

Biophysical Analysis of the N-Terminal Domain from the Human Protein Phosphatase 1 Nuclear Targeting Subunit PNUTS Suggests an Extended Transcription Factor TFIIS-Like Fold

Thomas Zacharchenko¹ · Igor Barsukov¹ · Daniel J. Rigden¹ · Daimark Bennett¹ · Olga Mayans^{1,2} 

Abstract Human protein phosphatase 1 nuclear targeting subunit (PNUTS) plays critical roles in DNA repair, cell growth and survival. The N-terminal domain of PNUTS mediates interactions with Tox4 and the phosphatase and tensin homolog PTEN, which are essential for the roles of this protein. To study this N-terminal domain, we have established its recombinant overproduction in *E. coli* utilizing NusA fusion. Upon removal of the tag, the remaining PNUTS sample is soluble and highly pure. We have characterized the domain using circular dichroism and nuclear magnetic resonance and analyzed its sequence using bioinformatics. All data agree in suggesting that the PNUTS N-terminal segment adopts a compact, globular fold rich in α -helical content, where the folded fraction is substantially larger than the previously annotated fold. We conclude that this domain adopts a single fold, likely being an extended form of the transcription factor S-II leucine/tryptophan conserved-motif. Thermal denaturation yielded a melting temperature of ~ 49.5 °C, confirming the stability of the fold. These findings pave the way for the molecular characterization of functional interactions mediated by the N-terminal region of PNUTS.

Keywords Recombinant protein overexpression · Secondary structure prediction · Thermal denaturation · Circular dichroism · Nuclear magnetic resonance

Abbreviations

CD	Circular dichroism
EDTA	Ethylenediaminetetraacetic acid
HSQC	Heteronuclear single quantum coherence
NMR	Nuclear magnetic resonance
PNUTS	Protein phosphatase 1-binding nuclear targeting protein
PP1	Protein phosphatase 1
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TCEP	Tris-2-carboxyethyl-phosphine
TFIIS	Transcription factor IIS
TFIIS	Structural motif of the transcription factor
LW-motif	IIS with conserved leucine and tryptophan residues

1 Introduction

Human PP1 nuclear targeting protein (PNUTS) is a protein phosphatase 1 (PP1) binding protein with critical functions in the response to cellular stresses, including DNA damage, and the regulation of RNA-polymerase II-mediated gene expression [1–5]. It forms a ternary complex with PP1, Tox4 and WDR82 that targets PP1 to the nucleus [3–10], and further interacts with the tumour suppressor phosphatase and tensin homolog PTEN [11]. Despite its significance to key transcriptional processes, PNUTS is poorly characterized. At the molecular level, PNUTS is a largely unstructured protein that contains two small folded domains, located at each of its termini. The N-terminal

✉ Thomas Zacharchenko
fbstz@leeds.ac.uk

✉ Olga Mayans
olga.mayans@uni-konstanz.de

¹ Department of Biochemistry, Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK

² Department of Biology, Universität Konstanz, Universitätstrasse 10, 78457 Constance, Germany

domain is predicted to be similar to the N-terminal transcription factor IIS (TFIIS) LW domain (so-called by the presence of invariant leucine and tryptophan residues; [12]) and binds to Tox-4 [6] and PTEN [11]. Such TFIIS LW domains are small four-helix bundles that are present in transcription factors such as MED26 and elongin A [12]. They are part of the larger TFIIS module that engages RNA-polymerase II [13, 14] and Tox-4 [6]. The PNUTS C-terminal region contains a zinc finger domain implying a possible interaction with nucleic acids [14], although this domain in PNUTS is not known to bind either RNA or DNA. The polypeptide region between the TFIIS LW and zinc finger domains in PNUTS is highly unstructured and plays a conserved role in binding to PP1 [2, 15, 16].

We have established the recombinant overproduction of the N-terminal region of human PNUTS in a soluble form and characterized it biophysically. Our analysis suggests that this N-terminal segment contains a larger fold than the currently annotated TFIIS LW-like domain. Knowledge of the correct boundaries of this domain provides now better guidance for molecular studies of PNUTS protein protein interactions.

2 Methods

2.1 Structure Prediction from Sequence Data

For the initial identification of structural homologs we employed the HHpred online server that uses hidden Markov models for comparative analysis of sequences [17–19]. The server also incorporates secondary structure prediction using the PSIPRED method [20]. In addition, we also used the Network Protein Sequence Analysis secondary structure prediction server (<https://npsa-prabi.ibcp.fr>) implementing the MLRC [21], DSC [22], and PHD predictive methods [23]. For identification of the putative fold we utilised the intensive search mode of the Phyre2 online server [24]. The latter employs hidden Markov models to generate multiple sequence alignments from protein structures deposited at the Protein Data Bank (www.rcsb.org) [24].

2.2 Molecular Biology

Plasmid DNA containing NusA-His₆-3C-His₆-PNUTS (UniProtKB Q96QC0) was purchased from the Medical Research Council Dundee Phosphorylation and Ubiquitination unit (product DU37545). NusA-His₆-3C-His₆-PNUTS^{1–158} was subcloned from the former using ligation independent cloning into the pOPINB vector (Oxford Protein Production Facility, UK). This vector incorporates an additional N-terminal His₆-tag prior to the insert,

resulting in the His₆-NusA-His₆-3C-His₆-PNUTS^{1–158} protein product. The clone was confirmed by sequencing (GATC-biotech).

2.3 Protein Expression and Purification

Protein expression was in *E. coli* strain BL21*(DE3) (Invitrogen) grown at 37 °C in Luria Bertani medium supplemented with 25 µg/ml kanamycin. At an OD₆₀₀ = 0.6, cultures were cooled to 18 °C, expression induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and cells further incubated for 16 h. Cells were harvested by centrifugation and resuspended in 20 mM sodium phosphate pH 7.4, 500 mM NaCl, 20 mM Imidazole, 3 mM β-Mercaptoethanol containing an ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor cocktail (Roche) and 1 mg/ml bovine deoxyribonuclease (Sigma). Cells were lysed using pressure homogenisation. Lysates were clarified by centrifugation and filtered using a 20 µm filter prior to fast liquid chromatography. Initial purification was by metal affinity chromatography in a 5 ml His-Trap HiPrep column (GE healthcare), with the protein eluted using a linear gradient of imidazole (0–500 mM). The sample was then buffer exchanged into 20 mM Tris pH 7.4, 150 mM NaCl, 3 mM β-Mercaptoethanol using a HiPrep 26/10 desalting column and, for tag removal, incubated overnight at 4 °C with PreScission proteaseTM (which was tagged with glutathione S-transferase; GE Healthcare). The cleaved His₆-NusA-His₆ fusion tag was removed by ion exchange capture in a 5 ml HiTrap Q HiPrep column (GE Healthcare), with precision protease and His₆-PNUTS eluting in the flow through. This flow through eluate was once again exchanged into 20 mM sodium phosphate pH 7.4, 500 mM NaCl, 20 mM Imidazole, 3 mM β-Mercaptoethanol and applied to a 5 ml His-Trap HiPrep column (GE healthcare), which captured His₆-PNUTS whilst precision protease did not bind to the column. His₆-PNUTS was then further purified on a 5 ml HiTrap S HP column to a purity of >95 % as revealed by SDS-PAGE.

For isotopic labelling, bacterial cultures were grown and induced in 2M9 media with 1 g of ¹⁵N-ammonium chloride (Sigma) added per 1 l medium.

2.4 Sample Preparation for Biophysical Analysis

Protein samples were buffer exchanged into the respective buffers using a PD-10 desalting column (GE Healthcare) and concentrated using a Millipore 3 kDa spin concentrator at 4400 rpm. Protein concentration was determined by A₂₈₀ using a Nano-drop 2000 spectrometer (Thermo Scientific).

2.5 Nuclear Magnetic Resonance (NMR) Spectroscopy

NMR spectroscopy was performed in 20 mM HEPES pH 7.4, 150 mM NaCl, 3 mM β -mercaptoethanol with 5 % [v/v] $^2\text{H}_2\text{O}$. Data were collected on an AVANCE II+ 800 MHz spectrometer (Bruker) equipped with CryoProbe at 298 K. For temperature titration, proton shifts were calibrated using trimethylsilyl propanoic acid (TSP) as an external standard. Figures were made using TopSpin 3.1 (Bruker).

2.6 Circular Dichroism (CD)

CD data were collected on a Jasco J-1100 spectrometer equipped with a JASCO PTC-348WI temperature control unit. Fresh protein samples were buffer exchanged into 10 mM sodium phosphate pH 7.4, 0.5 mM tris-2-carboxyethyl-phosphine (TCEP) and data collected at 0.5 mg/ml in a 0.2 mM path length quartz cuvette at a frame rate of 100 nm/min. Prior to deconvolution, control buffer spectra were subtracted and the data zeroed using the CD signal at $\lambda = 260$ nm. Data were deconvoluted using the Dichroweb server with the CDSSTR method [25, 26]. To measure thermal stability, CD spectra were collected in the spectral range $\lambda = 180$ –260 nm and in the temperature range 20–90 °C. Temperature was increased at a rate of 1 °C per minute and the sample equilibrated for 5 min at each integral degree before the recording of the corresponding spectrum. Data fit was performed using the Boltzman equation in PRISM 7.

3 Results

3.1 Prediction of the Existence of a Helical Domain at the N-Terminus of PNUTS

Human PNUTS is a 940-residue long protein with an annotated TFIIS LW domain close to its N-terminus. Its Interpro entry [27] reveals that different domain databases assign different regions to this domain: in Pfam (entry PF08711; [28]) it covers residues Q93-V143 while Smart (SM00509; [29]) and Prosite (PS51319; [30]) allocate it residues K73-Q147 approximately. Thus, there is currently no consensus on the start point of the domain. To identify the boundaries of the PNUTS TFIIS LW fold, we performed a secondary structure prediction from sequence data. The results highlighted two stretches (residues P8-F18 and V27-L56) of high helical propensity prior to the annotated motif (Fig. 1). In addition, both HHpred and Phyre2 servers identified the protein IWS1 from *E.uniculi*, of known atomic structure [13], as a distant homolog of the PNUTS N-terminal segment.

The structure of IWS1 is that of an extended TFIIS LW fold with an additional N-terminal helical fraction that packs against the core fold forming a helical bundle. All the predicted α -helices in PNUTS are amphipathic and might, therefore, pack against each other forming a compact helical bundle as that of IWS1. Thus, we considered possible at this stage that the additional N-terminal helices in PNUTS might also contribute to the tertiary fold of its domain.

3.2 Recombinant Production of the N-Terminal Domain of PNUTS

We produced the N-terminal fraction of PNUTS (residues 1–158) as a soluble and stable protein product in *E. coli* in the form of a fusion protein of the type His₆-NusA-His₆-3C-His₆-PNUTS^{1–158} (Fig. 2a). The His₆-NusA-His₆ tag was cleaved with PreScission protease and removed by ion exchange chromatography profiting from the differential pI values of NusA and PNUTS^{1–158} (PNUTS^{1–158} has pI = 9.43 and is positively charged at pH 6.5, while NusA has pI = 4.62 being negatively charged at that pH; pI values were calculated using Prot-param [31]). The remaining His₆-PNUTS^{1–158} sample, containing the non-cleavable N-terminal affinity tag, was separated from PreScission protease by nickel affinity chromatography and ion exchange chromatography. This protocol produced a PNUTS^{1–158} protein of high purity at a yield of ~10 mg/l *E. coli* culture. The sample migrated in SDS-PAGE at a molecular mass of approx. 19 kDa, consistent with the molecular mass calculated from sequence (19,091 Da) (Fig. 2b). The identity of the purified product was verified by mass spectrometry.

3.3 The N-Terminal Domain of PNUTS Folds into a Stable α -Helical Motif

To test bioinformatics predictions, we analysed the secondary structure composition of PNUTS^{1–158} experimentally using CD. The CD spectrum was characteristic of a helix-rich protein (Fig. 3a). We analysed the data using CDSSTR [26] available through Dichroweb (<http://dichroweb.cryst.bbk.ac.uk>). The spectral deconvolution suggested a protein fold consisting of approx. 56 % helical content and a negligible 8 % β -strand content. The normalized residual mean square difference of this estimation was 0.016 indicating a close fit of reconstructed and experimental spectra). This estimation of secondary structure content is in good agreement with sequence-based predictions (Fig. 1b) and supports the existence of an extended TFIIS LW fold. A shorter version of the domain fold as annotated in Pfam and Smart/Prosite would have resulted in strongly reduced helical contents of approx. 29 and 37 %, respectively.

Fig. 3 CD characterization of PNUTS¹⁻¹⁵⁸. **a** CD spectrum recorded in 10 mM sodium phosphate pH 7.4, 0.5 mM TCEP; **b** CD monitored thermal denaturation curve showing the change of CD signal at $\lambda = 208$ nm

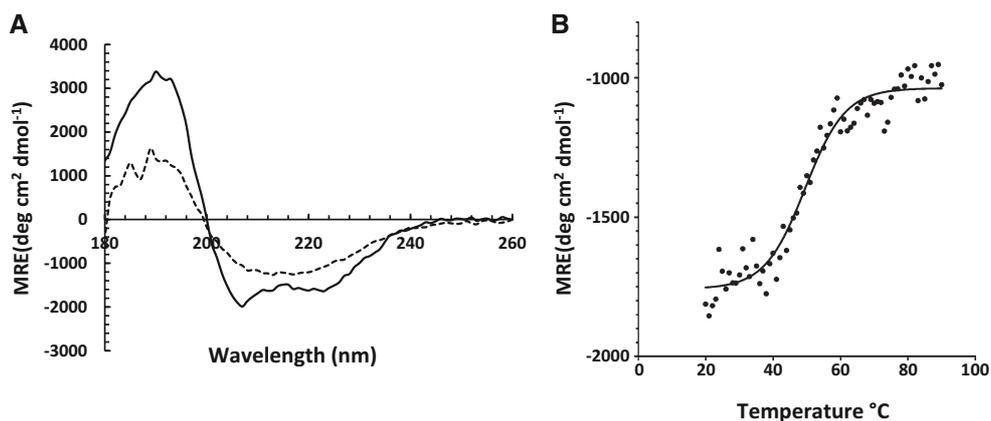
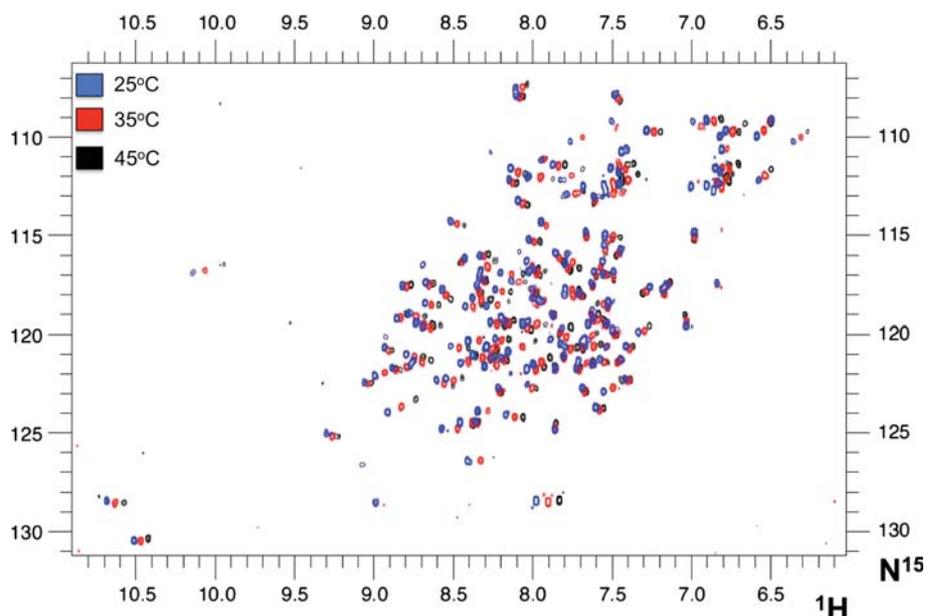


Fig. 4 (Colour online) Thermal denaturation of PNUTS¹⁻¹⁵⁸ monitored by NMR. Superposition of ¹H¹⁵N HSQC spectra of PNUTS¹⁻¹⁵⁸ at various temperatures in 20 mM HEPES pH 7.5, 150 mM NaCl, 3 mM β mercaptoethanol measured at 800 MHz. The linear change in chemical shifts of cross peaks upon temperature increase is clearly noticeable



Additional support for a single folded domain comes from the temperature dependence of the ¹H,¹⁵N-HSQC spectra (Fig. 4). Temperature increase resulted in uniform changes in the spectra that reflect a uniform thermal denaturation of a stable domain. Independent helices would have a different stability than a globular domain, manifesting in a selective broadening of a group of resonances upon temperature increase, which was not observed. On increasing the temperature from 25 to 35 °C, the resonances remained of equal intensity, with only resonance shift changes observed. At 45 °C, the majority of the resonances showed line broadening and a uniformly reduced intensity due to the increased exchange with solvent (Fig. 4). However, the chemical shift changes were linear in this temperature range, suggesting that the protein fold was uniformly destabilized. At this temperature the chemical shift changes were reversible and the original spectra detected upon reduction of the temperature to 25 °C, implying a reversible transitional state. A further

temperature increase to 55 °C led to irreversible protein denaturation with a complete loss of signals and visible precipitation of the sample. This suggests that the melting of this domain occurs completely within the temperature range 45–55 °C, in agreement with the T_m estimation from CD data. In summary, NMR-based findings support our conclusion that PNUTS¹⁻¹⁵⁸ forms a stable, single domain.

4 Discussion

Bioinformatics analysis of human PNUTS¹⁻¹⁵⁸ suggested that this N-terminal segment may contain a larger domain than the currently annotated TFIIS LW motif. To test this prediction, we expressed recombinantly and biophysically characterized the N-terminal region of PNUTS. Using CD and preliminary NMR data, we showed that PNUTS¹⁻¹⁵⁸ adopts a single helical fold, with a stability characterized by a T_m of ~ 49.5 °C. As the N-terminal segment of

PNUTS is thought to be involved in multiple protein-protein interactions, the accurate establishment of its domain boundaries is of central importance for future functional studies on PNUTS.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflicts of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Xing Z, Lin A, Li C, Liang K, Wang S, Liu Y, Park PK, Qin L, Wei Y, Hawke DH, Hung M, Lin C, Yang L (2014) lncRNA directs cooperative epigenetic regulation downstream of chemokine signals. *Cell* 159:1110–1125
- Kim YM, Watanabe T, Allen PB, Kim YM, Lee SJ, Greengard P, Nairn AC, Kwon YG (2003) PNUTS, a protein phosphatase 1 (PP1) nuclear targeting subunit. Characterization of its PP1 and RNA binding domains and regulation by phosphorylation. *J Biol Chem* 278:13819–13828
- Bennett D (2005) Transcriptional control by chromosome associated protein phosphatase 1. *Biochem Soc Trans* 33:1444–1446
- Ciurciu A, Duncalf L, Jonchere V, Lansdale N, Vasieva O, Glenday P, Rudenko A, Vissi E, Cobbe N, Alphey L, Bennett D (2013) PNUTS/PP1 regulates RNAPII mediated gene expression and is necessary for developmental growth. *PLoS Genet* 9:e1003885
- Dorn G (2013) miR-34a and the cardiomyopathy of senescence: SALT PNUTS, SALT PNUTS! *Cell Metab* 17:629–630
- Lee J, You J, Dobrota E, Skalnik DG (2010) Identification and characterization of a novel human PP1 phosphatase complex. *J Biol Chem* 285:24466–24476
- Allen PB, Kwon Y, Nairn AC, Greengard P (1998) Isolation and characterization of PNUTS, a putative protein phosphatase 1 nuclear targeting subunit. *J Biol Chem* 273:4089–4095
- Ramakrishnan R, Liu H, Donahue H, Malovannaya A, Qin J, Rice AP (2012) Identification of novel CDK9 and Cyclin T1 associated protein complexes (CCAPs) whose siRNA depletion enhances HIV-1 Tat function. *Retrovirology* 9:90
- Kapoor R, Sakshi A, Sanket SP, Binod K, Subbareddy M, Banerjee AC (2015) MicroRNA34a enhances HIV-1 replication by targeting PNUTS/PPP1R10 which negatively regulates HIV-1 transcriptional complex formation. *Biochem J* 470:293–302
- Landsverk HB, Mora Bermúdez F, Landsverk OJ, Hasvold G, Naderi S, Bakke O, Ellenberg J, Collas P, Syljuåsen RG, Küntziger T (2010) The protein phosphatase 1 regulator PNUTS is a new component of the DNA damage response. *EMBO Rep* 11:868–875
- Kavela S, Shinde SR, Ratheesh R, Viswakalyan K, Bashyam MD, Gowrishankar S, Vamsy M, Pattnaik S, Rao S, Sastry RA, Srinivasulu M, Chen J, Maddika S (2013) PNUTS functions as a proto-oncogene by sequestering PTEN. *Cancer Res* 73:205–214
- Ling Y, Smith AJ, Morgan GT (2006) A sequence motif conserved in diverse nuclear proteins identifies a protein interaction domain utilised for nuclear targeting by human TFIIIS. *Nucleic Acids Res* 34:2219–2229
- Diebold M, Koch M, Loeliger E, Cura V, Winston F, Cavarelli J, Romier C (2010) The structure of an Iws1/Spt6 complex reveals an interaction domain conserved in TFIIIS, Elongin A. *EMBO J* 29:3979–3991
- Klug A, Rhodes D (1987) Zinc fingers: a novel protein fold for nucleic acid recognition. *Cold Spring Harb Symp Quant Biol* 52:473–482
- Choy MS, Hieke M, Kumar GS, Lewis GR, Gonzalez DeWhitt KR, Kessler RP, Stein BJ, Hessenberger M, Nairn AC, Peti W, Page R (2014) Understanding the antagonism of retinoblastoma protein dephosphorylation by PNUTS provides insights into the PP1 regulatory code. *Proc Natl Acad Sci USA* 111:4097–4102
- Bennett D, Lyulcheva E, Alphey L, Hawcroft G (2006) Towards a comprehensive analysis of the protein phosphatase 1 interactome in *Drosophila*. *J Mol Biol* 364:196–212
- Hildebrand A, Remmert M, Biegert A, So J (2009) Fast and accurate automatic structure prediction with HHpred. *Proteins* 77:128–132
- Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33:244–248
- Söding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33:W244–W248
- Buchan DW, Minnici F, Nugent TC, Bryson K, Jones DT (2013) Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Res* 41(W1):W349–W357
- Guermeur Y, Geourjon C, Gallinari P, Deléage G, Lyon NSDe, Italie A, Pierre U (1999) Improved performance in protein secondary structure prediction by inhomogeneous score combination. *Bioinformatics* 15:413–421
- King RD, Sternberg MJE (1996) Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. *Protein Sci* 5:2298–2310
- Rost B, Sander C (1993) Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol* 232:584–599
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10:845–858
- Whitmore L, Wallace BA (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* 89:392–400
- Sreerama N, Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287:252–260
- Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, McAnulla C, McMenamin C, Nuka G, Pesseat S, Sangrador Vegas A, Scheremetjew M, Rato C, Yong SY, Bateman A, Punta M, Attwood TK, Sigrist CJ, Redaschi N, Rivoire C, Xenarios I, Kahn D, Guyot D, Bork P, Letunic I, Gough J, Oates M, Haft D, Huang H, Natale DA, Wu CH, Orengo C, Sillitoe I, Mi H, Thomas PD, Finn RD (2015) The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res* 43:D213–D221
- Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador Vegas A, Salazar GA, Tate J, Bateman A (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44:D279–D285
- Letunic I, Doerks T, Bork P (2015) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res* 43:D257–D260
- Sigrist CJ, de Castro E, Cerutti L, Cuče BA, Hulo N, Bridge A, Bougueleret L, Xenarios I (2013) New and continuing developments at PROSITE. *Nucleic Acids Res* 41:D344–D347
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy Server. In: Walker JM (ed) *The proteomics protocols handbook*. Humana Press, New York, pp 571–607