DNA Repair 35 (2015) 116–125

Contents lists available at ScienceDirect

DNA Repair

journal homepage: www.elsevier.com/locate/dnarepair

# Versatility in phospho-dependent molecular recognition of the XRCC1 and XRCC4 DNA-damage scaffolds by aprataxin-family FHA domains

Amy L. Cherry[a,1], Timothy J. Nott[a,2], Geoffrey Kelly[b], Stuart L. Rulten[b], Keith W. Caldecott[b], Stephen J. Smerdon[a,*]

[a] Francis Crick Institute, Mill Hill Laboratory, The Ridgeway, Mill Hill, London NW7 1AA, UK
[b] MRC Genome Damage and Stability Centre, University of Sussex, Brighton BN1 9RQ, UK

ARTICLE INFO

*Article history:*
Received 3 August 2015
Received in revised form 5 October 2015
Accepted 5 October 2015
Available online 23 October 2015

*Keywords:*
FHA domain
DNA break repair
DNA-damage signalling

ABSTRACT

Aprataxin, aprataxin and PNKP-like factor (APLF) and polynucleotide kinase phosphatase (PNKP) are key DNA-repair proteins with diverse functions but which all contain a homologous forkhead-associated (FHA) domain. Their primary binding targets are casein kinase 2-phosphorylated forms of the XRCC1 and XRCC4 scaffold molecules which respectively coordinate single-stranded and double-stranded DNA break repair pathways. Here, we present the high-resolution X-ray structure of a complex of phosphorylated XRCC4 with APLF, the most divergent of the three FHA domain family members. This, combined with NMR and biochemical analysis of aprataxin and APLF binding to singly and multiply-phosphorylated forms of XRCC1 and XRCC4, and comparison with PNKP reveals a pattern of distinct but overlapping binding specificities that are differentially modulated by multi-site phosphorylation. Together, our data illuminate important differences between activities of the three phospho-binding domains, in spite of a close evolutionary relationship between them.



## 1. Introduction

Genome integrity is under constant threat from a variety of endogenous and exogenous genotoxic agents, which create a broad spectrum of both single-stranded and double-stranded DNA lesions. Failure to repair such breaks can result in cell death or tumor development. Several complex repair pathways have evolved to resolve various types of DNA lesions; these involve stages of break detection, DNA end-processing, DNA gap filling and DNA ligation. XRCC1 is a crucial scaffolding protein in base excision repair that interacts with and coordinates many known components of that pathway. These include poly(ADP-ribose) polymerase 1 (PARP-1) [1,2], polynucleotide kinase phosphatase (PNKP) [3], aprataxin [4–8], aprataxin- and PNKP-like factor (APLF) [9,10], DNA polymerase β [1,11] and DNA ligase IIIα [12,13]. XRCC4 has a parallel role in the non-homologous end-joining (NHEJ) pathway for repair of double-stranded DNA breaks where it interacts with the catalytic subunit of DNA-dependent protein kinase (DNA-$PK_{cs}$) [14,15], PNKP [16], aprataxin [4], APLF [17,18] and DNA ligase IV [19,20].

Whilst many of the repair pathway components bind to different domains on the scaffolding proteins, three DNA-end modification proteins PNKP, aprataxin and APLF potentially compete for the same binding sites on XRCC1 and XRCC4. PNKP has dual activities; it phosphorylates 5′-OH termini and dephosphorylates 3′-phosphate termini which occur in >50% of breaks induced by oxidative stress [21,22]. Aprataxin removes AMP from 5′-adenylated DNA which can be formed when DNA ligation is aborted prematurely [23]. The precise role of APLF is yet to be determined although it does possess both endo- and exo-nuclease activity and its depletion is associated with impairment of NHEJ [18]. These three proteins share highly homologous forkhead-associated (FHA) domains (Fig. 1A), which have been shown to

*Abbreviations:* APLF, aprataxin- and PNK-like factor; PNKP, polynucleotide kinase phosphatase; FHA, forkhead-associated; XRCC1, X-ray repair cross-complementing protein 1; XRCC4, X-ray repair cross-complementing protein 4; PARP-1, poly(ADP-ribose)polymerase 1; NHEJ, non-homologous end-joining; DNA-PKcs, catalytic subunit of DNA-dependent protein kinase; CK2, casein kinase 2; HIT, Histidine triad; ZF, zinc finger; NTD, N-terminal domain; BRCT, BRCA1 C-terminal; CC, coiled-coil; ITC, isothermal titration calorimetry; APTX, aprataxin.

* Corresponding author.
*E-mail address:* Steve.Smerdon@crick.ac.uk (S.J. Smerdon).
[1] Institute of Science and the Environment, The University of Worcester, Henwick Grove, Worcester WR2 6AJ, UK.
[2] Baldwin Group, Physical & Theoretical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, UK.

http://dx.doi.org/10.1016/j.dnarep.2015.10.002

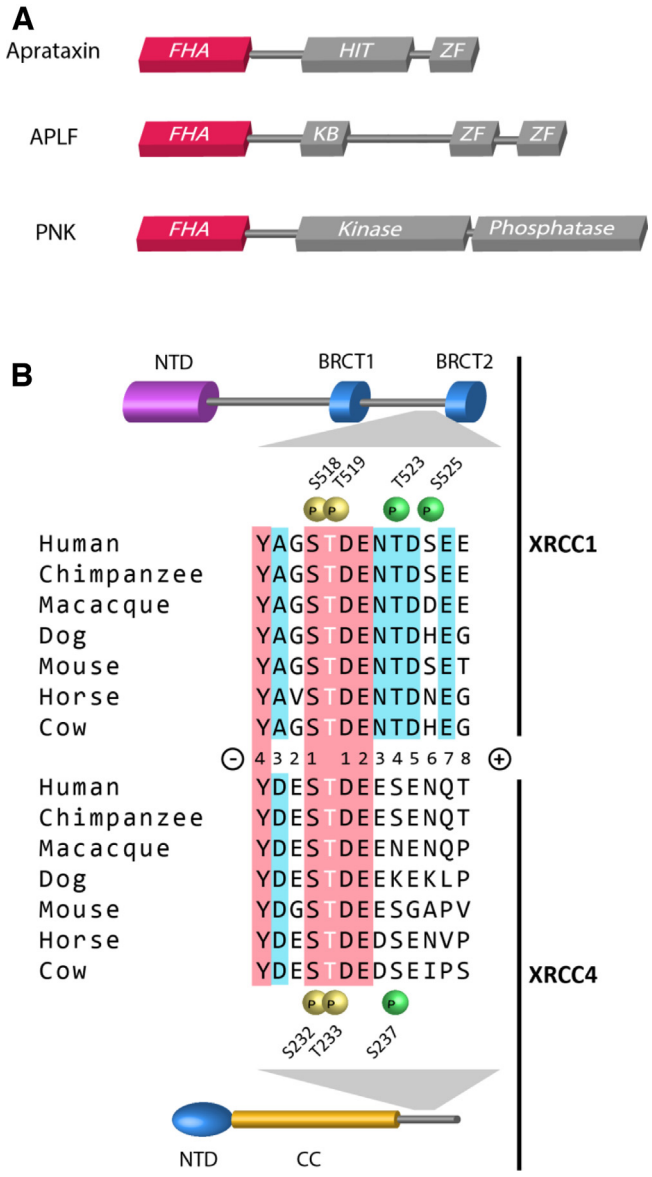


**Fig. 1.** Domain structure and CK2-phosphorylation sites of DNA repair proteins. (A) Schematic representation showing forkhead-associated (FHA), histidine triad (HIT), Ku-binding (KB), zinc finger (ZF), kinase and phosphatase domains of aprataxin, APLF and PNKP. (B) Schematic representation of N-terminal (NTD), BRCA1C-terminal (BRCT) and coiled-coil (CC) domains of XRCC1 and XRCC4 and interspecies sequence conservation of CK2 sites. Positions of CK2-phosphorylation sites in the core motif are denoted with yellow spheres and subsidiary CK2-sites denoted with green spheres. Core motif residues conserved in both XRCC1 and XRCC4 are highlighted in red and residues conserved within XRCC1 or XRCC4 are highlighted in light blue.

function as protein–protein interaction modules through their specific recognition of phosphothreonine-containing motifs on interacting partners [24–30]. All three bind XRCC1 and XRCC4 in a casein kinase 2 (CK2)-dependent manner [4,9,10,16–18,31]. Comparison of CK2 sites in XRCC1 and XRCC4 reveals a common YxxSTDE core motif, in which both serine and threonine are phosphorylated, with subsidiary sites present C-terminal to the core motif (Fig. 1B). Both PNKP and aprataxin FHA domains can bind a triphosphorylated peptide derived from this region of XRCC1 [7,32] and for PNKP each phosphate has been shown to contribute to binding affinity. Here, we systematically investigate the role of each of the XRCC1 and XRCC4CK2-phosphorylated residues in binding of the aprataxin and APLF FHA domains. The crystal structure of the APLF FHA domain bound to a triphosphosphorylated XRCC4 peptide, together with NMR titration experiments of aprataxin and XRCC1 peptides are used to explore the structural basis for multiple phospho-peptide binding. Together, these data reveal distinct but overlapping binding modes and specificities for this important family of DNA-damage responsive FHA domain proteins that are, in turn, differentially regulated by multi-site phosphorylation of their binding targets.

## 2. Materials and methods

### 2.1. Protein expression and purification

The genes for aprataxin residues 1–105 and APLF residues 1–106 were amplified using primer sets which incorporated 5′ BamHI and 3′ XhoI sites. PCR products and pGEX-6P-1 vector were digested with BamHI and XhoI and religated. All point mutations were generated using a QuickChange Site-Directed Mutagenesis Kit (Stratagene) following the manufacturer's recommended protocol.

For expression, the pGEX-6P-1/aprataxin-FHA constructs were transformed into *Escherichia coli* strain BL21 (DE3) and the pGEX-6P-1/APLF-FHA constructs transformed into the strain Rosetta2 (DE3). Cells were grown in LB at 37 °C to an $A_{600}$ of 0.5, induced with 0.4 mM isopropyl 1-thio-β-D-galactopyranoside and incubated at 18 °C for a further 16 h before harvesting. Cells were lysed by sonication and clarified by centrifugation (20,000 × *g* × 30 min). The supernatant was applied to a glutathione-4B resin (Amersham) and cleaved from the resin with rhinovirus 3C protease. Proteins were purified further by gel-filtration chromatography on a Superdex 75 matrix in 20 mM Tris pH 7.5, 150 mM NaCl, 5 mM DTT. Selenomethionine-labelled APLF L91 M protein was expressed in Rosetta2 (DE3) cells cultured in SelenoMet base media and nutrients supplemented with seleno-methionine solution (Molecular Dimensions Ltd.) and purified as the wild-type protein.

### 2.2. Isothermal titration calorimetry

Phosphopeptides based on the XRCC1 sequence 515-YAGSTDENTDSEEHQ-529 and the XRCC4 sequence 229-YDESTDEES-237 were synthesised with amidated C-termini to avoid potential end-effects of a free carboxy-terminus by Dr. W. Mawby (University of Bristol), purified by HPLC and characterised by mass spectrometry. FHA-phosphopeptide binding was quantified by isothermal titration calorimetry using a Microcal Omega VP-ITC calorimeter (MicroCal Inc., Northampton, MA). Protein was dialysed against ITC buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 5 mM β-mercaptoethanol) and peptides were dissolved in the dialysis buffer. Experiments were carried out at 22 °C and involved 30 successive 10 μl injections of peptide solution into a sample cell containing protein solution. Titrations with XRCC1 peptides used peptide at 1 mM and protein at 100 μM and titrations with XRCC4 peptides used peptide at 300 μm and protein at 30 μM. Heats of dilution were subtracted and binding isotherms were plotted and analysed with MicroCal origin version 7.0, assuming a single-site binding model.

### 2.3. Crystallization and structure determination

Selenomethionine-labelled protein and peptide were combined in a 1:3 ratio, with protein at a concentration of 10 mg/ml, in 50 mM HEPES pH 7.5, 150 mM NaCl, 5 mM β-mercaptoethanol. The complex crystallised from hanging drops set up at 18 °C with equal volumes of protein and reservoir solution 0.1 M Tris pH 8.0, 30% w/v PEG 3350, 0.2 M $MgCl_2$. Crystals grew within one week and were transferred into cryoprotectant (50 mM Tris pH 8.0, 15% w/v PEG 3350, 0.1 M $MgCl_2$, 75 mM NaCl, 25 mM HEPEs pH 7.5) and flash