on eigenvectors 2 and 1, 3 and 1 and 3 and 2 compared with wild-type and other mutant protein structures, demonstrating the higher structural flexibility (Fig. 6G–I). The wild-type (Fig. 6A–C), Cys645Arg (Fig. 6D–F) and Ser761Asn (Fig. 6J–L) show at least two different tertiary structure clusters. However, Val695Leu shows three different tertiary structural clusters, indicating higher structural dynamics (Fig. 6I). Fig. 6 (ESI†) shows displacements of components of

the wild-type and mutant proteins for the eigenvectors 1 and 2. The projection on first two eigenvectors of the residues for wild-type shows fluctuation in the c-terminal loop regions and the connecting loop between N-terminal BRCT and C-terminal BRCT domain, which are consistent with RMSF structure obtained (Fig. 7A, ESI†). In case of Cys645Arg mutant, projection on eigenvectors 1 and 2 of the residue indicates high fluctuation at 568–610 and 740–750 amino acid residue regions, which is also reflected in the RMSF structure (Fig. 6C and 7B, ESI†). Furthermore, projection on eigenvectors 1 and 2 in the structures of residues for BARD1 BRCT Val695Leu mutant protein specifies that the mutational effect on the residual fluctuation is not localized (Fig. 6E and F, ESI†). The RMSF structure for BARD1 BRCT Val695Leu also indicates that mutation has affected the structure at the global level (Fig. 7C, ESI†). Projection on eigenvectors 1 and 2 of residues for BARD1 BRCT Ser761Asn mutant protein structure reveals fluctuation in connecting loops at C-terminal BRCT (Fig. 6G and H and 1A, ESI†), which shows consistency with the RMSF structure (Fig. 7D, ESI†) (Table 3).

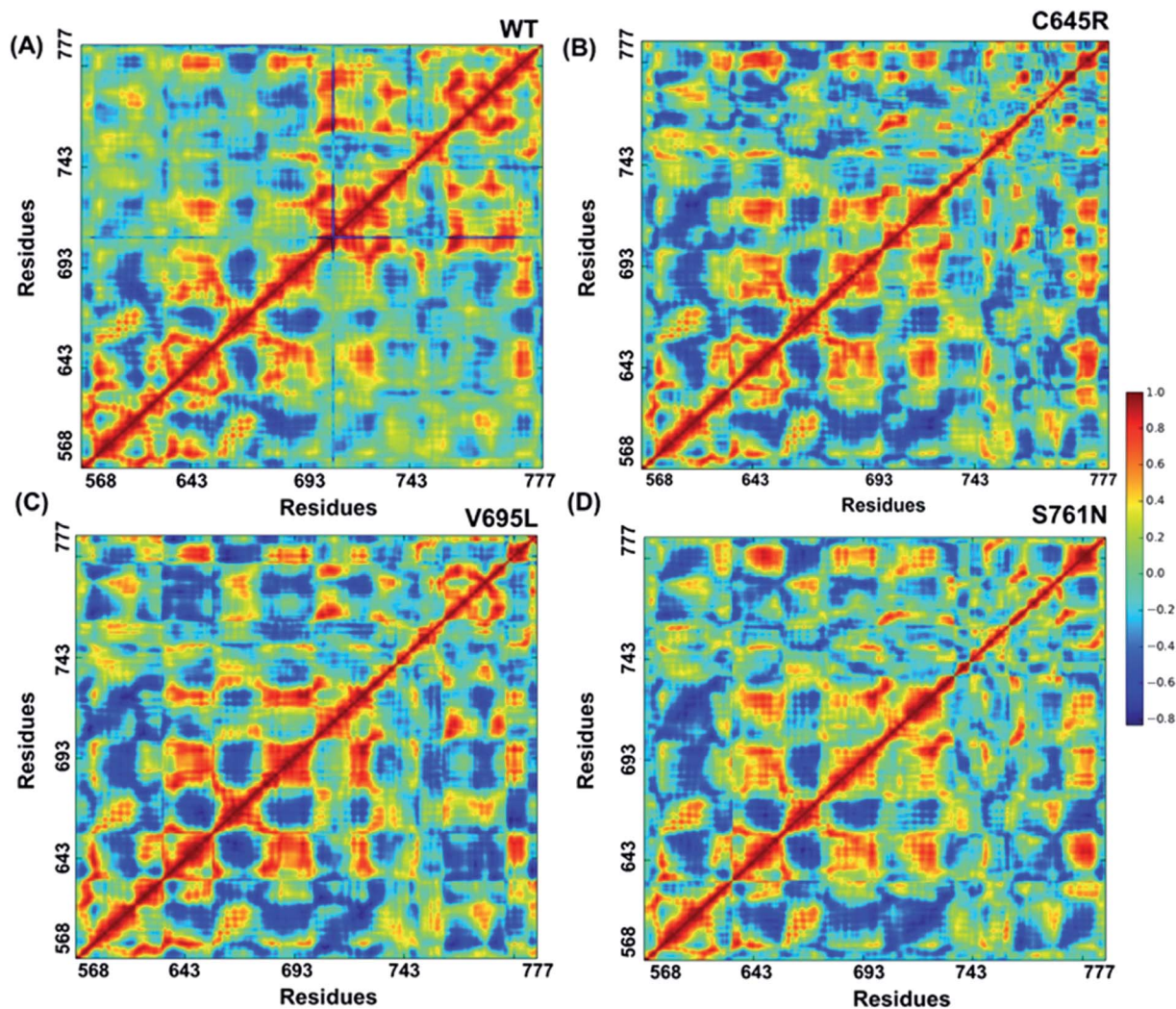


Fig. 7 Cross-correlations for PCA of wild-type and mutant proteins. (A) Cross-correlation for PCA of wild-type, (B) Cys645Arg, (C) Val695Leu and (D) Ser761Asn mutant proteins.



2.8 Covariance analysis

Cross-correlation matrixes of fluctuations for C-alpha atoms around their mean positions for wild-type and mutant proteins are represented in Fig. 7. The BARD1 BRCT wild-type protein shows a positive correlation within the N-terminal and C-terminal residues, but most of the N-terminal residues show a negative correlation with the C-terminal residues (Fig. 7A). When compared with wild-type, Cys645Arg and other mutant proteins show loss of positive residual correlation within N- and C-terminal BRCT residues (Fig. 7B–D). However, BARD1 BRCT Val695Leu protein structure shows a moderate decrease in the negative correlation within the N-terminal BRCT residues and increased positive correlation in atomic motion in C-terminal BRCT residues as compared with Cys645Arg and Ser761Asn (Fig. 7C). The BARD1 BRCT Ser761Asn mutant structure also shows localized changes in the positive correlation in the N-terminal (610–650 aa) and C-terminal (760–770 aa) BRCT region (Fig. 7D). These transformations reflect the difference in the internal dynamics of mutants in comparison with the wild-type structure.

2.9 Secondary structure characterization by DSSP

Comparative DSSP (dictionary of secondary structure of proteins) characterization of secondary structure has been presented in Fig. 8A–D (ESI†). It may be observed that the mutations have a local effect on the secondary structure of the BARD1 BRCT wild-type protein. Mutations Cys645Arg and Ser761Asn are present on the surface of BARD1 BRCT, while the mutation Val695Leu is present in the hydrophobic core of the protein. Though Cys645 substitution to arginine changes the chemical character of the amino acid from hydrophobic to basic, because of its presence in a loop region of BARD1 BRCT, it has no effect on the secondary structure of the protein. This observation is consistent with the results of CD spectroscopic analysis (Fig. 8A–D, ESI†). The BARD1 BRCT Val 695 Leu substitution is presented in the α 2 helix (Fig. 1A, ESI†), and the structure also shows no significant change in the structure as compared with the wild-type in CD spectroscopy as well as DSSP analysis. This intriguing change can be explained on the basis of the α -helix propensity of the amino acids.⁴² The BARD1 BRCT Ser761Asn mutation is present in the α 3 helix of BARD1 BRCT (Fig. 1A, ESI†) and does not show a drastic secondary structural change in the DSSP analysis. The propensity of alpha-helix formation by serine is higher than that by asparagine, which creates a possibility of losing the secondary structure. However, in the presence of strong neighboring α -helix former amino acids, the balance between enthalpic effects and conformational entropy is compensated.⁴² Therefore, even after serine-to-asparagine substitution, we did not observe any drastic secondary structure transformation in the CD spectroscopy as well as DSSP analysis.

2.10 Intra-molecular interactions analysis

BARD1 Cys645 in the wild-type structure was found to have non-covalent interactions with Leu592 and Leu570. However, in the molecular model of BARD1 BRCT Cys645Arg mutant structure, Arg645 shows loss of non-covalent interactions with Leu592 and gain of hydrogen bonding interactions with

Leu592 (Fig. 9A and B, ESI†). Similarly, in BARD1 BRCT wild-type, Val695 shows a hydrogen-bonding interaction with Gly698, Thr696, Ile692, and Leu691. Moreover, Val695 is also involved in the non-covalent interaction with Phe577, Phe763, Gly698, Lys693, and Ala697. In Val695Leu mutant, Leu695 shows one less hydrogen bonding interaction by limiting its interaction with Gly699, Gly700 and Leu691. Leu695 in mutant structure shows non-hydrogen bonding interaction similar to wild-type structure (Fig. 9C and D, ESI†). Ser761 in BARD1 BRCT wild-type is found to have very few interacting residues in its neighborhood. It forms a hydrogen bond with Ile764 and non-covalent interaction with Pro759, Asp741, and Phe763 (Fig. 9E, ESI†). Interestingly, in mutant structure, Asn761 forms a hydrogen bond with Asp741 and non-covalent interaction with Pro759 and Phe763 (Fig. 9F, ESI†). Loss of hydrogen bonding interaction with Ile764, which is an important residue suspected to interact with the phosphopeptide ${}_s\text{P-x-x-t-F}$, may lead to alteration of hydrophobic pocket architecture, which can lead to loss of phosphopeptide docking site for DNA damage repair proteins.

3 Discussion

In silico, biophysical and biochemical tools were used to demonstrate the pathogenicity of BARD1 Cys645Arg, Val695Leu, and Ser761Asn mutations. *In silico* prediction from I-mutant 3.0, SNP&GO and Align-GVGD servers clearly indicated that these mutations might be pathogenic in nature. Gel filtration, DLS, and chemical crosslinking profiles of the protein suggest that the monomeric property of BARD1 BRCT domain is not affected by mutations. The secondary structure investigation using CD spectroscopy indicates that these mutations do not drastically alter the secondary and tertiary structural composition of the BARD1 BRCT domain. Thermal denaturation studies using CD spectroscopy suggests that BARD1 BRCT wild-type and BARD1, Ser761Asn, Cys645Arg and Val695Leu mutants unfold *via* a two-state pathway. The T_m values suggest that Val695Leu and Ser761Asn mutants lose their thermodynamic stability. The tertiary structure characterization using fluorescence spectroscopy suggests that BARD1 BRCT wild-type and mutant proteins unfold *via* a three-state pathway. Moreover, the results from chemical denaturation are consistent with the results obtained from thermal denaturation by fluorescence spectroscopy.

According to molecular dynamics simulation, all the mutants show higher RMSD value as compared with the wild-type structure. The RMSF profile for wild-type and mutant structure indicates that the mutation has brought only local changes in the RMSF values. Comparative hydrogen bonding profile indicates that the BARD1 Val695Leu structure significantly loses hydrogen bonds as compared with the wild-type structure. Val695 residue is buried in the hydrophobic core of the BARD1 BRCT structure. Substitution of valine into leucine can lead to the addition of extra $-\text{CH}_2-$ groups, which can bring entropic changes in the hydrophobic core of the protein, and in turn result in an increase in the volume and the solvent accessible area. Intra-molecular hydrogen bonding is one of the major forces holding the secondary structure of the protein, and



the reduced number of H-bonds in the BARD1 BRCT Val695Leu and Ser761Asn mutant structures could be the possible reason behind the loss of compactness and thermodynamic stability.

Comparative circular dichroism spectroscopy profile and DSSP analysis suggest no significant change in the secondary structure of the mutant proteins. No drastic changes were observed in the structure of BARD1 Cys645Arg mutant due to its location in the BARD1 BRCT structure. Its presence in the random coil region may lead to high local conformational entropy and different rotameric states of arginine⁴³ without affecting the secondary structure of BARD1 BRCT. Moreover, no large changes are observed in the secondary structure in valine and leucine mutants because of the frequent existence of leucine in α -helix and it is a strong helix former than valine.⁴⁴ It has been reported that due to conformational entropy, leucine is thermodynamically more favored in α -helix than valine.^{42,45,46} However, in BARD1 Ser761Asn substitution, a small polar amino acid has been converted into asparagine, which has a comparatively large side chain. Serine has a slightly higher α -helix propensity compared with asparagine. Therefore, a small local change can be expected, but capturing this minute change is beyond the sensitivity of CD spectroscopy.⁴² The possibility of losing the secondary structure in BARD1 Ser761Asn mutant is also insignificant because of the presence of strong α -helix forming neighboring amino acids.⁴²

Loss of hydrogen bond can lead to a loss in structural compactness, which has been revealed in the PCA. BARD1

Val695Leu mutant shows high atomic motion and periodic jump in the tertiary structure when projected on the first three eigenvectors. Cross-correlation of PCA for wild-type and mutant proteins indicates a significant change in the correlated atomic motion of the mutant as compared with the wild-type protein structure. Intra-molecular interaction analysis performed over the minimum energy structures signifies that BARD1 Val695Leu has lost hydrogen bonding interaction. BARD1 Ser761Asn shows the loss in hydrogen bonding with Ile764, which is a critical residue in the hydrophobic peptide-binding pocket, and may lead to the structural alterations in phosphopeptide-binding pocket. The studies were performed on BRCA1 BRCT and other BRCT repeat domains containing proteins^{35,47–51} and the mutations in BARD1 BRCT lead to distinct changes in the hydrogen bonding and non-covalent interactions, which could be responsible for the loss of BARD1 BRCT-phosphopeptide and ADP-ribose binding and early recruitment of BRCA1 at the DNA damage site.²⁶ Studies of changes in the hydrogen bonding and non-covalent interactions, which drives the BARD1 BRCT–ligand interaction, will thus be helpful in the development of high-affinity small-molecule BARD1 BRCT inhibitors. In conclusion, an alteration in structure and dynamics of BARD1 BRCT native structure by different mutations impairs the functionality of BARD1 BRCT as an important DNA damage repair protein (Fig. 8).

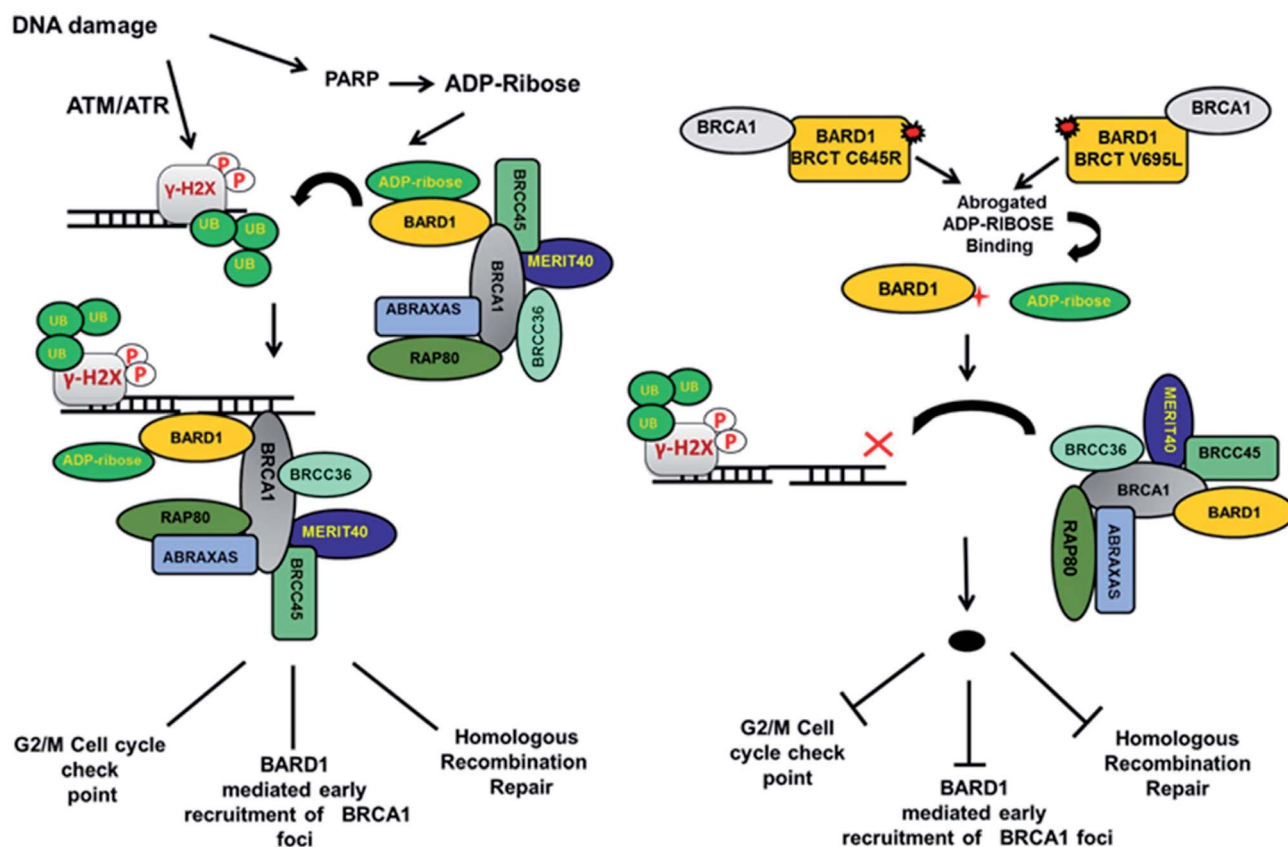


Fig. 8 Model representing the consequence of Cys645Arg and Val695Leu mutations on BARD1 BRCT function and DNA damage repair.



4 Experimental

4.1 Protein expression and purification of BARD1 BRCT wild-type and mutant proteins

The BARD1 BRCT (568–777) was expressed and purified as reported earlier.^{25,30,34} BARD1 BRCT mutants Cys645Arg, Val695Leu and Ser761Asn were generated by site-directed mutagenesis using mismatch forward and reverse primers. All the mutants were confirmed by DNA sequencing. The mutant proteins were expressed in *Escherichia coli* BL21 (DE3) cells grown at 37 °C until absorbance reached between 0.6 and 0.8 at 600 nm, followed by induction with 0.1 mM IPTG at 18 °C for 20 h. The His-tagged BARD1 BRCT proteins were affinity-purified in 50 mM Tris, 500 mM NaCl, 5% glycerol and 2 mM EDTA buffer (pH 7.5) supplemented with the protease inhibitors (Roche Applied Science). The protein of interest bound to Ni-NTA column was eluted by using a gradient concentration of imidazole and treated with TEV protease to remove the His tag. The mutant proteins were further purified using FPLC buffer (10 mM HEPES, 150 mM NaCl, pH 7.5) for biophysical characterization.

4.2 Chemical denaturation of BARD1 BRCT wild-type and mutant proteins

Chemical denaturation of BARD1 BRCT wild-type and mutant proteins Cys645Arg, Val695Leu and Ser761Asn was performed at 10 °C. By using GuHCl as a denaturant, tryptophan micro-environment was monitored using a fluorescence spectrophotometer (Horiba, USA) at excitation wavelength $\lambda = 295$ nm. The 10 M GuHCl (Sigma-Aldrich) stock solutions were diluted in the range of 0–6 M with 2 μ M of BARD1 BRCT wild-type and mutant proteins in buffer (10 mM HEPES, 150 mM NaCl, pH 7.5) with increasing GuHCl concentration. Fluorescence emission spectra were recorded with a concentration change from $\lambda = 310$ –400 nm and blank subtractions were performed. In chemical denaturation, a three-state model (folded (N), intermediate (I) and unfolded (U)) was used for curve fitting.^{34,38,39} Furthermore, circular dichroism spectra of BARD1 BRCT wild-type and mutant proteins were recorded using sealed quartz cuvettes of 1 mm path length using a JASCO J-815 spectropolarimeter (Jasco, Easton, MD). BARD1 BRCT wild-type and mutant proteins at concentrations of 10 μ M and 40 μ M were scanned in the far-UV ($\lambda = 200$ –240 nm) and near-UV range ($\lambda = 350$ –250 nm), respectively. An average of seven spectra at a scan speed of 20 nm min⁻¹ with a resolution of 1 nm and response time of 1 s were used for representation. A standard blank correction was applied to the raw data. The results are presented as mean residual ellipticity [θ] (deg cm² dmol⁻¹).

4.3 Thermal denaturation

Thermal denaturation of BARD1 BRCT wild-type and mutant proteins at a concentration of 10 μ M and 40 μ M in buffer A causes unfolding over a temperature range of 5–75 °C. Spectra were recorded both in far-UV range (200–240 nm) and near-UV range (350–250 nm) at a scan speed of 20 nm s⁻¹. The sample was heated at a rate of 1 °C min⁻¹, and change in the millidegree ellipticity was monitored at 222 nm. After blank correction, an averaged

normalized data of three independent experiments in the far UV range were considered for T_m calculation and fitting in a two-state transition model to obtain thermodynamic parameters.^{34,38,39}

4.4 Limited proteolysis and glutaraldehyde crosslinking of BARD1 BRCT wild-type and mutant proteins

Initially, 2 mg mL⁻¹ BARD1 BRCT wild-type and mutant proteins were incubated with 10 pg μ L⁻¹ trypsin and chymotrypsin (final concentration) each at 37 °C and 25 °C, respectively, for different time intervals of 0, 5, 10, 30, 60, 120 and 180 min. After incubation, the reaction was terminated individually by adding 1 mM PMSF (Sigma Aldrich), and the samples were heated after adding an equal volume of Laemmli buffer, and then analyzed over SDS-PAGE. Untreated BARD1 BRCT wild-type and mutant proteins with trypsin and chymotrypsin were considered as control, and experiments were performed in triplicates. Furthermore, wild-type and mutant proteins Cys645Arg, Val695Leu and Ser761Asn in buffer (10 mM HEPES, 150 mM NaCl, pH 7.5) were incubated with freshly prepared solution of glutaraldehyde (final concentration, 0.1%) for 0, 1, 2.5, 5, 10, 15, 30 and 60 minutes at 37 °C. The cross-linking reaction was terminated by adding 5 μ L of 1 M Tris-HCl of pH 8.0. The cross-linked samples were heated after adding an equal volume of Laemmli sample buffer before loading and analysis on 12% SDS-PAGE.

4.5 Dynamic light scattering experiment (DLS) of BARD1 BRCT wild-type and mutant proteins

BARD1 BRCT wild-type and Cys645Arg, Val695Leu and Ser761Asn mutant proteins at a concentration of 1 mg mL⁻¹ in buffer (10 mM HEPES, 150 mM NaCl, pH 7.5) were degassed and filtered (0.22 μ M filter, Millipore) before subjecting to hydrodynamic analysis using a Malvern particle size analyzer (Zetasizer μ V). All the proteins were scanned at 5 min intervals for 15 min. The experiment was repeated thrice and an average was considered to be the true value for calculating the effective diameter.

4.6 *In silico* prediction of BARD1 BRCT mutants

The possible pathogenic consequences of substitution mutations of BARD1 BRCT Cys645Arg, Val695Leu and Ser761Asn were studied using I-mutant 3.0,³² AGVGD^{52,53} and SNP&GO³³ software tools.

4.7 Molecular dynamics simulation

Molecular dynamics simulation was performed for BARD1 BRCT and mutant structure with GROMACS 4.5.5 (ref. 54 and 55) using OPLS force field.⁵⁶ Coordinates of BARD1 BRCT for MD simulation were taken from PDB ID: 2NTE.³⁰ The systems were embedded in the simple point charge (SPC) water model, with a margin of 1.5 nm in between the protein and the boundaries of the periodic cubic box. The ionization state of the residue was set to be consistent with neutral pH and counter ions were added to neutralize the system. All the bond lengths were constrained using the SHAKE algorithm⁵⁷ and electrostatic interactions were calculated by the particle mesh Ewald (PME)



method.^{58,59} The solvated system was then subjected to energy minimization, as reported earlier,^{34,60} using the steepest-descent algorithm to eliminate any bad contacts with added water until a tolerance of 1000 kJ mol⁻¹ was reached. The energy-minimized system was treated for equilibration by 100 ps NVT simulation at 300 K, followed by 100 ps of NPT simulation to achieve proper equilibration of the simulated system. Final production simulations were performed in the isothermal isobaric (NPT) ensemble at 300 K using an external bath with a coupling constant of 0.1 ps. The pressure was kept constant (1 bar) using pressure coupled with the time constant set to 1 ps. The trajectories were stored at every 2 ps during the simulation. Analysis of RMSD, RMSF, radius of gyration, number of hydrogen bonds, molecular volume and SASA were performed using the GROMACS 4.5.5 inbuilt commands.

4.8 Principal component analysis (PCA)

PCA was performed over the wild-type and mutant protein trajectories. Eigenvalues and eigenvectors were calculated by diagonalizing the covariance matrix. Furthermore, covariance analysis was performed for the BARD1 BRCT wild-type and Cys645Arg, Val695Leu and Ser761Asn mutant proteins with Gromacs 4.5.5. Cross-correlation for PCA was performed on the C α -atoms of the respective trajectories using the Prody software.⁶¹

Conflicts of interest

All the authors declare that is no conflict of interest.

Acknowledgements

Funding for this study was supported by DST (grant DST/INT/RUS/RSF/P-11) and the Russian Scientific Foundation (Grant 16-45-02011). We thank DBT-BTIS facility at ACTREC for providing software support.

References

- R. Baer and T. Ludwig, *Curr. Opin. Genet. Dev.*, 2002, **12**, 86–91.
- P. S. Brzovic, J. E. Meza, M. C. King and R. E. Klevit, *J. Biol. Chem.*, 2001, **276**, 41399–41406.
- P. S. Brzovic, P. Rajagopal, D. W. Hoyt, M. C. King and R. E. Klevit, *Nat. Struct. Biol.*, 2001, **8**, 833–837.
- R. Baer and T. Ludwig, *Curr. Opin. Genet. Dev.*, 2002, **12**, 86–91.
- R. Hashizume, M. Fukuda, I. Maeda, H. Nishikawa, D. Oyake, Y. Yabuki, H. Ogata and T. Ohta, *J. Biol. Chem.*, 2001, **276**, 14537–14540.
- R. A. Greenberg, B. Sobhian, S. Pathania, S. B. Cantor, Y. Nakatani and D. M. Livingston, *Genes Dev.*, 2006, **20**, 34–46.
- P. S. Brzovic, J. R. Keeffe, H. Nishikawa, K. Miyamoto, D. Fox, M. Fukuda, T. Ohta and R. Klevit, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 5646–5651.
- Y. Xia, G. M. Pao, H.-W. Chen, I. M. Verma and T. Hunter, *J. Biol. Chem.*, 2003, **278**, 5255–5263.
- C. F. Chen, L. W. Chen, C. T. Chien, M. S. Wu and T. J. Tsai, *Clin. Exp. Pharmacol. Physiol.*, 1996, **23**, 819–824.
- A. Feki, C. E. Jefford, P. Durand, J. Harb, H. Lucas, K. H. Krause and I. Irminger-Finger, *Biol. Reprod.*, 2004, **71**, 1614–1624.
- J. A. Rodriguez, S. Schuchner, W. W. Au, M. Fabbro and B. R. Henderson, *Oncogene*, 2004, **23**, 1809–1820.
- J. A. Rodriguez, W. W. Au and B. R. Henderson, *Exp. Cell Res.*, 2004, **293**, 14–21.
- M. Cohen, A. Meisser, L. Haenggeli, I. Irminger-Finger and P. Bischof, *Mol. Hum. Reprod.*, 2007, **13**, 111–116.
- Q. Wu, A. Paul, D. Su, S. Mehmood, T. K. Foo, T. Ochi, E. L. Bunting, B. Xia, C. V. Robinson and B. Wang, *Mol. Cell*, 2016, **61**, 434–448.
- T. H. Thai, F. Du, J. T. Tsan, Y. Jin, A. Phung, M. A. Spillman, H. F. Massa, C. Y. Muller, R. Ashfaq, J. M. Mathis, D. S. Miller, B. J. Trask, R. Baer and A. M. Bowcock, *Hum. Mol. Genet.*, 1998, **7**, 195–202.
- D. Fox, 3rd, I. Le Trong, P. Rajagopal, P. S. Brzovic, R. E. Stenkamp and R. E. Klevit, *J. Biol. Chem.*, 2008, **283**, 21179–21186.
- J. M. Glover, R. S. Williams and M. S. Lee, *Trends Biochem. Sci.*, 2004, **29**, 579–585.
- M. Stucki, J. A. Clapperton, D. Mohammad, M. B. Yaffe, S. J. Smerdon and S. P. Jackson, *Cell*, 2005, **123**, 1213–1226.
- A. K. Varma, R. S. Brown, G. Birrane and J. A. Ladias, *Biochemistry*, 2005, **44**, 10941–10946.
- C. C. Y. Leung and J. M. Glover, *Cell Cycle*, 2011, **10**, 2461–2470.
- I. Rappold, K. Iwabuchi, T. Date and J. Chen, *J. Cell Biol.*, 2001, **153**, 613–620.
- I. A. Manke, D. M. Lowery, A. Nguyen and M. B. Yaffe, *Science*, 2003, **302**, 636–639.
- M. Masson, C. Niedergang, V. Schreiber, S. Muller, J. Menissier-de Murcia and G. de Murcia, *Mol. Cell. Biol.*, 1998, **18**, 3563–3571.
- R. M. Taylor, A. Thistlethwaite and K. W. Caldecott, *Mol. Cell. Biol.*, 2002, **22**, 2556–2563.
- A. Thanassoulas, M. Nomikos, M. Theodoridou, D. Yannoukakos, D. Mastellos and G. Nounesis, *Biochim. Biophys. Acta*, 2010, **1804**, 1908–1916.
- M. Li and X. Yu, *Cancer Cell*, 2013, **23**, 693–704.
- M. Ishitobi, Y. Miyoshi, S. Hasegawa, C. Egawa, Y. Tamaki, M. Monden and S. Noguchi, *Cancer Lett.*, 2003, **200**, 1–7.
- C. Ghimenti, E. Sensi, S. Presciuttini, I. M. Brunetti, P. Conte, G. Bevilacqua and M. A. Caligo, *Genes, Chromosomes Cancer*, 2002, **33**, 235–242.
- M. Laufer, S. V. Nandula, A. P. Modi, S. Wang, M. Jasin, V. V. Murty, T. Ludwig and R. Baer, *J. Biol. Chem.*, 2007, **282**, 34325–34333.
- G. Birrane, A. K. Varma, A. Soni and J. A. Ladias, *Biochemistry*, 2007, **46**, 7706–7712.
- S. V. Tavtigian, M. S. Greenblatt, F. Lesueur and G. B. Byrnes, *Hum. Mutat.*, 2008, **29**, 1327–1336.



- 32 E. Capriotti, P. Fariselli, I. Rossi and R. Casadio, *BMC Bioinf.*, 2008, **9**(Suppl 2), S6.
- 33 R. Calabrese, E. Capriotti, P. Fariselli, P. L. Martelli and R. Casadio, *Hum. Mutat.*, 2009, **30**, 1237–1244.
- 34 R. K. Choudhary, Vikrant, Q. M. Siddiqui, P. S. Thapa, S. Raikundalia, N. Gadewal, N. S. Kumar, M. Hosur and A. K. Varma, *J. Biomol. Struct. Dyn.*, 2015, 1–29.
- 35 C. M. Ekblad, H. R. Wilkinson, J. W. Schymkowitz, F. Rousseau, S. M. Freund and L. S. Itzhaki, *J. Mol. Biol.*, 2002, **320**, 431–442.
- 36 C. M. Ekblad, A. Friedler, D. Veprintsev, R. L. Weinberg and L. S. Itzhaki, *Protein Sci.*, 2004, **13**, 617–625.
- 37 A. Thanassoulas, M. Nomikos, M. Theodoridou, P. Stavros, D. Mastellos and G. Nounesis, *Int. J. Biol. Macromol.*, 2011, **49**, 297–304.
- 38 C. N. Pace and K. L. Shaw, *Proteins*, 2000, (Suppl 4), 1–7.
- 39 Vikrant, P. Nakhwa, D. C. Badgular, R. Kumar, K. K. Rathore and A. K. Varma, *J. Biomol. Struct. Dyn.*, 2014, **32**, 2017–2032.
- 40 J. R. Lakowicz, *Principles of fluorescence spectroscopy*, Springer Science & Business Media, 2013.
- 41 N. Guex and M. C. Peitsch, *Electrophoresis*, 1997, **18**, 2714–2723.
- 42 C. N. Pace and J. M. Scholtz, *Biophys. J.*, 1998, **75**, 422–427.
- 43 M. J. Sternberg and J. S. Chickos, *Protein Eng.*, 1994, **7**, 149–155.
- 44 P. Y. Chou and G. D. Fasman, *Annu. Rev. Biochem.*, 1978, **47**, 251–276.
- 45 J. Hermans, *Curr. Opin. Struct. Biol.*, 1993, **3**, 270–276.
- 46 T. P. Creamer and G. D. Rose, *Proteins: Struct., Funct., Bioinf.*, 1994, **19**, 85–97.
- 47 N. Coquelle, R. Green and J. M. Glover, *Biochemistry*, 2011, **50**, 4579–4589.
- 48 J. Liu, Y. Pan, B. Ma and R. Nussinov, *Structure*, 2006, **14**, 1811–1821.
- 49 M. Li, L.-Y. Lu, C.-Y. Yang, S. Wang and X. Yu, *Genes Dev.*, 2013, **27**, 1752–1768.
- 50 V. M. Anisimov, A. Ziemys, S. Kizhake, Z. Yuan, A. Natarajan and C. N. Cavasotto, *J. Comput.-Aided Mol. Des.*, 2011, **25**, 1071–1084.
- 51 W. You, M. H. Yu-ming, S. Kizhake, A. Natarajan and A. C. Chia-en, *PLoS Comput. Biol.*, 2016, **12**, e1005057.
- 52 S. V. Tavtigian, A. M. Deffenbaugh, L. Yin, T. Judkins, T. Scholl, P. B. Samollow, D. de Silva, A. Zharkikh and A. Thomas, *J. Med. Genet.*, 2006, **43**, 295–305.
- 53 E. Mathe, M. Olivier, S. Kato, C. Ishioka, P. Hainaut and S. V. Tavtigian, *Nucleic Acids Res.*, 2006, **34**, 1317–1325.
- 54 H. Berendsen, J. Grigera and T. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269–6271.
- 55 H. J. Berendsen, D. van der Spoel and R. van Drunen, *Comput. Phys. Commun.*, 1995, **91**, 43–56.
- 56 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179–5197.
- 57 B. Hess, C. Kutzner, D. Van Der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
- 58 K. Kholmurodov, W. Smith, K. Yasuoka, T. Darden and T. Ebisuzaki, *J. Comput. Chem.*, 2000, **21**, 1187–1191.
- 59 T. I. Cheatham, J. Miller, T. Fox, T. Darden and P. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 4193–4194.
- 60 R. K. Choudhary, M. Q. Siddiqui, P. S. Thapa, N. Gadewal, S. K. Nachimuthu and A. K. Varma, *Sci. Rep.*, 2017, **7**, 3849.
- 61 A. Bakan, L. M. Meireles and I. Bahar, *Bioinformatics*, 2011, **27**, 1575–1577.

