nature
structural &
molecular biology



# Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1

R Scott Williams, Megan S Lee, D Duong Hau & J N Mark Glover

**The BRCT repeats in BRCA1 are essential for its tumor suppressor activity and interact with phosphorylated protein targets containing the sequence pSer-X-X-Phe, where X indicates any residue. The structure of the tandem BRCA1 BRCT repeats bound to an optimized phosphopeptide reveals that the N-terminal repeat harbors a conserved BRCT phosphoserine-binding pocket, while the interface between the repeats forms a hydrophobic groove that recognizes the phenylalanine. Crystallographic and biochemical data suggest that the structural integrity of both binding sites is essential for peptide recognition. The diminished peptide-binding capacity observed for cancer-associated BRCA1-BRCT variants may explain the enhanced cancer risks associated with these mutations.**

A large proportion of BRCA1-linked breast and ovarian cancers can be traced to truncation or missense mutations in the BRCA1 C-terminal (BRCT) domain (reviewed in ref. 1). The integrity of this region of the protein is essential for the normal functioning of BRCA1 in the repair of DNA double-strand breaks and homologous recombination[2].

The BRCT domain interacts with a diverse complement of proteins, suggesting a multifaceted role for this protein in the DNA damage response (reviewed in ref. 3). Several of its binding partners have been implicated in transcriptional regulation, such as the transcriptional co-repressor CtIP, histone deacetylases and the RNA polymerase holoenzyme. This is consistent with the idea that BRCA1 may affect the DNA damage response by regulating genes responsive to DNA damage, such as p21 and GADD45. A more direct role in DNA repair is indicated by the interaction of the BRCT domain with the DNA helicase BACH1; this interaction is required for efficient repair of double-strand breaks[4].

The BRCT domain consists of a pair of conserved BRCT repeats ~90–100 amino acids in length. BRCT repeats are found not only in BRCA1, but also in many other proteins that respond to DNA damage[5–7]. Several structural studies have revealed a conserved structure for the repeat, composed of a four-stranded parallel β-sheet flanked by a pair of α-helices on one face (α1 and α3), and a single α-helix (α2) on the opposite face[8–13]. The two BRCT repeats of BRCA1 pack in a head-to-tail manner involving α2 of the N-terminal repeat and α1 and α3 of the C-terminal repeat, as well as an inter-repeat linker[10,11]. The interface between the two repeats is highly conserved in several other proteins containing tandem BRCT repeats, and an essentially identical packing arrangement has been observed in the dual BRCT repeats of the DNA damage–responsive p53-binding protein, 53BP1 (refs. 11,13). Several BRCA1 missense mutations that have been linked to an enhanced risk for breast and ovarian cancer are localized at the interface between the two BRCT repeats and indicate that the correct packing of the two repeats is essential to BRCA1 function and tumor suppression[10,14,15].

It has recently been shown that the tandem BRCT repeats of BRCA1, but not the isolated, individual repeats, function as phosphopeptide-binding modules[16–18]. The BRCT repeats of several other proteins involved in the cellular response to DNA damage have also been suggested to bind phosphorylated peptide targets, indicating that this may be a common property of the BRCT protein family. The BRCA1 BRCT binds with high affinity to serine-phosphorylated peptides that contain phenylalanine at a position three residues C-terminal to the phosphoserine (+3 position)[17,18]. A pSer-X-X-Phe motif in BACH1 is essential for the interaction of BRCA1 and BACH1 in human cells and is required for the activation of the G2-M cell cycle checkpoint in response to DNA damage[17]. This may represent a primary means by which the BRCA1 BRCT domains, as well as many other BRCT proteins, interact with binding partners in response to DNA damage–induced cellular signaling.

Here we report the structure of a complex of the human BRCA1 BRCT domain bound to a high-affinity peptide target containing the recognition motif, pSer-X-X-Phe. The structure reveals the nature of binding pockets for both the target phosphoserine and phenylalanine residues. The conservation of residues that line these pockets indicates that this function is conserved not only in BRCA1 homologs, but also in several other BRCT proteins involved in DNA repair. The phosphopeptide-binding properties of a large set of clinically derived BRCA1 BRCT variants, and the structures of missense variants that are defective in peptide binding, highlight the importance of the integrity of the peptide-binding surface to BRCA1 function and tumor suppression.

Department of Biochemistry, University of Alberta, Edmonton, Alberta, T6G 2H7, Canada. Correspondence should be addressed to J.N.M.G. (mark.glover@ualberta.ca).

Published online 9 May 2004; doi:10.1038/nsmb776

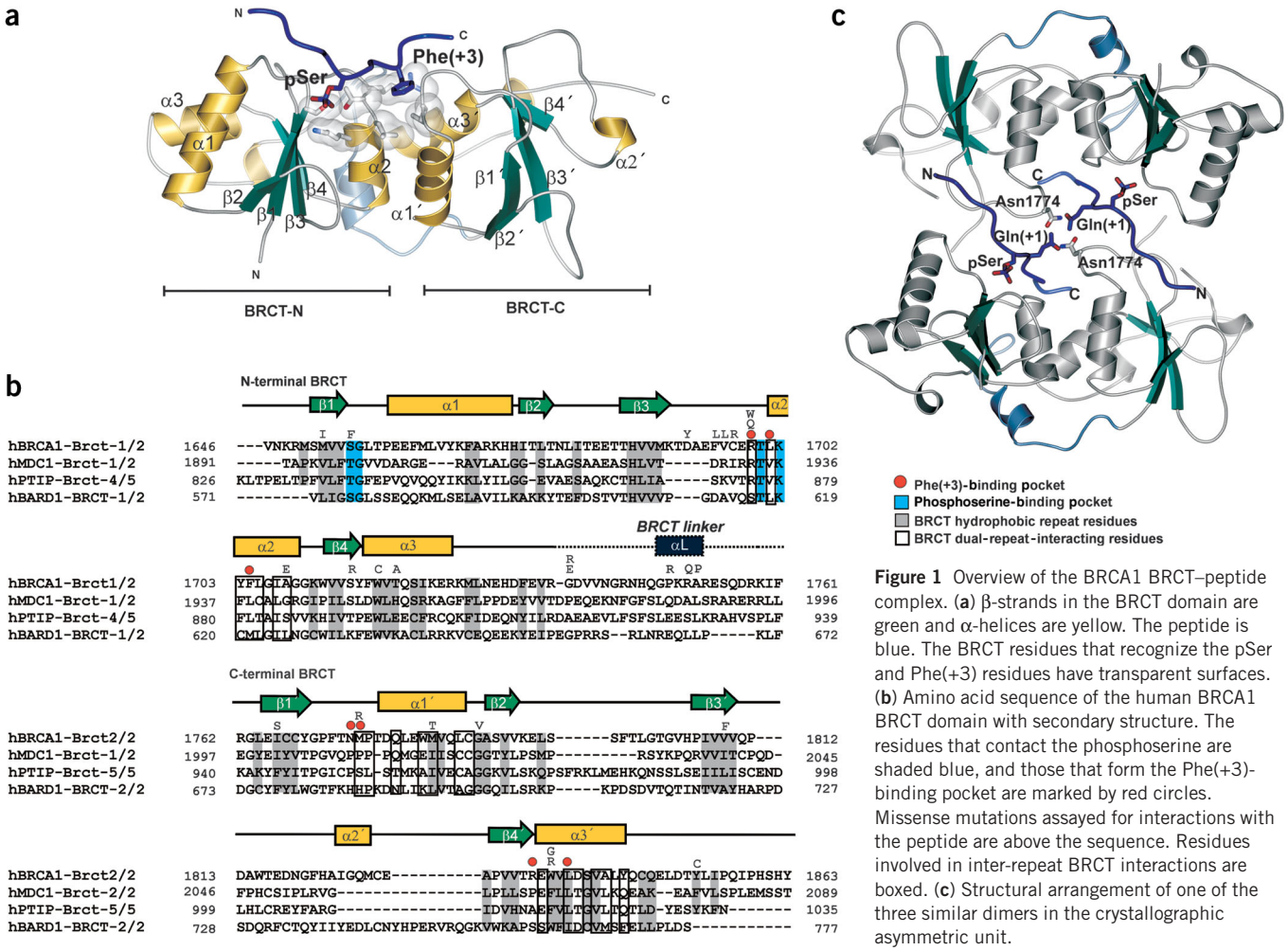





**Figure 1** Overview of the BRCA1 BRCT–peptide complex. (**a**) β-strands in the BRCT domain are green and α-helices are yellow. The peptide is blue. The BRCT residues that recognize the pSer and Phe(+3) residues have transparent surfaces. (**b**) Amino acid sequence of the human BRCA1 BRCT domain with secondary structure. The residues that contact the phosphoserine are shaded blue, and those that form the Phe(+3)-binding pocket are marked by red circles. Missense mutations assayed for interactions with the peptide are above the sequence. Residues involved in inter-repeat BRCT interactions are boxed. (**c**) Structural arrangement of one of the three similar dimers in the crystallographic asymmetric unit.

## RESULTS

### Structure determination

To define the structural basis for the recognition of serine-phosphorylated peptides by the BRCA1 BRCT, we crystallized and determined the structure of the human BRCA1 BRCT domain (encompassing residues 1649–1859) bound to an optimized pSer-X-X-Phe-containing peptide selected from an oriented peptide library[18] (**Fig. 1** and **Table 1**). The peptide, GAAYDIpSQVFPFAKKK (termed BRCTtide-7pS), binds the BRCT domain tightly with a $K_d$ of 400 nM (ref. 18). The structure was solved by molecular replacement with the structure of the unbound human BRCA1 BRCT domain. The crystallographic asymmetric unit contains five BRCT domain–peptide complexes. Non-crystallographic symmetry (NCS) enhanced the quality of the electron density maps and facilitated crystallographic refinement (see Methods).

### Overall structure

The five copies of the BRCT–peptide complex are arranged as two similar dimers, with a third, equivalent dimer generated by the rotation of the fifth BRCT–peptide complex about a crystallographic two-fold axis (**Fig. 1c**). Each dimer buries a large solvent-accessible surface (~2,000 Å$^2$ buried in total in the two BRCT–peptide complexes in each dimer). Molecular contacts between the BRCT repeats in the dimer are largely mediated through the bound peptides, which are sandwiched between the two domains. At the center of the dimer, a cage of hydrogen bonds is formed between the Gln(+l) residues from the peptides and Asn1774 of the BRCT domains. However, fixed-angle light scattering experiments and gel filtration chromatography indicate that the dimerization in solution is very weak, with a $K_d$ substantially >10 μM (data not shown).

The refined structures of the five peptide–BRCT complexes show that each peptide is positioned in a groove that extends across both BRCT repeats. The N-terminal portion of each peptide (residues Gly(–6) to pSer) lies over the N-terminal BRCT repeat, whereas the C-terminal portion (residues Asn(+1) to Lys(+7)) lies over the C-terminal BRCT repeat (**Figs. 1a** and **2a**). Tight interactions between the peptide and the BRCT domain involve residues from the pSer to the Phe(+3); residues further toward the N and C termini contact the BRCT domain less closely and show conformational variability in the five independent complexes (**Fig. 2b** and **Supplementary Fig. 1** online). These observations are consistent with the peptide selection studies, which have shown that positions other than +3 in this peptide scaffold are not strongly selected[18]. The phosphoserine is recognized by a shallow pocket composed of residues from the N-terminal BRCT repeat, whereas the Phe(+3) is recognized by a deeper and largely hydrophobic pocket at the interface between the N- and C-terminal repeats (**Fig. 2c**). This explains why both repeats are needed for efficient peptide binding[18]. In total, ~1,100 Å$^2$ is buried in the peptide-protein interface.

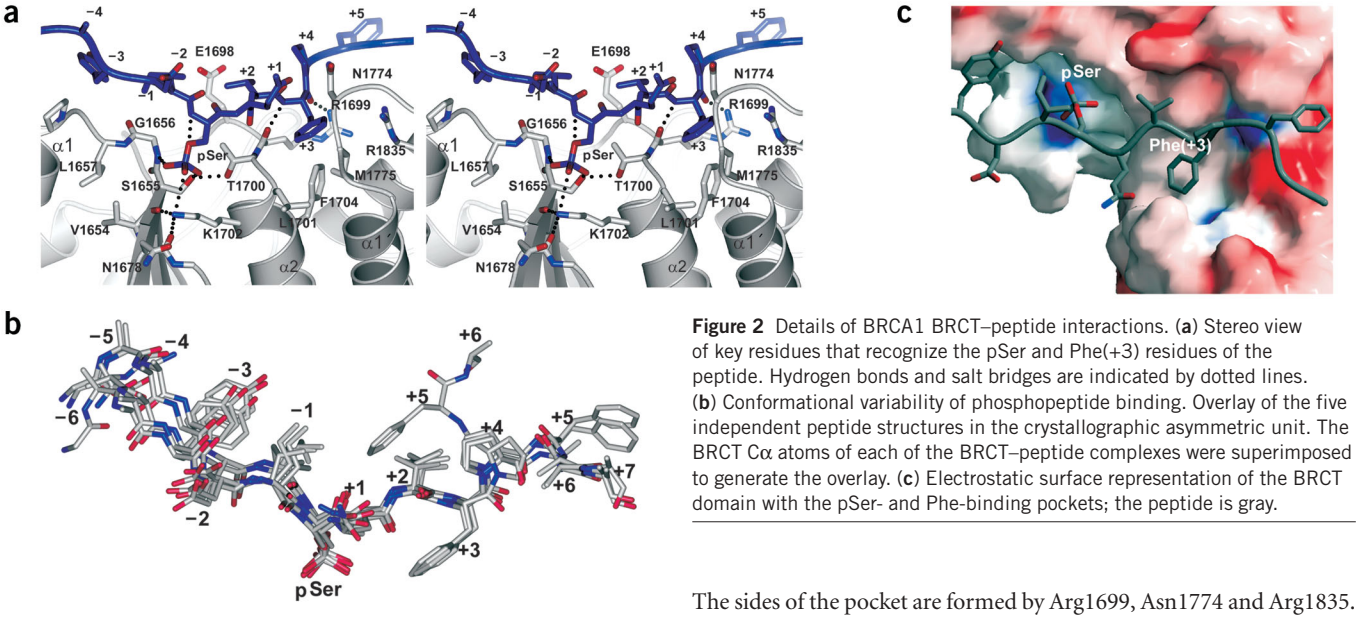


**Figure 2** Details of BRCA1 BRCT–peptide interactions. (**a**) Stereo view of key residues that recognize the pSer and Phe(+3) residues of the peptide. Hydrogen bonds and salt bridges are indicated by dotted lines. (**b**) Conformational variability of phosphopeptide binding. Overlay of the five independent peptide structures in the crystallographic asymmetric unit. The BRCT Cα atoms of each of the BRCT–peptide complexes were superimposed to generate the overlay. (**c**) Electrostatic surface representation of the BRCT domain with the pSer- and Phe-binding pockets; the peptide is gray.

## Phosphoserine recognition

A comparison of the structures of the BRCT domain both free and bound to peptide shows that the orientation of the hydrogen bond donors that make up the phosphoserine-binding pocket are prealigned for recognition of the phosphate (data not shown). The main chain NH of Gly1656 and the hydroxyl group of Ser1655 both donate hydrogen bonds to the phosphate (**Fig. 2a**). The serine hydroxyl is supported by further hydrogen-bonding interactions with Thr1700. The phosphate is also recognized by a salt-bridging interaction with Lys1702, which in turn is buttressed by hydrogen bonds to the main chain carbonyls of Val1654 and Asn1678. The orientation of the phosphate is further supported by a hydrogen bond between the phosphate Oγ and the main NH of the phosphoserine. The coordination of the phosphate oxygens by main chain NH, serine OH, and positively charged side chains is reminiscent of the way in which other protein modules, such as the FHA and 14-3-3 families, recognize phosphorylated residues[19].

We confirmed the importance of the phosphate-contacting residues to overall peptide binding by assessing the ability of BRCA1 BRCT variants containing mutations in the phosphate-binding pocket to bind a biotinylated peptide library containing the Ser-X-X-Phe motif, immobilized on streptavidin beads, in a phosphorylation-dependent manner (**Fig. 3a**). Wild-type BRCA1 BRCT bound tightly to the phosphopeptide library such that essentially all the *in vitro*–transcribed and translated protein was bound to the beads. This interaction was absolutely dependent on phosphorylation of the serine. In contrast, a S1655A mutation completely abolished the interaction, indicating that the hydroxyl group of Ser1655 is essential for peptide binding in solution. The patient-derived missense mutant, S1655F, also shows no detectable binding to the phosphopeptide (**Fig. 3**), indicating that this previously unclassified variant may be associated with an increased risk for cancer. Mutation of Lys1702, either to alanine or the isosteric methionine, also obliterated binding, verifying the importance of the side chain amino group for phosphopeptide recognition.

## Structure of the +3 binding pocket

The hydrophobic +3 binding pocket occupies the interface between the N- and C-terminal BRCT repeats (**Figs. 1a,b** and **2a**). The floor of the pocket is composed of Leu1701, Phe1704, Met1775 and Leu1839. The sides of the pocket are formed by Arg1699, Asn1774 and Arg1835. Arg1699 also has an important role in positioning the main chain of the Phe(+3). The main chain carbonyl of Arg1699 forms a hydrogen bond with the main chain NH of the residue at the +3 position, and the guanidinium group of the side chain makes contacts with the main chain carbonyl of the residue at the +3 position.

## BRCT mutations affect peptide binding

Truncation and missense mutations in the BRCA1 BRCT domain that have been isolated from patients in breast and ovarian cancer screening programs are recorded in a large database (available through the Breast Cancer Information Database http://research.nhgri.nih.gov/bic/, and reviewed in ref. 15). All of the truncation mutations and a large proportion of the missense mutations destabilize the protein fold[15,20]. We probed the ability of a large panel of these BRCA1 BRCT variants to bind biotinylated pSer-X-X-Phe peptides as described above (**Fig. 3b,c**). The missense mutations that cause little or no folding defects showed a range of phosphopeptide-binding activities. None of the highly destabilizing truncation or missense mutants specifically bound the phosphorylated peptides, demonstrating that correct folding of the BRCT domain is essential for specific recognition of the phosphopeptide target.

Several mutants (F1695L, T1720A, R1751Q, M1783T and M1652I) showed peptide-binding properties that were essentially indistinguishable from those of the wild type. M1652I has been shown to be a polymorphism not associated with an increased cancer risk[21] whereas there is insufficient pedigree data for the other mutations to directly assess their association with cancer risk. One of the moderately destabilizing mutants, D1692Y, showed a slight but reproducible defect in peptide binding, which may account for its association with disease (**Fig. 3c**). Most notably, several of these mutations (R1699Q, R1699W, V1696L, M1775R and V1809F) showed a complete loss of phosphopeptide-binding activity. R1699W would impair the ability of the BRCT domain to hydrogen-bond with the backbone of the peptide at the +3 position and this mutation has been suggested to be associated with breast and ovarian cancer[22]. Notably, R1699Q showed an equally marked peptide-binding defect, even though the glutamine side chain might be expected to hydrogen-bond in a similar manner to the wild-type arginine. M1775R, V1696L and V1809F also did not specifically bind the phosphopeptide.

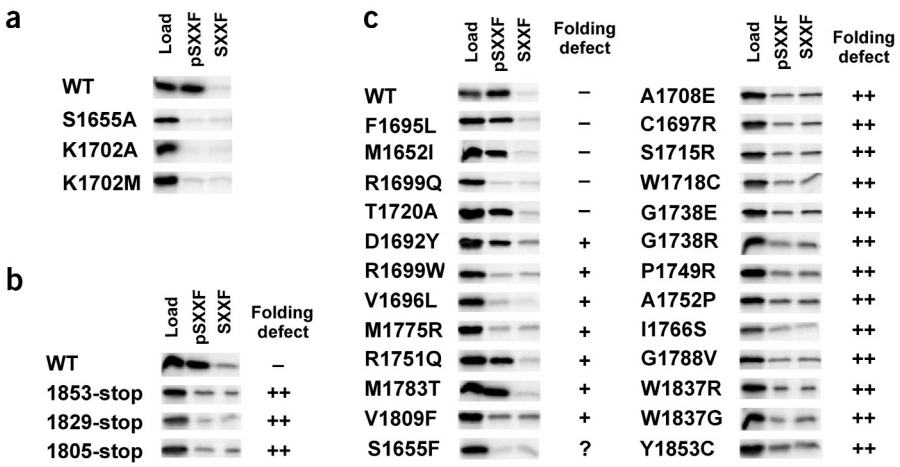


**Figure 3** Ability of mutant forms of the BRCA1 BRCT domain to bind pSer-containing peptides. (**a**) Effect of mutations in residues that directly contact pSer. Each of the mutants was produced by *in vitro* transcription and translation and was assayed for binding to biotinylated peptides containing Ser-X-X-Phe (SXXF) or pSer-X-X-Phe (pSXXF) motifs (see Methods). The far left lane of each set of three reactions shows 100% of the load material. The stability of the protein fold for most of the mutants was tested using a proteolytic assay[15] where ++ indicates a severe folding defect, + indicates a more modest defect, and – indicates that the stability of the mutant is indistinguishable from wild type. (**b**) Effect of truncation mutations associated with hereditary breast and ovarian cancer. (**c**) Effect of missense mutations derived from breast cancer screening programs.

### Structures of BRCT missense variants

We set out to crystallize BRCT missense variants to gain structural insights into the nature of the mutation-induced peptide-binding defects. Of the five intermediately destabilizing BRCT variants that were defective in peptide binding (D1692Y, R1699W, V1696L, M1775R and V1809F), four (R1699W, V1696L, M1775R and V1809F) could be produced in soluble form in *Escherichia coli*; three (R1699W, M1775R and V1809F) yielded crystals; and two (M1775R and V1809F) produced usable diffraction data beyond 3.0 Å with synchrotron radiation. The crystal structures of the M1775R and V1809F variants provide an explanation for their inability to bind pSer-X-X-Phe targets.

To understand how the V1809F mutation blocks peptide binding, we determined the crystal structure of this variant to a resolution of 2.8 Å (**Fig. 4** and **Table 1**). The structure of this mutant shows that a set of rearrangements in the protein hydrophobic core leads to the disruption of the +3 binding pocket (**Figs. 4a** and 5). In this mutant, the larger substituted phenylalanine side chain contacts Leu1780, causing an adjustment in this side chain that in turn brings it into contact with Met1775. This contact pushes the Met1775 side chain out into the +3 pocket, explaining the loss of peptide binding in this mutant. Notably, in the unbound structures of the V1809F variant, a strong electron density peak overlays with the position of the phosphate from the phosphoserine in the peptide-bound structure (**Fig. 4**). This feature is consistent with a bound sulfate anion, present in the crystallization solution, that may mimic phosphate binding, and suggests that the mutation does not disrupt phosphate anion coordination, but only the +3 binding pocket.

The M1775R mutation blocks the ability BRCA1 to interact with various protein partners including BACH1 (ref. 4), histone deacetylases[23] and CtIP[24]. The mutation is also associated with defects in DNA repair[2] and linked to an enhanced cancer risk[25]. It has been shown to cause substantial destabilization of the BRCA1 fold[14,20]. The structure of M1775R, described previously[14] (PDB entry 1N5O), reveals a subtle alteration in the protein structure at the site of the substitution. Superimposition of the structure of the BRCT–peptide complex on that of the M1775R variant determined in the absence of bound peptide shows that the guanidinium group of the substituted arginine occludes the +3 pocket, thereby blocking access by the Phe(+3) (**Fig. 5c**). Thus, the combined effects of protein destabilization and direct obstruction of the peptide interaction surface by the M1775R substitution probably contribute to BRCA1 tumor suppressor inactivation and ultimately to disease predisposition in individuals carrying this mutation.

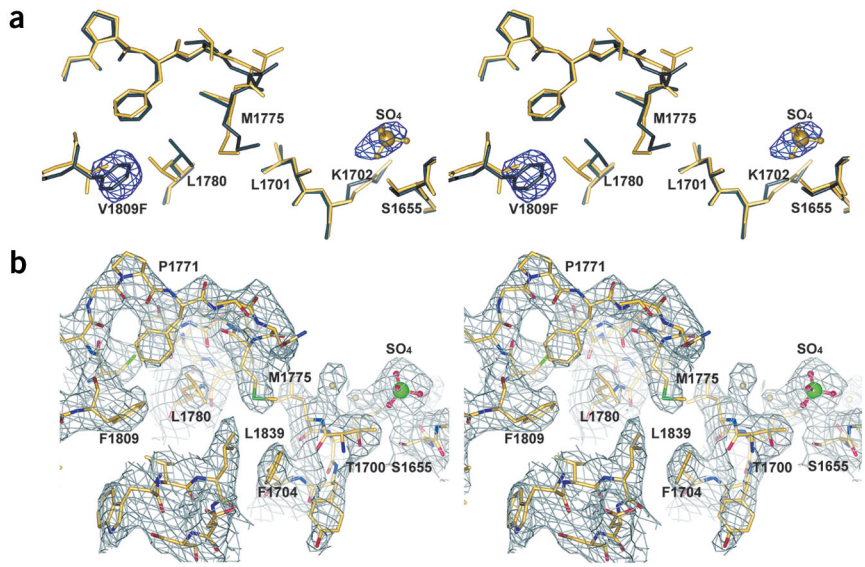


**Figure 4** Structure determination of variant V1809F. (**a**) $\sigma_A$-weighted $F_o - F_c$ difference electron density at 4.0 σ with a structural overlay of wild type (yellow) and V1809F (gray) structures. Electron density peaks for Phe1809 (gray) and a bound sulfate (gold ball and sticks) are observed. Sδ of Met1775 in the V1809F structure moves ~1.8 Å into the Phe(+3)-binding pocket. (**b**) $\sigma_A$-weighted $2F_o - F_c$ density for the final model at 2.8 Å is contoured at 1.0 σ.

## DISCUSSION

### A conserved phosphoserine-binding motif

The structures of proteins containing BRCT domains have revealed that the conserved set of BRCT residues, which contribute to the hydrophobic core and overall fold of these domains, are found in the central β-sheet, in helix α3 and at the C terminus of helix α1 (refs. 8,10–12). In general, there are few

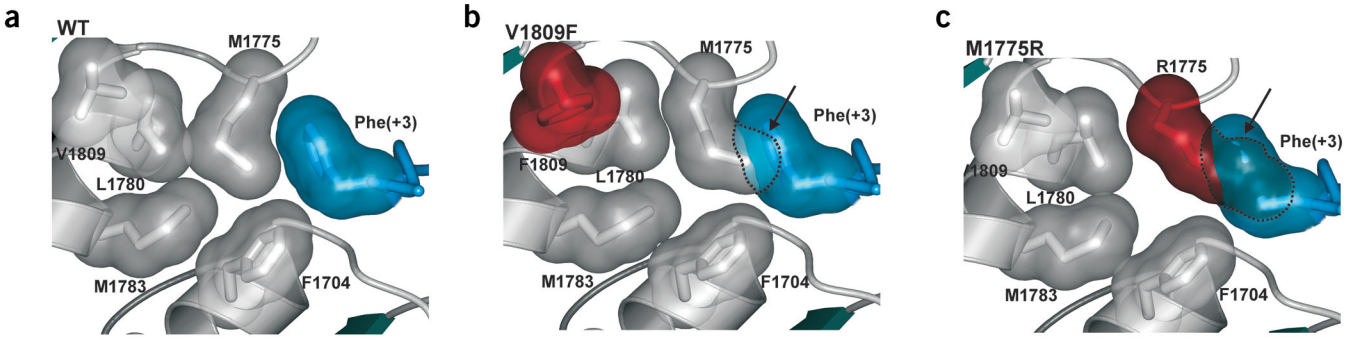


**Figure 5** Occlusion of the Phe(+3)-binding pocket by BRCA1 missense mutations V1809F and M1775R. (**a**) Wild-type amino acid packing environment near the Phe(+3)-binding pocket. Complementary van der Waals packing is observed between the protein (gray surfaces) and the Phe(+3) (blue surface). (**b**) The V1809F substitution (red surface) causes a relay of van der Waals packing rearrangements that permeate to the protein surface. The +3 binding pocket is blocked by Met1775 atoms Cγ, Sδ and Cε, which are found within 2–3 Å of the overlayed Phe(+3) aromatic ring. Black arrows and dotted lines indicate the BRCT surface incompatible with peptide binding. (**c**) Cancer-causing mutation M1775R (red surface) fills the +3 binding pocket and blocks peptide binding.



invariant residues among the BRCT proteins, and the sequences diverge greatly in regions that form the BRCT α2, and in two hypervariable insertion regions found at the β1-α1 and β2-β3 junctions[6,7,9]. These variable regions were originally hypothesized to form adaptable protein-protein interaction surfaces[8]. Indeed, both the β1-α1 loop and helix α2 participate in phosphoserine recognition in BRCA1.

Several BRCT proteins have now been tested for phosphopeptide binding, although the results found so far are controversial[16–18]. The structure of the BRCT domain–peptide complex reveals BRCT residues that have a key role in these interactions and provides a means of further identifying BRCT domains that may function as phosphopeptide-binding modules. Sequence analysis suggests that residues composing a conserved phosphoserine-binding pocket include a Ser/Thr-Gly motif in the variable β1-α1 connecting loop (Ser1655 and Gly1656 of hBRCA1) and a Thr/Ser-X-Lys motif at the N terminus of α2 of the same repeat (residues 1700–1702 of hBRCA1) (**Fig. 1b** and **Supplementary Fig. 2** online). Notably, the motif, which has not been included in previous descriptions of BRCT domain homology, is present in 57 of 518 unique BRCT sequences in the current Pfam database (http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00533), and in the N-terminal BRCT of dual-repeat-containing BRCT proteins reported to bind phosphopeptides, including all genuine BRCA1 homologs, hMDC1 (ref. 17), hPTIP[18] and hBARD1 (refs. 16,17; **Fig. 1b**). Moreover, single BRCT-containing proteins with apparent phosphoserine-binding activity, such as FCP1 (ref. 17) also contain this motif. Thus, we suggest that phosphoserine binding is a common function of BRCT domains, and has a critical role in regulating the assembly of protein complexes during the DNA damage response.

The dual-repeat BRCT domains of both BRCA1 and the transcriptional activator PTIP have been shown to bind not only peptides containing phosphoserine, but also those containing phosphothreonine, albeit with reduced affinity[18]. The γ-methyl group of a phosphothreonine residue would point away from the BRCA1 BRCT domain; however, it would sterically clash with a phosphate oxygen that is in a similar position in each of the five complexes in the asymmetric unit. Thus, replacing phosphoserine with phosphothreonine in the phosphopeptide would require a readjustment of the conformation of the peptide that might lead to a reduced binding affinity. Phosphotyrosine peptides do not seem to bind to these proteins[18], and indeed, it would be difficult to accommodate this large side chain in the phosphate-binding pocket without large changes in the conformation of either the phosphopeptide backbone or the phosphate-binding pocket.

### Specificity of binding at position +3

The crystal structures of BRCT protein complexes have revealed a role for dual-repeat domains as a scaffold for variable BRCT-BRCT linker structures that participate directly in ligand binding[11,13,26]. For example, the dual-repeat BRCT domain of 53BP1 anchors a β-hairpin BRCT-linker structure that contacts the DNA-binding domain of p53 (refs. 11,13). In an analogous manner, specific recognition of a bound phosphopeptide by the BRCA1 BRCT domain is dependent on both the dual-repeat scaffold structure and a BRCT variable loop (β1′-α1′), which together create a hydrophobic groove that binds the Phe(+3).

As expected, the interface between the two BRCT repeats and the +3 binding pocket is highly conserved in BRCA1 homologs[10,13]. Residues predicted to form the repeat interface are also conserved in other dual-repeat BRCT proteins known to bind phosphoserine-containing peptides (**Fig. 1b**). However, the amino acids lining the walls of this pocket, formed by the variable β1′-α1′ loop and the N terminus of α3′ from the C-terminal BRCT of a tandem repeat, seem less well conserved. Such sequence variability suggests that different BRCT domains may have diverse binding specificities.

In an independent study in this issue, the structure of the BRCA1 BRCT domain bound to the specific BACH1 phosphopeptide is reported[27]. Descriptions of the key determinants of phosphopeptide recognition, including the phosphoserine- and Phe(+3)-binding pockets, are in good agreement with the observations reported here. The mode of phosphoserine recognition observed for BRCA1 is probably conserved in other BRCT-containing proteins that participate in the coordination of the DNA damage response. Molecular snapshots of cancer-associated mutant BRCT proteins provide a structural explanation for the catastrophic inactivation of the BRCA1 tumor suppressor. The observation that the majority of mutations mapping to the BRCT of BRCA1 impair peptide binding helps to explain the elevated cancer risks associated with these variants.

## METHODS

**Cloning, protein expression and purification.** Mutations S1655A, S1655F, K1702A and K1702M were introduced into the human BRCA1 coding sequence for residues 1646–1859 using PCR methods and cloned into the *Eco*RI and *Bam*HI sites of pLM1. Vector construction of all other pLM1-based T7 expression vectors for BRCA1-BRCT variants was as described[15]. For structural studies, wild-type hBRCA1(1646–1859) or hBRCA1(1646-1859) variants were overexpressed in *E. coli* BL21(DE3) at 25 °C (wild type) or 20 °C (missense variants) for 12–15 h and purified as described[10].

**Table 1 Crystallographic data and refinement statistics**

| | BRCT–peptide complex | BRCT V1809F |
|---|---|---|
| Space group | $C222_1$ | $P6_122$ |
| Unit cell dimensions (Å) | | |
| *a* | 97.12 | 113.92 |
| *b* | 138.36 | 113.92 |
| *c* | 198.27 | 120.91 |
| Resolution range (Å) | 50–3.3 | 20–2.8 |
| No. of observed reflections | 81,261 | 116,289 |
| No. of unique reflections | 20,317 | 11,828 |
| Completeness (%)[a] | 99.4 (99.9) | 99.3 (96.3) |
| $R_{sym}$ (%)[a,b] | 12.5. (50.0) | 5.1 (43.4) |
| Overall $I / \sigma I$[a] | 11.1 (2.7) | 37.9 (3.0) |
| $R_{cryst}$ (%)[c] | 25.8 | 27.5 |
| $R_{free}$ (%)[c] | 30.2 | 29.4 |
| No. of atoms | 8,380 | 1,647 |
| $H_2O$ | 0 | 51 |
| $SO_4^{2-}$ | 0 | 10 |
| $Co^{2+}$ | 0 | 1 |
| Ramachandran plot (%) | | |
| Most favored | 90.0 | 85.7 |
| Allowed | 9.7 | 14.3 |
| Generous | 0.3 | 0 |
| R.m.s. deviation | | |
| Bonds (Å) | 0.008 | 0.019 |
| Angles (°) | 1.252 | 1.636 |

[a]Values in parentheses are for the highest-resolution shell, 3.42–3.30 Å for the BRCT–peptide complex, and 2.90–2.80 Å for BRCT V1809F. [b]$R_{sym} = 100 \times \Sigma_{hkl} |<I> - I| / \Sigma_n <I>$. [c]$R_{cryst} = \Sigma_n |F_o(h) - F_c(h)| / \Sigma_n |F_o(h)|$, where $F_o(h)$ and $F_c(h)$ are the observed and calculated structure factors, respectively, for the resolution range 50–3.3 Å (BRCT–peptide complex) or 20–2.8 Å (BRCT V1809F). $R_{free}$ was calculated with 5% (BRCT–peptide complex) or 7% (BRCT V1809F) of all reflections excluded from refinement.

**Crystallization and data collection.** For the BRCT–peptide complex, protein stocks were dialyzed into protein buffer (400 mM NaCl, 10mM Tris-HCl, pH 7.5, 1mM DTT) and concentrated to 20 mg $ml^{-1}$ for use in crystallization trials. For peptide complex formation, wild-type hBRCA1(1646–1859) at 20 mg $ml^{-1}$ was mixed with the optimized BRCT-binding phosphopeptide[18] (BRCTtide-7pS, Ac-GAAYDIpSQVFPFAKKK-$NH_2$) at a 1.5:1 phosphopeptide/protein molar ratio. Crystals were grown at 20–22 °C using the hanging-drop vapor diffusion by mixing 1 µl of complex in protein buffer with 1 µl of well solution 1 (0.2 M ammonium acetate, 0.1 M tri-sodium citrate dehydrate, pH 5.6, 30% (w/v) PEG 4000). Small crystals, which were observed after 3–4 weeks, grew to a maximal size of ~30 × 30 × 30 µm within 5–6 weeks.

Hexagonal crystals of BRCT missense variant V1809F, with unit cell dimensions similar to those of the wild-type protein crystals, crystallized at 4 °C within 2-4 d after the mixing of 2 µl of 20 mg $ml^{-1}$ hBRCA1(1646–1859) V1809F in protein buffer with well solution 2 (1.2–1.4 M ammonium sulfate, 50 mM MES, pH 6.7–6.8, 10 mM $CoCl_2$). For cryopreservation, single crystals were soaked in the appropriate well solution supplemented with 26% (v/v) glycerol and then flash-frozen in liquid nitrogen. All data were collected at beamline 8.3.1 of the Advanced Light Source, Lawerence Berkeley National Laboratory. Data reduction and scaling were done using the HKL package[28] (**Table 1**).

**Structure determination.** Molecular replacement trials using MOLREP[29] (version 7.3) were carried out with coordinates of the unliganded human BRCT dual-repeat structure[10] (PDB entry 1JNX). An initial four-molecule solution was obtained from a rotation and translation search using data in the range 10.0–4.0 Å. After rigid body refinement of this solution, inspection of $\sigma_A$-weighted $2F_o - F_c$ and $F_o - F_c$ electron density maps revealed clear density for a fifth BRCT domain that was fit manually into the electron density using O[30] (version 7.0). After a second round of rigid body fitting, electron density in $2F_o - F_c$ and $F_o - F_c$ maps for the corresponding five phosphopeptide chains was apparent, but discontinuous in some regions. Subsequent five-fold NCS averaging and density modification using RESOLVE[31] improved the overall quality of the $2F_o - F_c$ model-phased experimental electron density map and allowed for the unambiguous trace of the phosphopeptide chains. Further model building and refinement was carried out using O and REFMAC[32] (version 5.0). Five-fold NCS restraints applied to the BRCA1 protein chains allowed for positional refinement of coordinate parameters using data to a resolution of 3.3 Å. Differences between the BRCA1 protein chains were limited to regions that were ordered or disordered depending on crystallographic packing interactions. To allow for conformational variability of the peptides observed between the different complexes, NCS restraints were not used for the phosphopeptide chains. Refinement of thermal parameters was limited to overall *B*-factor refinement with a separate group TLS anisotropic thermal parameter model applied to each of the five protein–peptide complexes. The final model contains 1,101 amino acid residues, with good geometry. Refinement statistics are summarized in **Table 1**.

Structure solution and refinement of BRCA1 variant V1809F was similar to that described for variant M1775R[14]. Residue 1809 was mutated to alanine before rigid body fitting of the model to the V1809F hexagonal dataset. $\sigma_A$-weighted $F_o - F_c$ difference density contained two large 4-σ peaks corresponding to Phe1809 and a sulfate bound in the phosphoserine-binding pocket near Lys1702 and Ser1655. Further iterative model building and refinement was done using O[30] and REFMAC[32] and the model was assessed with PROCHECK[33]. Molecular surfaces were drawn with GRASP[34]. Structural figures were created using BobScript[35] and rendered with POV-Ray (http://www.povray.org).

**Peptide-binding assays.** pLM1 plasmid (0.1–0.4 µg) encoding mutant BRCT domains was used directly as template for coupled transcription and translation reactions in a reticulocyte lysate. For binding assays, the BRCT variants were translated and labeled with [$^{35}$S]methionine for 2 h at 30 °C using the TNT-Quick system (Promega). Bead-immobilized peptide affinity resin was prepared in binding buffer (150 mM NaCl, 50 mM Tris-HCl, pH 7.5, 0.1% (v/v) NP-40, 1 mM EDTA, 1 mM DTT) by incubating a ten-fold molar excess of a partially selected biotinylated phosphopeptide library (pSXXF, biotin-ZGZGGAXXXpSXXFXXAYKKK, where Z is aminohexanoic acid, X is any amino acid except cysteine, and pS is phosphoserine) and the corresponding dephosphorylated peptide (SXXF, biotin-ZGZGGAXXXSXXFXXAYKKK) with streptavidin agarose beads (25 pmol $µl^{-1}$ stock, Sigma-Aldrich) for 30 min at 4 °C. Excess peptide was removed by washing five times with ten bed volumes of binding buffer. [$^{35}$S]methionine-labeled BRCT-variant (1 µl) containing lysate was added to 20 µl of affinity resin in a total of 150 µl binding buffer supplemented with 0.3 mg $ml^{-1}$ BSA, which was used to reduce background resin binding observed for highly destabilizing BRCT variants. Following incubation for 2 h at 4 °C, the resin was washed three times with 200 µl binding buffer. Bound BRCT variants were eluted from the resin with the addition of 20 µl SDS-PAGE loading buffer, run on a 15% (w/v) SDS-PAGE gel, and visualized using a Molecular Dynamics phosphorimager.

**Coordinates.** Coordinates of the BRCT V1809F mutant and BRCT–peptide complex have been deposited in the Protein Data Bank (accession numbers 1T2U and 1T2V, respectively).

*Note: Supplementary information is available on the Nature Structural & Molecular Biology website.*

ACKNOWLEDGMENTS

We are grateful to J. Parrish, E. Bergmann, T. Moraes, J. Holton and the Berkeley Centre for Structural Biology staff for discussions and technical support during data collection. Data was collected on Advanced Light Source beamline 8.3.1 with funding from the Alberta Synchrotron Institute as part of a participating research team. We thank M. Yaffe and S. Smerdon for sharing unpublished data and the gift of biotinylated peptide libraries, and R. Boyko for help with BRCT sequence alignments. This work was supported by a grant from the Canadian Breast Cancer Research Alliance. J.N.M.G. is a Canada research chair in structural molecular biology.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 2 April; accepted 23 April 2004
Published online at http://www.nature.com/nsmb/

1. Nathanson, K.L., Wooster, R., Weber, B.L. & Nathanson, K.N. Breast cancer genetics: what we know and what we need. *Nat. Med.* **7**, 552–556 (2001).
2. Scully, R. *et al.* Genetic analysis of BRCA1 function in a defined tumor cell line. *Mol. Cell* **4**, 1093–1099 (1999).
3. Venkitaraman, A.R. Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* **108**, 171–182 (2002).
4. Cantor, S.B. *et al.* Bach1, a novel helicase-like protein, interacts directly with brca1 and contributes to its DNA repair function. *Cell* **105**, 149–160 (2001).
5. Bork, P. *et al.* A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins. *FASEB J.* **11**, 68–76 (1997).
6. Koonin, E.V., Altschul, S.F. & Bork, P. BRCA1 protein products ... Functional motifs... *Nat. Genet.* **13**, 266–268 (1996).
7. Callebaut, I. & Mornon, J.P. From BRCA1 to RAP1: a widespread BRCT module closely associated with DNA repair. *FEBS Lett.* **400**, 25–30 (1997).
8. Zhang, X. *et al.* Structure of an XRCC1 BRCT domain: a new protein-protein interaction module. *EMBO J.* **17**, 6404–6511 (1998).
9. Huyton, T., Bates, P.A., Zhang, X., Sternberg, M.J. & Freemont, P.S. The BRCA1 C-terminal domain: structure and function. *Mutat. Res.* **460**, 319–332 (2000).
10. Williams, R.S., Green, R. & Glover, J.N. Crystal structure of the BRCT repeat region from the breast cancer–associated protein BRCA1. *Nat. Struct. Biol.* **8**, 838-842 (2001).
11. Joo, W.S. *et al.* Structure of the 53BP1 BRCT region bound to p53 and its comparison to the Brca1 BRCT structure. *Genes Dev.* **16**, 583–593 (2002).
12. Krishnan, V.V., Thornton, K.H., Thelen, M.P. & Cosman, M. Solution structure and backbone dynamics of the human DNA ligase IIIα BRCT domain. *Biochemistry* **40**, 13158–13166 (2001).
13. Derbyshire, D.J. *et al.* Crystal structure of human 53BP1 BRCT domains bound to p53 tumour suppressor. *EMBO J.* **21**, 3863–3872 (2002).
14. Williams, R.S. & Glover, J.N. Structural consequences of a cancer-causing BRCA1-BRCT missense mutation. *J. Biol. Chem.* **278**, 2630–2635 (2003).
15. Williams, R.S. *et al.* Detection of protein folding defects caused by BRCA1-BRCT truncation and missense mutations. *J. Biol. Chem.* **278**, 53007–53016 (2003).
16. Rodriguez, M., Yu, X., Chen, J. & Songyang, Z. Phosphopeptide binding specificities of BRCA1 COOH-terminal (BRCT) domains. *J. Biol. Chem.* **278**, 52914–52918 (2003).
17. Yu, X., Chini, C.C., He, M., Mer, G. & Chen, J. The BRCT domain is a phospho-protein binding domain. *Science* **302**, 639–642 (2003).
18. Manke, I.A., Lowery, D.M., Nguyen, A. & Yaffe, M.B. BRCT repeats as phosphopeptide-binding modules involved in protein targeting. *Science* **302**, 636–639 (2003).
19. Yaffe, M.B. & Smerdon, S.J. Phosphoserine/threonine binding domains: you can't pSERious? *Structure* **9**, R33–R38 (2001).
20. Ekblad, C.M. *et al.* Characterisation of the BRCT domains of the breast cancer susceptibility gene product BRCA1. *J. Mol. Biol.* **320**, 431–442 (2002).
21. Deffenbaugh, A.M., Frank, T.S., Hoffman, M., Cannon-Albright, L. & Neuhausen, S.L. Characterization of common BRCA1 and BRCA2 variants. *Genet. Test.* **6**, 119–121 (2002).
22. Vallon-Christersson, J. *et al.* Functional analysis of BRCA1 C-terminal missense mutations identified in breast and ovarian cancer families. *Hum. Mol. Genet.* **10**, 353–360 (2001).
23. Yarden, R.I. & Brody, L.C. BRCA1 interacts with components of the histone deacetylase complex. *Proc. Natl. Acad. Sci. USA* **96**, 4983–4988 (1999).
24. Yu, X., Wu, L.C., Bowcock, A.M., Aronheim, A. & Baer, R. The C-terminal (BRCT) domains of BRCA1 interact *in vivo* with CtIP, a protein implicated in the CtBP pathway of transcriptional repression. *J. Biol. Chem.* **273**, 25388–25392 (1998).
25. Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66–71 (1994).
26. Sibanda, B.L. *et al.* Crystal structure of an Xrcc4–DNA ligase IV complex. *Nat. Struct. Biol.* **8**, 1015–1019 (2001).
27. Clapperton, J.A. *et al.* Structure and mechanism of BRCA1 BRCT domain recognition of phosphorylated BACH1 with implications for cancer. *Nat. Struct. Mol. Biol.* advance online publication, 9 May 2004 (doi:10.1038/nsmb775).
28. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–325 (1997).
29. Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022–1025 (1997).
30. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta. Crystallogr. A* **47**, 110–119 (1991).
31. Terwilliger, T.C. SOLVE and RESOLVE: automated structure solution and density modification. *Methods Enzymol.* **374**, 22–37 (2003).
32. Winn, M.D., Isupov, M.N. & Murshudov, G.N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta. Crystallogr. D* **57**, 122–133 (2001).
33. Laskowski, R.A., MacArthur, M.W. & Thornton, J.M. Validation of protein models derived from experiment. *Curr. Opin. Struct. Biol.* **8**, 631–639 (1998).
34. Merritt, E.A. & Bacon, D.J. Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524 (1997).
35. Esnouf, R.M. An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Model.* **15**, 112–113, 132–134 (1997).