Sequence analysis

A common ancestry for BAP1 and Uch37 regulators

Luis Sanchez-Pulido*, Lesheng Kong and Chris P. Ponting

MRC Functional Genomics Unit, Department of Physiology, Anatomy and Genetics, University of Oxford, South Parks Road, Oxford OX1 3QX, UK

Associate Editor: Alfonso Valencia

ABSTRACT

Summary: To reveal how the polycomb repressive–deubiquitinase (PR–DUB) complex controls substrate selection specificity, we undertook a detailed computational sequence analysis of its components: additional sex combs like 1 (ASXL1) and BRCA1-associated protein 1 (BAP1) proteins. This led to the discovery of two previously unrecognized domains in ASXL1: a forkhead (winged-helix) DNA-binding domain and a deubiquitinase adaptor domain shared with two regulators of ubiquitin carboxyl-terminal hydrolase 37 (Uch37), namely adhesion regulating molecule 1 (ADRM1) and nuclear factor related to kappaB (NFRKB). Our analysis demonstrates a common ancestry for BAP1 and Uch37 regulators in PR–DUB, as a transmitter between H2A ubiquitin and the chromatin remodeler.

Contact: luis.sanchezpulido@dpag.ox.ac.uk

Supplementary information: Supplementary data are available at Bioinformatics online.

Received on January 12, 2012; revised on May 18, 2012; accepted on May 23, 2012

1 INTRODUCTION

BRCA1-associated protein 1 (BAP1) and ubiquitin carboxyl-terminal hydrolase 37 (Uch37) (also known as UCHL5) are two deubiquitinating proteases (DUBs) from the same peptidase family that share a high degree of sequence similarity not only in their catalytic ubiquitin carboxy-terminal hydrolase (UCH) domain but also in an additional conserved carboxy-terminal region, previously called the Uch37-like domain (ULD) (Misaghi et al., 2009; Yao et al., 2006).

The controlled removal of ubiquitin or ubiquitin-like proteins by DUBs is required in diverse cellular processes. The catalytic domains of DUBs often require adaptor domains to control the specificity of substrate selection and thus to avoid cleavage of inappropriate substrates (Komander et al., 2009).

The human PR–DUB complex contains BAP1 and additional sex combs like 1 (ASXL1). The PR–DUB complex represses polycomb group target genes by removing monoubiquitin from histone H2A in nucleosomes. Recently, it was found that the ASXL1 N-terminal region (residues 2–365) is essential for its interaction with BAP1, although the domains responsible for this interaction have been determined in neither ASXL1 nor BAP1 protein sequences (Bott et al., 2011; Scheuermann et al., 2010).

The PR–DUB complex has increasingly attracted medical interest as somatic mutations in ASXL1 frequently occur in different myelodysplastic syndromes (Bejar et al., 2011; Boulwood et al., 2010; Caruccia et al., 2009; Gelsi-Boyer et al., 2009; Hellström-Lindberg, 2010). ASXL1 mutations have been also associated with glioblastoma, ovarian serous cystadenocarcinoma, cervical and colon cancer (Bell et al., 2011; McLendon et al., 2008; Scotto et al., 2008; Williams et al., 2010).

De novo nonsense mutations in human ASXL1 have also been recently associated with Bohring–Opitz syndrome (also known as Oberklaid–Danks or C-like syndromes), characterized by severe intellectual disability, distinctive facial features and multiple congenital malformations (Heissche et al., 2011).

Human ASXL1 belongs to a family of three vertebrate paralogs (Fisher et al., 2003, 2006; Katoth and Katoth, 2003, 2004) and it is orthologous to Drosophila Axs (additional sex combs), whose product is required for long-term repression of HOX genes during development and whose mutants enhance homotic phenotypes of both polycomb and trithorax mutations (Breen and Duncan, 1986; Milne et al., 1999; Smaclait et al., 1992). Mouse ASXL1 is required for normal hematopoiesis and its knockout results in simultaneous anterior and posterior transformations of the axial skeleton (Fisher et al., 2010a, b).

In addition, several lines of evidence suggest that ASXL1 has a genuine transcription factor role by regulating retinoic acid receptors and peroxisome proliferator-activated receptors (Cho et al., 2006; Lee et al., 2010; Park et al., 2011). ASX proteins have been previously reported to contain two evolutionarily conserved regions in metazoans, from fly to humans: a region termed the ASX homology (ASXH) domain and a C-terminal plant homeo domain (PHD) zinc finger (Fisher et al., 2003, 2006).

However, with the exception of its C-terminal PHD domain, significant sequence similarity was not detected with any other known protein domain. In order to trace its evolutionary history and to identify functional domains, we embarked on a computational protein sequence analysis of the ASX protein family.

2 RESULTS AND DISCUSSION

2.1 Computational protein sequence analysis. The DEUBiquitinase ADaptor domain

Our iterative profile-based database searches identified previously unreported homologues of ASXL1 in fungi and land plants. Animal ASX HMMe3 (Finn et al., 2011) profile searches (based on Supplementary Figure S1 multiple sequence alignment corresponding to the ASXH conserved region) revealed ASX homologues in fungi, identifying the Aspergillus fumigatus (Unipro:Q4WDB3_ASPFU) sequence with an E-value (HMMer3- best first domain) of 3.5 × 10^{-4}.

*To whom correspondence should be addressed.
Fig. 1. Domain architecture of DEUBAD domain-containing proteins and their known and predicted interactions with BAP1 and Uch37 DUBs. Experimentally determined interactions are shown with brown arrows corresponding to: A) An ASXL1 N-terminal region (amino acids 2 to 365) interacts with full length BAPI (Scheuermann et al., 2010; Bott et al., 2011); B) The hRpn13/ADRM1 ULD-interacting region was localized to its second domain (amino acids 278-391), here termed its DEUBAD domain (Hamazaki et al., 2006; Yao et al., 2006; Qiu et al., 2006; Chen et al., 2010); and C) The Uch37 ULD-interacting region of NFRKB was localized to its first 465 residues (Yao et al., 2008). Summary of the experimental evidences supporting each interaction: Scheuermann et al., 2010: Reconstituted complex in humans and Affinity Capture-Western for D. melanogaster orthologs (ASX and Calypso). Bott et al., 2011: Coimmunoprecipitation. Hamazaki et al., 2006: Affinity Capture-Western and reconstituted complexes. Yao et al., 2006: Yeast two-hybrid confirmed by Affinity Capture-Western and reconstituted complexes. Qiu et al., 2006: Affinity Capture-Western. Chen et al., 2010: Intermolecular NOE (Nuclear Overhauser Effect) interactions. Yao et al., 2008: Anti tag communoprecipitation and reconstituted complexes. Purple arrows indicate predicted interactions between DEUBAD domains and C-terminal ULDs in BAPI and Uch37 DUBs.

These approaches, however, failed to identify potential ASX orthologues in the nematode worm Caenorhabditis elegans or in yeasts (Saccharomyces cerevisiae and Schizosaccharomyces pombe). A profile derived from animal and fungal ASX sequences identified the Arabidopsis thaliana subfamily IV of GATA zinc fingers (Manfield et al., 2007): GATA27 (Tair:At5g47140, Uniprot: Q5PP38) and GATA26 (Tair:At4g17570, Uniprot: Q8W4H1) with E-values of 1.6×10^{-4} and 2.3×10^{-3}, respectively.

Consistent with the conservation of the Polycomb H2A ubiquitination machinery in plants (Bratzel et al., 2010; Sanchez-Pulido et al., 2008), the finding of ASX and BAPI homologues in Arabidopsis suggests that the histone H2A deubiquitination machinery (PR–DUB complex) could also be conserved between plants and animals. During a search for more remote evolutionary relationships, we were surprised to find statistically significant similarities with ASXL1 for two further human proteins. HMMer3 profile searches using the ASX family profile (corresponding to Supplementary Figure S1 multiple sequence alignment marked with the red bar) revealed the arthropod Pediculus humanus (UniProt: E0VM25_PEDHC) sequence (which is a known orthologue of the human NFRKB protein) with an E-value of 0.0073. The ASX/NFRKB profile (based on Supplementary Figure S1 multiple sequence alignment marked with the red and yellow bars) localized the Trichoplax adhaerens (UniProt: B3S476_TRIAD) sequence (a known orthologue of the human ADRM1 protein) with an E-value of 0.0067. In addition, profile-versus-profile comparisons (using HHpred) (Soding et al., 2005) confirm the previously described evidences of sequence similarity among these families: the profile of the ASX family identified the NFRKB and ADRM1 families with E-values of 2×10^{-7} and 5×10^{-6}, respectively. Statistically significant E-values of sequence similarity connected these human ASXL1,2,3, NFRKB and ADRM1 proteins and reciprocal searches produced convergent results; no further human homologues were identified.

NFRKB is a subunit of the INO80 chromatin-remodeling complex (Conaway and Conaway, 2009; Papamichos-Chronakis et al., 2011; Yao et al., 2008). The region of NFRKB that we observe to
be homologous to ASXL1 is located within a region previously described as binding the C-terminal Uch37 ULD (Fig. 1). The second human protein homologous to ASXL1 is ADRM1, whose tertiary structure shows it to consist of a bundle of eight alpha-helices (Chen et al., 2010). ADRM1 is a proteasome subunit that serves as a receptor for both ubiquitin and Uch37 (Chen et al., 2010; Hamaraki et al., 2006; Qua et al., 2006; Yao et al., 2006). Again, this ADRM1 region is known to bind the C-terminal Uch37 ULD and hence was initially called the Uch37-binding domain (Yao et al., 2006). However, owing to its affinity for a specific portion (the ULD of Uch37) and its presence as a DEUBiquitinate ADaptor (DEUBAD) domain in other proteins, we term these DEUBAD domains. The presence of the DEUBAD domain in ASXL1 and its presence within a region known to interact with BAP1 (Scheuermann et al., 2010) strongly implicates the ASXL1 DEUBAD domain in binding to the C-terminal BAPI ULD (Fig. 1).

2.2 Identifying a putative DNA-BINDING domain

Although these observations help to explain how the ASXL1 DEUBAD domain regulates BAPI deubiquitinase activity within PR-DUB, they fail to explain the proposed transcription factor function of ASXL1 (Cho PR–DUB, they fail to explain the proposed transcription factor DEUBAD domain regulates BAP1 deubiquitinase activity within a region known to interact with BAP1 (Scheuermann et al., 2010) and its presence as a DEUBiquitinase ADaptor (DEUBAD) domain in other proteins, we term these DEUBAD domains. The presence of the DEUBAD domain in ASXL1 and its presence within a region known to interact with BAP1 (Scheuermann et al., 2010) strongly implicates the ASXL1 DEUBAD domain in binding to the C-terminal BAPI ULD (Fig. 1). The second forkhead superfamily of transcription factors (AFX, PDB-ID:1e17, rate 0.68) (Boura et al., 2008) with a significant -value of 0.028 (estimated error rate <9%). Moreover, in support of the first match, the next five most statistically significant matches are members of the forkhead superfamily of transcription factors (AFX, PDB-ID:1e17, E-value = 0.089; FOXX1.a, PDB-ID:2cy2, E-value = 0.16; FOXX4, PDB-ID:3ic2, E-value = 0.35 and; NFSG, PDB-ID:1vm, E-value = 0.63) (Boura et al., 2010; Clark et al., 1995; Tsai et al., 2006). The consistency of secondary structure predictions and corroboration by profile-to-profile comparison (HHpred) methods, taken together provide strong evidence that animal ASX proteins contain a forkhead domain located at their N-termini.

By inspecting previously characterized forkhead domain structures in interaction with DNA (Brent et al., 2008; Boura et al., 2010; Clark et al., 1993; Stroud et al., 2006; Tsai et al., 2006), we identified two types of positions whose conservation would be in agreement with a DNA-binding role for the ASXL1 forkhead domain (Supplementary Fig. S2): (i) residues that have been implicated in non-specific DNA binding (DNA backbone contacts) are indicated by blue triangles above the alignment and (ii) indicated with black triangles and boxes are two residues located in alpha-helix 3 (Asn51 and His61; human ASXL1) which correspond to Asn165 and His169 of HNF3G (Clark et al., 1993) that are predicted to confer DNA-binding sequence specificity (major groove DNA contacts).

In agreement with a DNA-binding role for the ASXL1 forkhead domain, candidate ASX orthologues from plants contain, rather than a forkhead domain, a different, GATA zinc finger DNA-binding domain at their N-termini (Supplementary Fig. S2) (subfamily IV of Arabidopsis thaliana GATA zinc fingers; GATA26 and GATA27) (Manfield et al., 2007).

The human ASXL1 protein, therefore, likely binds DNA through its N-terminal forkhead domain thereby conferring DNA sequence specificity to the PR–DUB complex activity of histone H2A deubiquitination. In a similar manner, a forkhead domain that has recently been identified in the trithorax protein ASH2L appears to link sequence-specific DNA binding to histone H3K4 methylation (Chen et al., 2011; Sarvan et al., 2011).

Although we were finalizing this article, a bioinformatic analysis of ASXL1 was published (Aravind and Iyer, 2012), mainly focused on the identification of the N-terminal winged helix-turn-helix/forkhead domain. Our article is consistent with this report, but further extends the ASX family analysis by showing the structural and functional predictions for the DEUBAD domain.

3 CONCLUSION

In summary, the computational identification of the DEUBAD and forkhead domains in the human ASXL1 protein should now greatly assist in experimentally delineating its critical function in regulating the histone code under non-pathological and pathological conditions.

Funding: Medical Research Council UK (to L.s.-P. and C.P.P) and European Research Council (to L.K.).

Conflict of Interest: none declared.

REFERENCES


Common ancestry for BAP1 and Uch37 regulators

1955