Structure and Architecture of BRCT Domains

Subjects: Parasitology | Cell Biology

Contributor: José Peña-Guerrero , Celia Fernández-Rubio , Alfonso T. García-Sosa , Paul A. Nguewa

The human BRCT domain was first resolved from the crystal structure of the N-terminal BRCT of the X-ray repair cross-complementing protein 1 (XRCC1), determined by X-ray crystallography to a 3.2 Å resolution. Its tertiary structure features a central core of four-stranded parallel β-sheet (β1, β2, β3, and β4) flanked by two α-helices (α1 and α3) on the C-terminal end, a single α-helix (α2) on the N-terminal end, and two surface loops connecting β1 with α1 and α2 with β3 (the overall structure being β1-α1-β2-β3-α2-β4-α3). BRCT domains have been identified in a wide group of living organisms (from bacteria, parasites to mammals) and viruses. As mentioned above, those domains take part in a variety of important cell processes including DDR and cell cycle control. In addition, a few of these protein modules have been shown to be involved in pathologies such as cancer or infectious diseases including leishmaniasis. Therefore, significant efforts have been made towards finding compounds able to specifically inhibit the functions of these protein domains.

BRCT infection disease cancer leishmania

1. Introduction

The protein encoded by the breast cancer susceptibility gene (*BRCA1*) contains at its C-terminal end two copies of a conserved 90–100 amino acids segment that was named BRCT (BRCA1 C-Terminus) domain. Over the years, BRCT domains have been identified in other proteins as well, located not only at C-terminal ends, but also in multiple or single copies [1]. Biochemical and sequence analyses suggest that BRCT domains mostly support protein-protein interactions and are typically linked with DNA repair, recombination, and cell cycle control [1][2][3]. Although there are BRCT domains present in Archaea, Bacteria, Eukarya, and even viruses, BRCT-containing proteins are more commonly observed in bacteria and eukaryotes [4], as it has been shown that the number of BRCT-bearing proteins within a genome increases according to the genome complexity [4].

Proteins, from bacteria to mammals, that bear BRCT domains are mostly implicated in cell processes such as DNA Damage Response (DDR), DNA repair, and/or cell cycle control [1][5]. They are also involved in the regulation of the metabolism of fatty acids, rRNA processing [6], and ribosome biogenesis [7][8]. BRCT domains can be found both as isolated single domains and assembled into complexes with a high variety of protein domains [4]. Accordingly, BRCT-bearing proteins are involved in diverse cell processes, including those related to diseases development including oncogenesis. BRCT domains were first identified in the breast cancer suppressor BRCA1. It is well known that protein domains such as the BRCA1 C-terminal (BRCT) domain mediate in DNA damage response (DDR) signalling events. Moreover, variations in the DDR pathway are linked to certain cancer development.

BRCT domains are not only related to cancer development but are also linked to other pathological processes. It was demonstrated that the expression of BRCA1 plays a neuroprotective role during brain ischemia/reperfusion (I/R) episodes [9]. These domains also play a role in pathogen biology and during infectious processes caused by viruses, bacteria, fungi, and parasites. In the human pathogen *Toxoplasma gondii* database, there are at least three putative BRCT-bearing proteins. However, it is still not known if these proteins are functional in the DSB repair pathway of this parasite.

In *Leishmania major*, the BRCT-harbouring protein *LmjPES,* a homolog of the oncogene *PES1,* has been shown to be linked to parasite infectivity. *LmjPES* is demonstrated to be overexpressed in the metacyclic promastigotes, the infective form for vertebrates. In fact, *LmjPES*-overexpressed parasites showed higher infectivity rates and increased their virulence capability in in vivo models. The footpads of mice inoculated with such parasites exhibited higher and faster swelling compared to those of animals infected with wild-type parasites. In addition, an increased induction of iNOS was detected in the inoculation area [10]. Therefore, the BRCT domain from LmjPES has been used as a target when searching for new treatments against leishmaniasis. Using a structure-based drug discovery strategy, a battery of new specific inhibitors targeting BRCT from LmjPES was reported. After in vitro validation, one of such compounds exhibited leishmanicidal activity against promastigotes and amastigotes of three *Leishmania* species without harming macrophages [11]. As previously described, BRCT domains are protein modules involved in cell cycle and DDR, and the relation between mutations in molecules involved in such mechanisms and the development of malignancies has been reported. Links between cancer and infectious diseases, including parasitic pathologies, have also been described. Recently, BRCT from *LmjPES* was expressed in mammalian cells. These transgenic mammalian cells expressing the BRCT domain from *Leishmania* parasites dramatically increased their growth rate and their resistance to DNA-damaging treatments such as etoposide and 5-fluorouracil [12]. In addition, those cells exhibited altered expressions of mitochondrial, proliferation, and chemoresistance genes. In vivo experiments in athymic nude mice revealed a significant capability of those transgenic cells to generate highly proliferative tumours. The aforementioned study reinforced the existing link between parasitism and cancer development through potential horizontal gene transfer [12]. In addition, it shed some light on the relevance of BRCT domains in parasites.

BRCT domains are DNA/protein binding modules linked to several functions such as DNA repair, cell cycle, or protein interaction, among others. To perform their roles, these domains interact with their partners in both phosphate-dependent and phosphate-independent manners.

2. BRCT Domains: Structure and Architecture

The human BRCT domain was first resolved from the crystal structure of the N-terminal BRCT of the X-ray repair cross-complementing protein 1 (XRCC1), determined by X-ray crystallography to a 3.2 Å resolution. Its tertiary structure features a central core of four-stranded parallel β-sheet (β1, β2, β3, and β4) flanked by two α-helices (α1 and α3) on the C-terminal end, a single α-helix (α2) on the N-terminal end, and two surface loops connecting β1 with α1 and α2 with β3 (the overall structure being β1-α1-β2-β3-α2-β4-α3) (**Figure 1**).

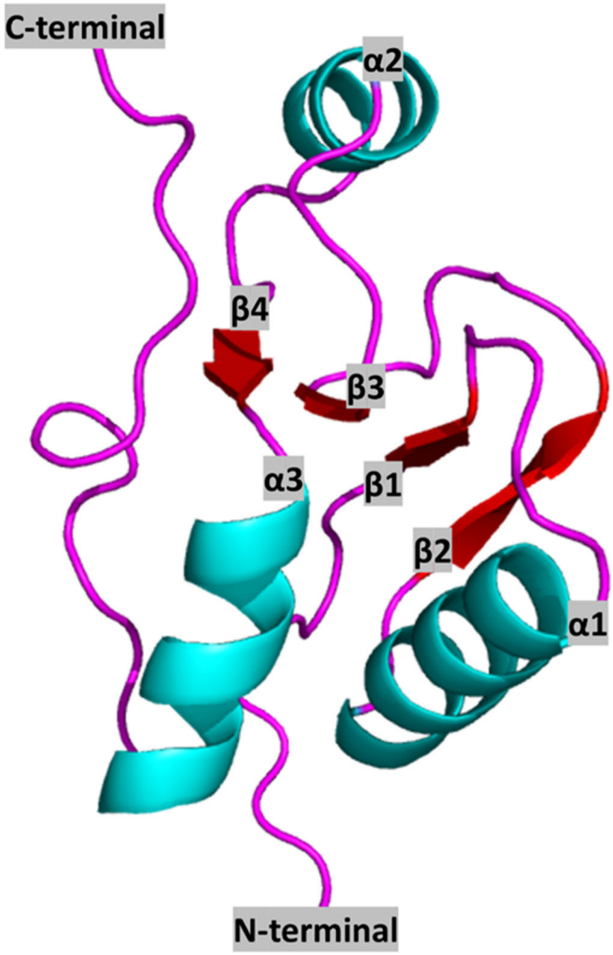

Figure 1. Representation of a canonical BRCT domain (from human XRCC1 protein) (PDB ID: 1CDZ) [13]. Colours represent secondary structure elements: red: β-sheets; cyan: α-helixes; pink: Loops. N and C terminals, as well as secondary structure elements, are annotated in black letters.

Hydrophobic interactions are formed between residues from α1 with those from β1 and β2 as well as residues from α2 with those from β4 [13]. In addition, five conserved hydrophobic motifs, named A, B, C, D, and E, can be found in β1, α1 and α1-β2 loop, β3, α3, and C-terminal end, respectively [4][5].

Analyses of alignments between representative BRCT sequences reveal very low identity, despite the predicted conserved structure. According to the amino acid sequence, the conserved hydrophobic amino acid clusters are located within the central β-sheet, on α1 and α3, and at the N- and C-terminal region of the domain [1][5]. Moreover, other highly conserved residues within the family are a double Gly-Gly motif located in the short loop between α1 and β2, and those which form helix α3 surface and are involved in the interaction between α1 and α3, suggesting that this two-stranded helical bundle is a crucial component of the domain (**Figure 1**).

Variations in BRCT structure and sequence are mainly located in α2, which varies in both size and amino acid composition [13] and is even absent in some members such as DNA Ligase III BRCT [14]. Numerous modifications are also detected in the surface loops connecting sheets and helices. These variable regions are supposed to form

stable protein-protein interaction surfaces [13]. In summary, the most conserved residues are located in central β-sheet, α1 and α3, while α2 is the most variable region within the BRCT family.

The phylogenetic relationships of the BRCT domain have been explored using PES BRCT as a model [11]. It is known that BRCT domains bear two motifs, I and II, and contain clusters of conserved residues. Motif I features a glycine-(glycine/arginine) doublet, while motif II includes a highly conserved tryptophan [1]. Both motifs are expected to be identifiable within the PES1 BRCT domain [15]. Although the highly conserved Gly-Gly position, which corresponds to the aforementioned motif I, is present in the alignment, the researchers were not able to identify the conserved tryptophan corresponding to motif II in trypanosomatids and *Plasmodium*, as it was substituted by the aromatic tyrosine. However, the presence of a conserved cluster of tyrosine-valine (short nonpolar amino acid), followed by a fully conserved core of glutamine-proline-glutamine and the previously mentioned tryptophan/tyrosine-leucine (short nonpolar amino acid), was found. Considering the aforementioned differences between human and trypanosomatid PES BRCT domains, as well as the gap in positions 35–45, the generation of selective LmjPES BRCT inhibitors that may not interact with human PES1 BRCT seems feasible [11].

In addition, the architecture of BRCT domains is highly variable. The top five common architectures are as follows: isolated; single domain + FHA (Forkhead associated); tandem domain + RING; two single domains + two homeodomain; and tandem domain + pleckstrin homology (PH) domain-like + Dbl homology (DH) domain [16]. Regarding the typical BRCT tandem—for example, the one found in BRCA1 protein—the domains fold in a head-to-tail orientation, engaging a wide hydrophobic interface composed of the α2 helix of the N-terminal repeat, which is packed against α1 and α3 of the C-terminal repeat. The conservation of α1 and α3 residues may indicate that this specific folding is shared by other proteins bearing this domain [17]. Nonetheless, there are variations of the typical tandem BRCT folding such as, for instance, the tandem BRCT domains of DNA Ligase IV (LigIV), which are separated by a significantly longer link [14].

BRCT domains have been associated with phosphopeptide-linking. A conserved phosphoserine-binding pocket which includes a Ser/Thr-Gly motif in the variable β1-α1 connecting loop (Ser1655 and Gly1656 of human BRCA1) and a Thr/Ser-X-Lys motif at the N terminus of α2 of the same repeat (residues 1700–1702 of human BRCA1) is present in those BRCTs predicted to have a phosphopeptide binding function [2]. In BRCA1, this structure creates a cleft in which the ligand (bearing a pSXXF motif) can be bound following a “two-anchor” interaction. pSer maintains polar interactions with conserved residues in β1 and α2, whereas +3 F sidechain inserts into a deep hydrophobic pocket of the interface between the first and second BRCT domain. Furthermore, the residues that coordinate the binding to pSer are highly conserved, in contrast to the residues that interact with pSer+3. Such variations of the residues interacting with pSer+3 may explain why different BRCT domains prefer different phosphopeptides, as well as their location at the pSer+3 position [18]. Additionally, residues located outside this pSXXF motif also play a role in this interaction, as different phosphopeptides sharing the same core DpSPVF sequence and peptide length, but bearing distinct N-terminal and C-terminal sequences, interact with BRCA1 tandem BRCT domains with significantly different affinities. Due to the higher number of hydrophobic interactions, those residues located in the N-terminal side contribute to the binding affinity [19].

Interactions with phosphopeptides have also been described in single BRCT, specifically the contact of phosphorylated Bloom syndrome RecQ-like helicase (BLM) with BRCT5 of murine Topoisomerase IIβ binding protein 1 (TopBP1) [20]. In this case, the phosphorylated serine of BLM may interact with Lys707, Ser 657, and Gln 658 of TopBP1, whereas the conserved motif Phe, Val, and Pro-Pro (situated -4, -3 and -2, and -1, respectively) can interact with Phe 681, Phe 682, Ala 710, Trp 714, and Met 692 from TopBP1 BRCT5 c2 loop which connects β2 and β3 sheets. Sun et al. also reported the binding of phosphorylated Mediator of DNA damage checkpoint protein 1 (MDC1) with TopBP1 in a special manner. The peptide from MDC1 appears to be trapped between two BRCT4/5 domains and interactions occur in the same amino acids from TopBP1, but with a reversed orientation in the phosphorylated MDC1 peptide [20]. Furthermore, single-BRCT interactions with its phosphorylated target occur with a dissociation constant (Kd) significantly higher than measured for tandem BRCT repeats. These data confirm the interactions of TopBP1 BRCT4/5 with BLM as a monomer and with higher affinity than MDC1 [20]. A similar mechanism is likely involved in other interactions such as TopBP1 with p-53 Binding Protein 1 (53BP1) [21].

Gómez-Cavazos et al. recently described a new BRCT interaction in the binding between the BRCT module from the Epithelial Cell Transforming Sequence 2 (ECT2) and the Rho family Gtpase CYK4 [22]. The basic surface, which is required for hosphor-CYK4 interaction and cytokinesis, is primarily composed of residues from the second domain (BRCT-1) of the triple BRCT module (BRCT-0, 1, 2) of ECT2, supporting the critical role of this single domain in the interactions with phosphorylated CYK4 [22]. Interestingly, this binding mode differs from that observed in tandem BRCT modules where the phosphopeptide binds across the interface formed between the two BRCT domains [17].

The presence of BRCT modules has been noted in many signalling proteins, such as the triple BRCT module present in DNA damage response scaffolding protein TopBP1. These results support the hypothesis that non-canonical interaction modes may mediate BRCT module functions in diverse contexts [22].

With regards to the structure of novel BRCT domains, the tertiary structure of the leishmanial PES BRCT domain was recently calculated via homology modelling [11]. BRCT domains usually show a hydrophobic core of β-sheets trapped between α-helixes with a typical β1, α1, β2, β3, α2, β4, α3 pattern [13] and exhibiting α2 variable among BRCT domains [13][14]. Overall, the constructed models maintained the aforementioned general structure. Furthermore, the hydrophobic motifs present in conventional BRCT domains [5], which would support the typical BRCT folding, were also found in those constructed models. It is well known that such hydrophobic interactions are important for the correct BRCT folding [13].

3. New and Ongoing Perspectives

Considering their relevant implications in key biological processes related to DNA repair and replication, the role of BRCT domains in human cancer development has been assessed extensively. Nevertheless, due to their structural plasticity and presence in other organisms, future research is required to study their potential pharmacological exploitation in cancer and other diseases including Neglected Tropical Diseases (NTDs) (https://doi.org/10.3390/pharmaceutics15071839). Those pathologies, such as leishmaniasis or African

trypanosomiasis, cause significant social and economic burdens and yet they still lack treatment with effective and selective drugs.

References

1. Bork, P.; Hofmann, K.; Neuwald, P.F.A.; Altschul, S.F.; Koonin, E.V. A Superfamily of Conserved Domains in DNA Damage-Responsive Cell Cycle Checkpoint Proteins. FASEB J. 1997, 11, 68–76.
2. Leung, C.C.Y.; Glover, J.N.M. BRCT domains: Easy as one, two, three. Cell Cycle 2011, 10, 2461–2470.
3. Gerloff, D.L.; Woods, N.T.; Farago, A.A.; Monteiro, A.N.A. BRCT domains: A little more than kin, and less than kind. FEBS Lett. 2012, 586, 2711–2716.
4. Sheng, Z.-Z.; Zhao, Y.-Q.; Huang, J.-F. Functional Evolution of BRCT Domains from Binding DNA to Protein. Evol. Bioinform. Online 2011, 7, 87–97.
5. Callebaut, I.; Mornon, J.P. From BRCA1 to RAP1: A widespread BRCT module closely associated with DNA repair. FEBS Lett. 1997, 400, 25–30.
6. Hölzel, M.; Grimm, T.; Rohrmoser, M.; Malamoussi, A.; Harasim, T.; Gruber-Eber, A.; Kremmer, E.; Eick, D. The BRCT domain of mammalian Pes1 is crucial for nucleolar localization and rRNA processing. Nucleic Acids Res. 2007, 35, 789–800.
7. Muhseena, N.K.; Mathukkada, S.; Das, S.P.; Laha, S. The repair gene BACH1—A potential oncogene. Oncol. Rev. 2021, 15, 519.
8. Grimm, T.; Hölzel, M.; Rohrmoser, M.; Harasim, T.; Malamoussi, A.; Gruber-Eber, A.; Kremmer, E.; Eick, D. Dominant-negative Pes1 mutants inhibit ribosomal RNA processing and cell proliferation via incorporation into the PeBoW-complex. Nucleic Acids Res. 2006, 34, 3030–3043.
9. Xu, P.; Liu, Q.; Xie, Y.; Shi, X.; Li, Y.; Peng, M.; Quo, H.; Sun, R.; Li, J.; Hong, Y.; et al. Breast cancer susceptibility protein 1 (BRCA1) rescues neurons from cerebral ischemia/reperfusion injury through NRF2-mediated antioxidant pathway. Redox Biol. 2018, 18, 158–172.
10. Algarabel, M.; Fernández-Rubio, C.; Musilova, K.; Peña-Guerrero, J.; Vacas, A.; Larrea, E.; Nguewa, P.A. In Leishmania major, the Homolog of the Oncogene PES1 May Play a Critical Role in Parasite Infectivity. Int. J. Mol. Sci. 2021, 22, 12592.
11. Peña-Guerrero, J.; Fernández-Rubio, C.; Burguete-Mikeo, A.; El-Dirany, R.; García-Sosa, A.T.; Nguewa, P. Discovery and Validation of Lmj_04_BRCT Domain, a Novel Therapeutic Target: Identification of Candidate Drugs for Leishmaniasis. Int. J. Mol. Sci. 2021, 22, 10493.

12. Larrea, E.; Fernández-Rubio, C.; Peña-Guerrero, J.; Guruceaga, E.; Nguewa, P.A. The BRCT Domain from the Homologue of the Oncogene PES1 in Leishmania major (LmjPES) Promotes Malignancy and Drug Resistance in Mammalian Cells. Int. J. Mol. Sci. 2022, 23, 13203.
13. Zhang, X.; Moréra, S.; Bates, P.A.; Whitehead, P.C.; Coffer, A.I.; Hainbucher, K.; Nash, R.A.; Sternberg, M.J.E.; Lindahl, T.; Freemont, P.S. Structure of an XRCC1 BRCT domain: A new protein-protein interaction module. EMBO J. 1998, 17, 6404–6411.
14. Glover, J.N.M.; Williams, R.S.; Lee, M.S. Interactions between BRCT repeats and phosphoproteins: Tangled up in two. Trends Biochem. Sci. 2004, 29, 579–585.
15. Haque, J.; Boger, S.; Li, J.; Duncan, S. The murine Pes1 gene encodes a nuclear protein containing a BRCT domain. Genomics 2000, 70, 201–210.
16. Wilson, D.; Pethica, R.; Zhou, Y.; Talbot, C.; Vogel, C.; Madera, M.; Chothia, C.; Gough, J. SUPERFAMILY—Sophisticated comparative genomics, data mining, visualization and phylogeny. Nucleic Acids Res. 2009, 37, 380–386.
17. Williams, R.S.; Green, R.; Glover, J.N.M. Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. Nat. Struct. Biol. 2001, 8, 838–842.
18. Rodriguez, M.C.; Songyang, Z. BRCT domains: Phosphopeptide binding and signaling modules. Front. Biosci. 2008, 13, 5905–5915.
19. Liu, X.; Ladias, J.A.A. Structural basis for the brca1 brct interaction with the proteins atrip and baat1. Biochemistry 2013, 52, 7618–7627.
20. Sun, L.; Huang, Y.; Edwards, R.A.; Yang, S.; Blackford, A.N.; Niedzwiedz, W.; Glover, J.N.M. Structural Insight into BLM Recognition by TopBP1. Structure 2017, 25, 1582–1588.e3.
21. Mer, G.; Botuyan, M.V. A New BRCT Binding Mode in TopBP1-BLM Helicase Interaction. Structure 2017, 25, 1471–1472.
22. Gómez-Cavazos, J.S.; Lee, K.Y.; Lara-González, P.; Li, Y.; Desai, A.; Shiau, A.K.; Oegema, K. A Non-canonical BRCT-Phosphopeptide Recognition Mechanism Underlies RhoA Activation in Cytokinesis. Curr. Biol. 2020, 30, 3101–3115.e11.