

---

# Structural Basis for Phosphorylation-dependent Signaling in the DNA Damage Response

J. N. Mark Glover

Department of Biochemistry, University of Alberta, Edmonton

## Abstract

The response of eukaryotic cells to DNA damage requires a multitude of protein-protein interactions that mediate the ordered repair of the damage and arrest of the cell cycle until repair is complete. Two conserved protein modules, BRCT and FHA domains, play key roles in the DNA damage response as recognition elements for nuclear Ser/Thr phosphorylation induced by DNA damage-responsive kinases. BRCT domains, first identified at the C-terminus of BRCA1, often occur as multiple tandem repeats of individual BRCT modules. Our recent structural and functional work has revealed how BRCT repeats recognize phosphoserine protein targets, and have revealed a secondary binding pocket at the interface between tandem repeats that recognizes the amino acid 3 residues C-terminal to the phosphoserine. We have also studied the molecular function of the FHA domain of the DNA repair enzyme, PNK. This domain interacts with threonine-phosphorylated XRCC1 and XRCC4, which is responsible for the recruitment of PNK to the sites of DNA strand break repair. Our studies have revealed a flexible mode of recognition that allows PNK to interact with numerous negatively charged substrates.

## Introduction

Cells have evolved to deal with a bewildering complexity of DNA damage, from relatively small base lesions to single and double strand breaks in the DNA backbone. The repair of DNA damage in general involves multiple steps and several enzymatic activities. The DNA intermediates generated en route to repair are believed to often be more mutagenic than the original damage, and it appears

that scaffold proteins are utilized to efficiently shuttle these DNA intermediates between the various repair enzymes that are required at successive steps in the repair process. In addition, specific cell cycle checkpoints can effectively delay progression through the cell cycle in response to DNA damage, to allow the cell time to repair damage before replication.

Very recent work has uncovered unique protein signaling modules that are specific to the DNA damage response. Here I review recent structural and functional work from our laboratory that reveals fundamental principles of phospho-protein recognition utilized by BRCT and FHA proteins – modules that are key to the regulated cellular response to DNA damage. Our studies reveal the structural basis for why specific mutations in BRCA1 BRCT repeats are associated with breast and ovarian cancer, and could be used to develop new cancer therapies that modulate these interactions.

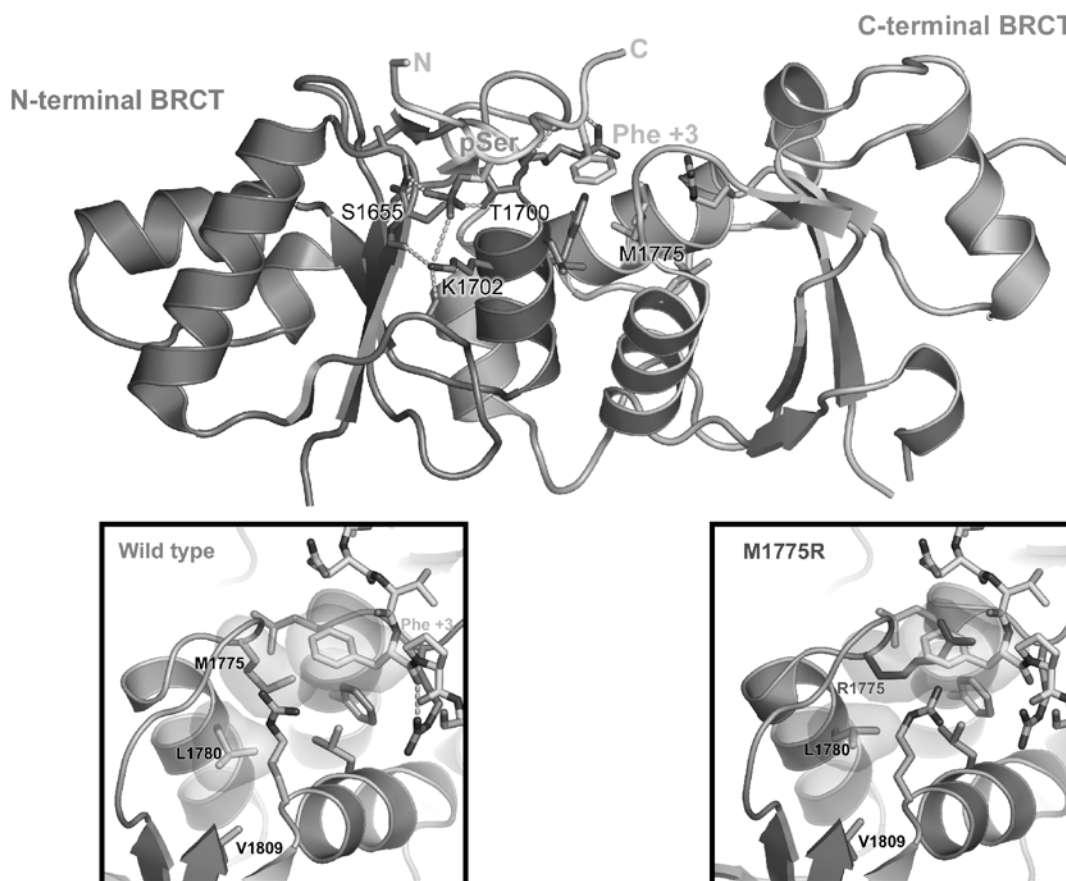
## BRCT repeats

BRCT repeats were first discovered at the extreme C-terminus of the breast cancer-associated protein, BRCA1, but were later shown to exist in a large family of proteins that are linked to the cellular response to DNA damage (1-4) (see <http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00533> for the Pfam listing of the BRCT family). The BRCT repeats in BRCA1 are essential for the tumor suppressor function of the protein as protein truncation and missense variants within the BRCT have been shown to be associated with human breast and ovarian cancers (reviewed in (5)).

We determined the X-ray crystal structure of the two BRCT repeats of BRCA1 ((6), Figure 1). The structure revealed that the two repeats adopt similar folds, and pack together in a head-to-tail manner. This interaction is stabilized by the packing of a single helix from the N-terminal repeat against a pair of helices from the C-terminal repeat and the inter-repeat linker. While this structure alone did not reveal the molecular function of the protein, we nevertheless could use it as a basis to predict the effects of specific mutations, with the underlying hypothesis that the function would be disrupted by mutations that disrupt the structure of the domain. We also developed a simple proteolytic

assay to directly test the conformational stability of BRCA1 BRCT variants (5).

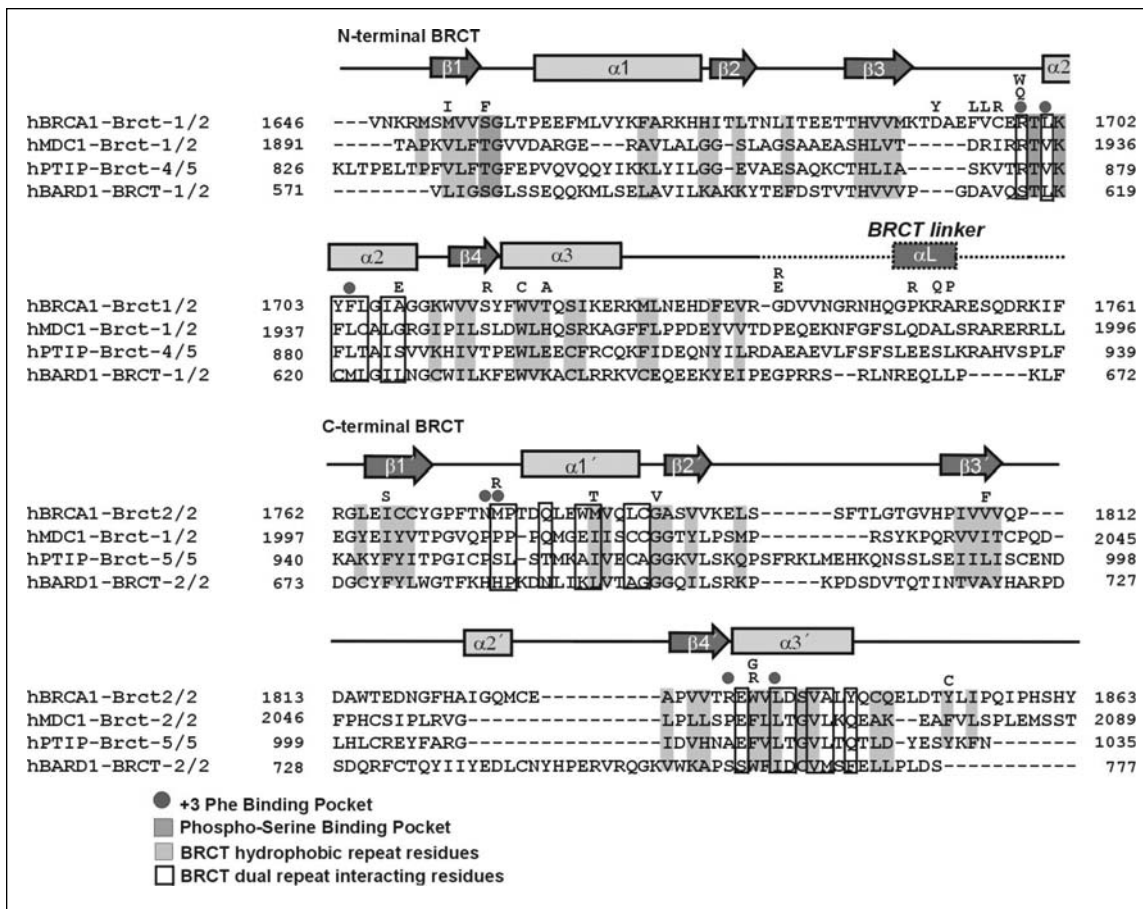
While most of the cancer-causing missense variants cause a dramatic destabilization of the BRCT fold, some of the mutations led to a more modest increase in proteolytic susceptibility. One of these mutations, M1775R, was particularly interesting as this was one of the first characterized BRCT mutations to be linked to cancer(7, 8). We were able to crystallize and determine the structure of this variant, which revealed a subtle rearrangement of side chains around the site of the mutation, and a perturbation of the surface features of the protein (Figure 1) (9). However, a key question remained:



**Figure 1. Structures of the BRCA1 BRCT domain and interactions with phosphopeptides**

Top panel. Structure of the BRCA1 BRCT domain bound to a pSer-x-x-Phe peptide target. The N- and C-terminal BRCT repeats, the inter-repeat linker and the bound peptide are distinguished by different shadings. Residues important for peptide recognition are shown as sticks.

Bottom panels. Details of the Phe +3 binding pocket in the wild type structure (left), the M1775R variant (right). Note that the M1775R and V1809F structures were determined in the absence of a bound peptide. The peptide shown in the M1775R panel is overlaid from the wild type structure.



**Figure 2. Sequence alignment of BRCT repeat proteins**

The amino acid sequence alignment of the tandem BRCT repeats from BRCA1, MDC1, PTIP and BARD1 is shown, with the secondary structure from the BRCA1 BRCT crystal structure. The residues that contact the pSer are shaded, and the residues that form the Phe binding pocket are marked by circles. Missense mutations which have been assayed for interactions with the peptide are displayed above the sequence. Residues involved in inter-repeat BRCT interactions are boxed. Adapted from (14).

is the defect associated with this mutation due to the somewhat reduced stability of this variant, or does the alteration of the protein surface perturb an important interaction surface?

Late in 2003, a major breakthrough in our understanding of BRCT function came when work from the laboratories of Junjie Chen and Mike Yaffe revealed that the tandem BRCT repeats of BRCA1 functions as a phospho-peptide binding module (10, 11). The BRCA1 BRCT is highly selective for the sequence pSer-X-X-Phe. These interactions were shown to mediate the association of BRCA1 with the DNA helicase BACH1, which is essential for the correct functioning of the G2/M cell cycle checkpoint (10, 12), and with the tran-

scriptional co-repressor, CtIP (13). These authors also showed that BRCT repeats from other proteins also functioned as phospho-peptide binding modules, implying that phospho-peptide binding might be a conserved function for BRCT repeats in the DNA damage response.

To understand how the BRCA1 BRCT recognizes its specific phospho-peptide target, we determined the structure of the BRCA1 BRCT bound to a high affinity peptide derived from an in vitro peptide selection experiment containing the pSer-X-X-Phe motif (14) (Figures 1 and 2). Similar structures have also been determined by other groups (15, 16). The structure revealed a phospho-serine binding pocket in the N-terminal BRCT repeat,

---

and a phenylalanine binding pocket in a groove formed at the interface between the two repeats. Key residues that constitute the phospho-serine pocket include Ser 1655, Gly 1656, and Lys 1702, which all directly recognize the phosphate moiety. Thr 1700 also plays an important role in phosphate recognition, as it hydrogen bonds with Ser 1655, keeping the serine hydroxyl in a rigid orientation appropriate for phosphate recognition. This [Ser-Gly...Thr-X-Lys] motif is conserved in a variety of other BRCT repeats, suggesting that many of these proteins will also bind phospho-serine containing peptides (4, 14) (Figure 2). Indeed, a number of other BRCT-containing proteins, including PTIP, 53BP1, and BARD1, have now been shown to have phospho-peptide binding activity(10, 11, 17).

To directly test the phospho-serine binding capacity of our set of 25 BRCT missense variants, we assayed the ability of a pSer-X-X-Phe peptide to specifically pull down *in vitro* transcribed/translated BRCT variants, compared to a non-phosphorylated control. Our results demonstrated that the structural integrity of the BRCT domain was required for phospho-peptide recognition, as none of the missense variants that highly destabilized the protein fold specifically bound the pSer peptide. The results confirmed the importance of the residues of the phosphate recognition pocket as mutation of any of these residues to alanine completely destroyed the ability of the protein to bind phospho-peptide (14).

The Phe at the +3 position relative to the pSer is bound in a deep groove at the interface between the two repeats (Figure 1). This interaction explains the fact that both BRCT repeats are needed for peptide binding, and loss of part of the C-terminal repeat is associated with hereditary breast cancer. BRCT domains are most commonly found as tandem repeats in other proteins involved in the DNA damage response, and sequence analysis suggests that the head-to-tail packing of the BRCT repeats in BRCA1 is conserved in other repeats (4, 6). Thus it is likely that this secondary recognition groove at the repeat interface is conserved in other BRCT repeat proteins. However, lack of sequence conservation of residues that line

the groove suggest that the peptide binding specificity of other BRCT repeat binding proteins may be different.

The importance of several of the residues that line the phenylalanine binding pocket was also tested using the pull down assay (14). Mutations of many of these residues were found to disrupt specific phospho-peptide interactions. For example, mutation of Arg 1699, which interacts with the backbone of the phenylalanine, to either Trp or Gln, completely abrogated binding. The cancer-associated mutation, M1775R, also resulted in a complete loss of peptide binding. Interestingly, M1775 lies at the bottom of the inter-repeat groove and makes close van der Waals contact with the phenylalanine. Superposition of the M1775R mutant structure on the structure of the BRCT-peptide complex reveals that the substituted arginine side chain occupies the phenylalanine binding pocket, suggesting that this mutant is impaired in binding pSer -X-X-Phe targets (14) (Figure 1).

Significant advances have been made in the understanding of the molecular function of the BRCA1 BRCT repeats, and these studies have allowed us to understand the molecular defects associated with a number of previously uncharacterized sequence variants in this region of the protein. However, it remains to be shown what the function of phospho-serine binding is in the context of the intact protein. Do these interactions facilitate rearrangements in the overall structure and organization of the BRCA1 complex? Do these interactions affect that ubiquitin ligase activity of BRCA1? Answers to these questions will not only help to detail molecular mechanisms that underlie hereditary breast cancer, but will help to define mechanisms that regulate the cellular response to DNA damage.

### **FHA domains**

FHA (Forkhead-associated) domains are ~80-100 amino acid domains that play key signaling roles in multiple cellular processes such as signal transduction, transcription, vesicular transport, and protein degradation, in addition to the DNA damage response (18). The structures of a number of FHA domains are now known, both free and in complex with phospho-peptide targets, and studies of the

FHA domains from the DNA-damage response proteins Rad53 and Chk2 have been particularly informative in revealing the peptide binding specificities of these proteins. In general, FHA proteins differ from BRCT domains in that they are highly specific for phospho-threonine, as opposed to phospho-serine-containing targets.

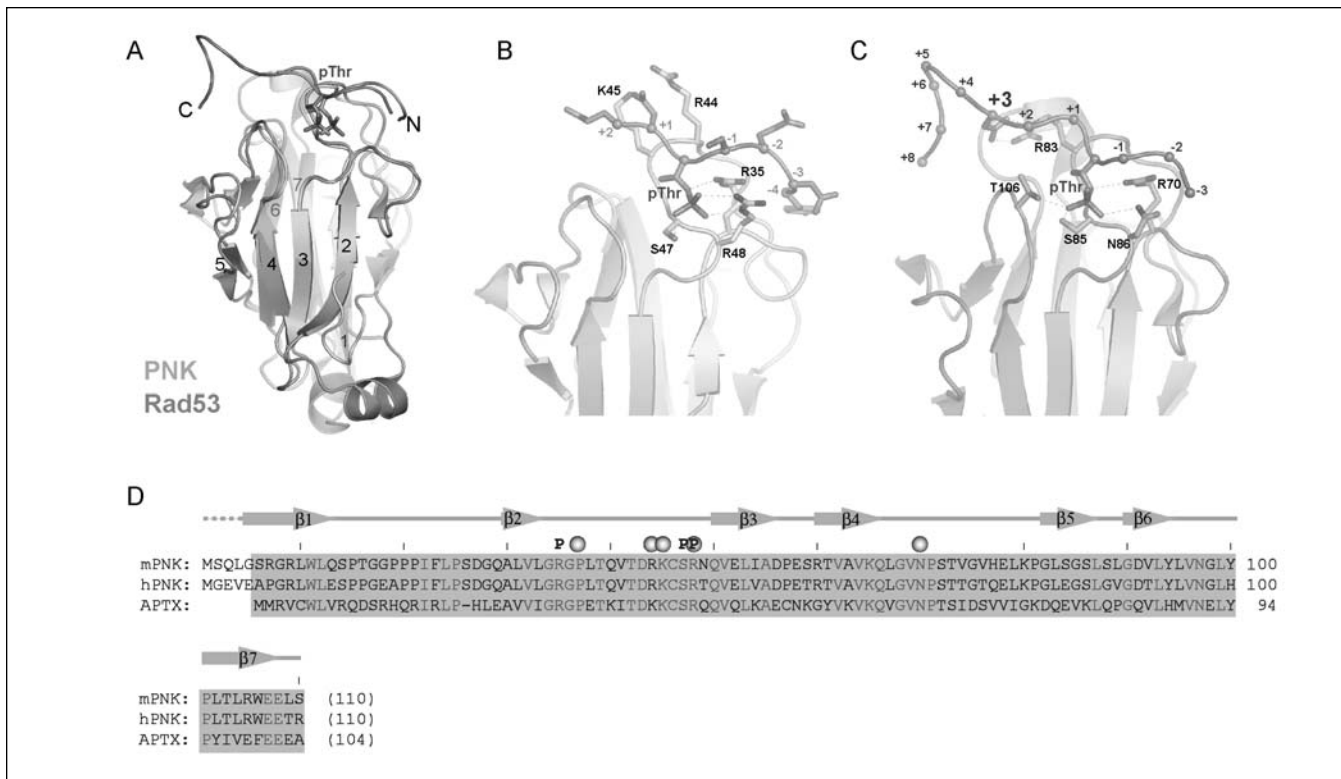
However, like BRCT domains, the Rad53 and Chk2 FHA domains exhibit marked selectivity for specific sidechains 3 residues C-terminal to the phosphorylated residue. In the case of the N-terminal FHA domain from Rad53, the specificity is for negatively charged residues at this position, which is recognized by an arginine residue (Arg 83, Figure 3) (19). In contrast, the Chk2 FHA shows selectivity for hydrophobic residues at the pThr+3 position, which is bound in a largely hydrophobic

pocket (20).

### FHA function in Polynucleotide kinase (PNK)

PNK is a bifunctional 5'-kinase/3'-phosphatase that is responsible for the processing of damaged DNA ends at both double and single stranded breaks, as well as in base excision repair pathways(21-24). PNK activity is often essential to restore the 5'-phosphate/3'-hydroxyl termini that are required for DNA polymerase and ligase activities to complete the repair of these lesions.

PNK contains, in addition to its catalytic kinase and phosphatase domains, an N-terminal FHA domain, which targets the enzyme to sites of repair and exhibits a different substrate selectivity than previously studied FHA domains. The FHA



**Figure 3. Phospho-threonine peptide recognition by the FHA domain of PNK**

A. Structural alignment of the FHA domain – peptide complexes from Rad53 (dark grey) and PNK (light grey). Strands are numbered according to the order in the PNK FHA.

B and C. Details of PNK-phospho-peptide (B) and Rad53-phospho-peptide (C) recognition . FHA sidechains that make critical interactions with the peptide are shown as sticks and the FHA domains are in the same orientation for comparison. Note that the critical interaction involving Arg 83 which specifies an aspartate at the pThr+3 position in Rad53 is missing in PNK. D. Sequence alignment of FHA domains from PNK and APTX. P indicates residues that recognize the target pThr, while the green balls indicate other residues that contact the peptide.

---

domain of PNK specifically recognizes CK2 phosphorylated forms of the single strand break repair scaffold protein, XRCC1, and the double strand break repair protein XRCC4, interactions which are required to direct PNK to sites of damage(21, 23). Peptide array binding experiments have demonstrated that the PNK FHA, unlike other FHAs, exhibits its highest degree of selectivity for sequences N-terminal to the phospho-threonine, rather than C-terminal (23).

To understand the structural basis for the recognition of phosphorylated XRCC4 and XRCC1 by the PNK FHA, we determined the structure of this domain bound to an XRCC4-derived phosphopeptide (Ac-YDES(pT)DEESEKK-CONH<sub>2</sub>, Figure 3)(25). The structure reveals that the phospho-threonine residue is recognized in the same way as in other FHA domains, utilizing the invariant Arg 35 and Ser 47, as well as the highly conserved Arg 48, which each ligate the phosphate group.

Comparison of the sequences surrounding the phospho-threonine in the XRCC1 and XRCC4 target sequences reveal a predominance of acidic residues both N- and C-terminal to the phospho-threonine. In XRCC1, there is the additional possibility of multiple phosphorylations by the acidophilic kinase CK2 in the vicinity of the primary site of phosphorylation (21). This, together with peptide binding array studies, has indicated that the FHA preferentially binds acidic target peptides. The structure of the PNK FHA explains this preference (Figure 3B). The peptide binding surface is highly positively charged; key basic residues within this surface are Arg 44, Lys 45 and Arg 48, which are presented on two loops which together comprise a peptide-binding cradle. Arg 48, in addition to its role in phosphate recognition, is also poised to select for the acidic Asp at the peptide pThr-3 position. Arg 44 adopts different orientations in each of the 3 distinct FHA-peptide complexes in the crystallographic asymmetric unit, indicating a significant degree of flexibility in this residue that could allow it to interact with either the pThr-2 or pThr+1 positions. In the XRCC4 target, both the -2 and +1 residues are negatively charged. Arg 44 and Arg 48 are essential for peptide recognition as mutation of these residues (to alanine for Arg 44,

or asparagine for Arg 48) completely destroys the ability of the PNK FHA to bind the XRCC4 phospho-peptide. Lys 45 could provide additional electrostatic recognition for negatively charged residues at pThr+1 and pThr+2, however, mutation of this residue to alanine has a negligible affect on the peptide binding affinity, suggesting that this residue plays only a minor role, if any, in phospho-peptide selection. There is no interpretable electron density for the peptide chain C-terminal to the pThr+2 position, indicating that this region of the target is highly mobile and, in contrast to other the peptide targets of other FHAs, is not bound by the FHA domain.

The PNK FHA is quite distinct at the amino acid sequence level from most other FHA domains, and this is reflected in its unusual mode of peptide recognition. The PNK FHA is however, highly similar to the FHA of aprataxin (APTX), a protein associated with the neurological disorder ataxia-oculomotor apraxia and likely also involved in DNA repair (26, 27) (Figure 3D). Interestingly, APTX also associates with both XRCC1 and XRCC4 and it has been proposed that competition between APTX and PNK for XRCC1 or XRCC4 may provide a mechanism for the regulation of PNK activity (28, 29).

The peptide-contacting residues, including Arg 44, Lys 45, and Arg 48, are all conserved in APTX, suggesting that the mode of peptide recognition in this protein could be similar to that seen in PNK. However, the pThr+3 position appears to be important in the APTX:XRCC1 interaction, since mutation of Glu to Ala at this site in XRCC1 abolished the binding to APTX (29). APTX has a single positively charged residue, Lys 75 (corresponding to Pro 81 in PNK), which could approach the pThr+3 residue to provide sequence selectivity at this position.

Interestingly, while the PNK FHA recognizes its negatively charged targets using a complementary, positively charged binding surface, all of the electrostatic interactions that we have identified between the FHA domain and the acidic residues N- or C-terminal to the pThr are relatively long (> 3.5 Å) and are expected to be relatively weak

(Figure 3B). This may allow the PNK FHA to recognize several similar but non-identical acidic target peptides in both XRCC1 and XRCC4. It may also be that we have not yet identified the highest affinity target for the PNK FHA. Peptide selection studies have not revealed dramatic binding preferences such as the preference of the BRCA1 BRCT for phenylalanine/tyrosine at the pSer+3 position. The fact that CK2, the kinase responsible for phosphorylation of both XRCC1 and XRCC4, phosphorylates clusters of residues in these targets in an autocatalytic manner, presents the possibility that the PNK FHA may bind most tightly to multiply phosphorylated targets.

### Implications for cancer therapy

Inhibitors of proteins involved in DNA repair and DNA damage-associated cell cycle checkpoints offer potential leads for new anti-cancer therapeutics. For example, inhibitors of BRCA1 could provide an avenue of attack for breast and ovarian tumours which are BRCA1+, but have been disabled in an alternative checkpoint system. PNK inhibitors might provide a means to increase the efficacy of traditional, DNA-targeting drugs by reducing the ability of the tumours to repair the damage induced by the therapy. Signaling processes involving BRCT and FHA domains have been shown to be essential in these systems and the detailed analysis of the principles that underlie phospho-peptide recognition in these systems could provide a basis for the rational design of inhibitors.

### References

1. E. V. Koonin, S. F. Altschul, P. Bork , *Nature Genet.* 13, 266 (1996).
2. P. Bork et al., *FASEB J* 11, 68 (1997).
3. I. Callebaut, J. P. Mornon, *FEBS Lett* 400, 25 (1997).
4. J. N. Glover, R. S. Williams, M. S. Lee , *Trends Biochem Sci* 29, 579. (2004).
5. R. S. Williams et al. *J Biol Chem* 278, 53007. (2003).
6. R. S. Williams, R. Green, J. N. Glover , *Nat Struct Biol* 8, 838. (2001).
7. P. A. Futreal et al., *Science* 266, 120 (1994).
8. Y. Miki et al., *Science* 266, 66 (1994).
9. R. S. Williams, J. N. Glover , *J Biol Chem* 278, 2630. (2003).
10. X. Yu, C. C. Chini, M. He, G. Mer, J. Chen , *Science* 302, 639. (2003).
11. I. A. Manke, D. M. Lowery, A. Nguyen, M. B. Yaffe , *Science* 302, 636. (2003).
12. S. B. Cantor et al., *Cell* 105, 149 (2001).
13. X. Yu, J. Chen , *Mol Cell Biol* 24, 9478. (2004).
14. R. S. Williams, M. S. Lee, D. D. Hau, J. N. Glover , *Nat Struct Mol Biol* 11, 519. (2004).
15. J. A. Clapperton et al. *Nat Struct Mol Biol* 11, 512. (2004).
16. E. N. Shiozaki, L. Gu, N. Yan, Y. Shi , *Mol Cell* 14, 405. (2004).
17. M. Rodriguez, X. Yu, J. Chen, Z. Songyang , *J Biol Chem* 278, 52914. (2003).
18. D. Durocher, S. P. Jackson , *FEBS Lett* 513, 58. (2002).
19. D. Durocher et al. *Mol Cell* 6, 1169. (2000).
20. J. Li et al. *Mol Cell* 9, 1045. (2002).
21. J. I. Loizou et al. *Cell* 117, 17. (2004).
22. C. Chappell, L. A. Hanakahi, F. Karimi-Busheri, M. Weinfeld, S. C. West , *EMBO J* 21, 2827. (2002).
23. C. A. Koch et al. *EMBO J* 23, 3874. (2004).
24. A. Rasouli-Nia, F. Karimi-Busheri, M. Weinfeld , *Proc Natl Acad Sci U S A* 101, 6905. (2004).
25. N. K. Bernstein et al. *Mol Cell* 17, 657. (2005).
26. N. Gueven et al. *Hum Mol Genet* 13, 1081. (2004).
27. M. C. Moreira et al. *Nat Genet* 29, 189. (2001).
28. P. M. Clements et al. *DNA Repair (Amst)* 3, 1493. (2004).
29. H. Luo et al. *Mol Cell Biol* 24, 8356. (2004).