

The structure of TAX1BP1 UBZ1+2 provides insight into target specificity and adaptability

M. Angeles Ceregido^{1,2,3*}, Mercedes Spínola Amilibia^{4,5*}, Lieven Buts^{2,3}, José Rivera-Torres⁶, Abel García-Pino^{2,3}, Jerónimo Bravo⁴ and Nico A.J. van Nuland^{2,3}

¹Departamento de Química Física e Instituto de Biotecnología, Facultad de Ciencias, Universidad de Granada, Granada, Spain

²Jean Jeener NMR Centre, Structural Biology Brussels, Vrije Universiteit Brussel (VUB), Brussels, Belgium

³Molecular Recognition Unit, Department of Structural Biology, VIB, Brussels, Belgium

⁴Instituto de Biomedicina de Valencia (IBV-CSIC), Valencia, Spain

⁵Departamento de Biología Físico-Química, Centro de Investigaciones Biológicas (CIB-CSIC), Madrid, Spain

⁶Department of Epidemiology, Atherothrombosis and Imaging, Spanish National Cardiovascular Research Center (CNIC), 28029, Madrid, Spain

* Joint first authors.

To whom correspondence should be addressed: J. Bravo, Instituto de Biomedicina de Valencia (IBV-CSIC), C/ Jaime Roig 11, Valencia, Spain. Tel.: +34-963391760; Fax: +34-963690800; Email: jbravo@ibv.csic.es; N.A.J. van Nuland, Department of Structural Biology, VIB, Pleinlaan 2, 1050 Brussels, Belgium. Tel. +32-26293553; Fax: +32-26291962; Email: nvnuland@vub.ac.be

ABSTRACT

TAX1BP1 is a novel ubiquitin-binding adaptor protein involved in the negative regulation of the NF-kappaB transcription factor, which is a key player in inflammatory responses, immunity and tumorigenesis. TAX1BP1 recruits A20 to the ubiquitinated signalling proteins TRAF6 and RIP1, leading to their A20-mediated deubiquitination and the disruption of IL-1 and TNF-induced NF-kappaB signalling, respectively. The two zinc fingers localized at its C-terminal, function as novel ubiquitin binding domains (UBZ, ubiquitin binding zinc finger). Here we present for the first time both the solution and crystal structures of two classical UBZ domains in tandem within the human TAX1BP1. The relative orientation of the two domains is slightly different in the X-ray structure with respect to the NMR structure, indicating some degree of conformational flexibility, which is rationalized by NMR relaxation data. The observed degree of flexibility and stability between the two UBZ domains might have consequences on the recognition mechanism of interacting partners.

KEYWORDS: NMR, X-ray, SAXS, Ubiquitin-binding domains, Inflammation.

INTRODUCTION

Nuclear Factor kappa B (NF-kappaB) acts as a regulator of genes encoding molecules that control innate and adaptive immune responses^{1,2} inflammation, cell proliferation, apoptosis and oncogenesis¹. NF-kappaB is transiently activated by bacterial lipopolysaccharide (LPS) and cytokines such as tumor necrosis factor (TNF) and interleukin (IL)-1. The uncontrolled activation of NF-kappaB leads to pathologic

states such as chronic inflammation, septic shock and autoimmunity^{3,4,5,6,7}. The cell controls NF-kappaB activity through a negative feedback loop where the A20-ubiquitin editing complex, composed by A20, TAX1BP1 and the E3 ligases Itch and RNF11, plays an essential role^{8,9}.

A20 is an enzyme with dual-function, working as deubiquitinating and E3 ligase¹⁰. A20 is a major regulator of NF-kappaB signalling¹¹; indeed, A20-deficient mice develop severe inflammation and die prematurely¹² and the importance of A20 in limiting inflammation is evidenced by the wide number of human autoimmune diseases associated with polymorphisms in the A20 genomic region^{13,14,15}.

Tax1-binding protein1 (TAX1BP1; also known as TXBP151 or T6BP) functions as an ubiquitin-binding adaptor protein for the enzyme A20¹⁶ being pivotal in the termination of NF-kappaB signalling and the anti-apoptotic activity of A20^{8,16,17,18}. Indeed, it has been recently reported that the kinase IKK α phosphorylates TAX1BP1 at two serine residues, which initiates the assembly of the A20-ubiquitin editing complex¹⁹. Itch and RNF11 (ring finger protein 11) E3 ligases interact with the A20-TAX1BP1 complex in a stimulus-dependent manner and are also required for the termination of NF-kappaB signalling^{8,9}.

The human T cell leukemia virus type (HTLV-I) genome encodes an oncogenic protein, Tax that triggers persistent activation of NF-kappaB^{20,21}. Tax works as an antagonist of the A20 ubiquitin-editing complex, by interacting with TAX1BP1 and inhibiting the function of A20^{8,22,23}. It has recently been proposed that Tax can suppress the IKK α -mediated TAX1BP1 phosphorylation, acting as a negative

regulator of TAX1BP1¹⁹.

TAX1BP1 is localized in the cytoplasm and trans-locates to nucleus or plasma membrane in a stimulus dependent manner^{17,18,24,25}. The N-terminal part of TAX1BP1 (Fig.1) contains a SKIP carboxyl homology (SKICH) domain. SKICH domains have been described as membrane targeting domains, however, this property has not been reported for TAX1BP1²⁶. The central part contains three putative coiled-coil sequences and an overlapping homo-dimerization region^{27,28}. TAX1BP1 contains two C-terminal ubiquitin-binding zinc finger domains in tandem, UBZ1 and UBZ2 (Fig.1). Both contain highly conserved “PPXY” motifs targeting WW domain containing proteins²⁹, like Itch⁸. It is known that TAX1BP1 interacts with Itch, in a stimulus-dependent manner⁸, being Itch an essential subunit of the A20 ubiquitin-editing complex. Notably, removal of Itch impairs the function of A20⁸.

The UBZ domains of TAX1BP1 were shown to bind Lys63-ubiquitinated proteins such as TRAF6 and RIP1 leading to their A20-mediated deubiquitination and the disruption of IL-1 and TNF-induced NF-kappaB signalling, respectively^{17,18}. Thus, it is possible that TAX1BP1 mediates the recognition of substrates by the A20 ubiquitin-editing complex³⁰. Moreover, the UBZ domains are also involved in the binding of TAX1BP1 to the motor protein myosin VI³¹ implicated in many cellular processes as endocytosis, secretion, membrane ruffling and cell motility.

Structural characterization of the tandem of UBZ domains (UBZ1+2) in TAX1BP1 is thus important for a deep understanding of the interaction network between TAX1BP1 and its different targets. We thus determined the high-resolution structures of UBZ1+2

of the human TAX1BP1 in crystal and solution environment by X-Ray and NMR, respectively. The three-dimensional structures show all characteristics that are typically found in classical zinc finger domains, including the tetrahedral Cys₂His₂ coordination with the Zinc atoms. The crystal is densely packed, resulting in a slightly different relative orientation of the two UBZ domains compared to the structure in solution. We show by the use of Small Angle X-ray Scattering (SAXS) that both UBZ1 and UBZ2 remain structured when separated.

Up to date, there are no structures published of UBZ domains in tandem. Until now, the structure of only two different UBZ domains have been reported, one NMR structure of the UBZ domain of the human DNA Y-polymerase eta (DNA pol eta)³² and one X-ray structure of the UBZ domain of the human Werner helicase interacting protein 1 (WRNIP1)³³. Based on the sequence alignment of the UBZ domains of TAX1BP1, DNA pol eta and WRNIP1, we have observed phylogenetic differences that might point to the existence of distinct binding mechanisms within these families of UBZ domains to its common partner, ubiquitin.

RESULTS

Crystal structure of TAX1BP1 UBZ1+2

Figure 2 shows the arrangement of the 11 X-ray structures of UBZ1+2 present in the asymmetric unit. Each molecule is tightly packed between two neighboring molecules in an anti-parallel fashion resulting in two groups (ACEGIK and BDFHJ) with their N- and C-terminus respectively, pointing in the same direction. Figure 3 shows an

overlay of the 11 UBZ1+2 molecules highlighting the tetrahedral coordination of the zinc atoms with the Cys₂His₂ of UBZ1 and UBZ2. The zinc atoms are coordinated by Cys10, Cys13, His29 and His33 in the UBZ1 domain and by Cys37, Cys40, His56 and His60 in UBZ2. Both domains, UBZ1 as well as UBZ2, show the classical Cys₂His₂ zinc finger domain fold^{34,35}, consisting of two short β strands (residues 8-9 and 16-17 in UBZ1 but not observed in all structures, and 35-36 and 43-44 in UBZ2) followed by an α -helix (residues 23-31 in UBZ1 and 50-61 in UBZ2) forming a left-handed $\beta\beta\alpha$ -unit, in agreement with the 3D structures of the two other UBZ domains published till the moment^{32,33}. The tetrahedral coordination of each zinc atom stabilizes the 3D structure of both zinc finger domains by holding together the β -sheet with the α -helix.

Table 1 summarizes the crystallographic data collection and refinement statistics. The 11 structures are very similar with a backbone RMSD of 0.16 Å and show good Ramachandran statistics with only 0.46 % of the residues (3 residues located at the extreme of the chains) in energetically unfavorable areas. The structural coordinates have been deposited in the Protein Data Bank (PDB) with accession number 4BMJ.

Solution structure of TAX1BP1 UBZ1+2

The semi-automatic NMR assignment protocol resulted in a full backbone assignment and a ¹H side-chain assignment of 98%. Figure 4 shows the assigned ¹H-¹⁵N-HSQC spectrum of UBZ1+2 that resulted from the procedure. The ¹H, ¹⁵N and ¹³C assigned resonances have been deposited in the BioMagResBank (<http://www.bmrwisc.edu/>) under accession number 19201.

Figure 5 shows a cartoon representation of the ensemble of 20 lowest-energy water-refined structures derived from the automatic NOE assignment protocol in CYANA and the subsequent refinement using RECOORD. The ensemble fulfills the experimental data very well and shows excellent Ramachandran statistics (Table 2) reflecting the high quality of the 3D structures. UBZ1 as well as UBZ2 show the same Cys₂His₂ zinc finger domain fold as in the X-ray structure. The structural disorder observed for the N- and C-terminus is consistent with the higher flexibility in this region of the protein as indicated by the smaller heteronuclear NOE values (see below). The structural coordinates and experimentally derived restraints have been deposited in the PDB with accession number 2M7Q.

Backbone dynamics of TAX1BP1 UBZ1+2

Backbone ¹⁵N R₁, R₂ and hetNOE were measured for the 50 non-overlapping cross peaks of UBZ1+2 (Fig. 6). The hetNOE values (Fig. 6A) are almost all lower than the theoretical upper limit value of 0.84 at 600 MHz³⁶. The low (negative) NOE values of the residues at the N-terminus (UBZ1 domain) and at the C-terminus (UBZ2 domain), indicate that the terminal regions are much more flexible than the other regions of the protein. Elevated R₂ values (Fig. 6C) are observed especially for the residues located in the UBZ1 domain (residues 16-33), whereas the R₁ values are at about the same in the whole UBZ1+2 molecule (Fig. 6B). Elevated R₂ values are indicative of conformational exchange on the microsecond-millisecond time scale³⁷.

Analysis of the relaxation data of UBZ1+2 using TENSOR2³⁸ indicates an average ¹⁵N R₂/R₁ ratio (Fig. 6D) in the most ordered regions of 6.84, corresponding to an apparent rotational correlation time τ_c of 7.73 ns. This correlation time is higher than

the expected 5.15 ns for a globular protein of the same molecular weight (8.3 kDa) using the empirical formula based on a best fit of data from the NESG website, (www.nmr2.buffalo.edu/nescg/wiki/NMR_determined_Rotational_correlation_time), most likely due to the anisotropy of the system but also because the flexible termini of the tandem can significantly enhance the τ_c ³⁹. TENSOR2³⁸ analysis shows that UBZ1+2 presents a rotational anisotropy of $D_{\text{par}}/D_{\text{perp}}=1.37 \pm 0.1$. This means that the rotational diffusion tensor is significantly asymmetric and exhibits a prolate axial symmetry with $D_{\text{par}}=D_{\text{xx}}=D_{\text{yy}}=(1.93 \pm 0.09) \cdot 10^7 \text{ s}^{-1}$ and $D_{\text{perp}}=D_{\text{zz}}=(2.65 \pm 0.12) \cdot 10^7 \text{ s}^{-1}$, where the α -helix of the UBZ1 domain is parallel to the unique axis of the diffusion tensor. Having this into account, the higher R2 values measured for the residues in the α -helix of UBZ1 than those in UBZ2 (Fig. 6C) can be explained by its parallel alignment to the longitudinal axis of the prolate diffusion tensor³⁷.

Incorporation of this tensorial description into the model-free analysis^{40,41} of the local internal mobility affecting the backbone amides, has allowed us to describe the dynamics of the molecule. The measured relaxation data can be adequately described for most residues by a single parameter, the order parameter S^2 (Fig. 6E), which is related to the local mobility of the N-H vector on the subnanosecond timescale. The profile of the S^2 (Fig. 6E) confirms the increase in internal mobility at the N- and C-terminal part of the protein, as can be seen from a large decrease in S^2 . The residues located at the loop connecting the β -sheet and the α -helix of UBZ2 present a small decrease of the S^2 values in comparison with the values of the rest of secondary structure elements, which point to an increase in flexibility in this region of UBZ1+2. The local internal mobility of the residues showing conformational exchange (Fig. 6F) can only be described with the TENSOR2³⁸ model using an additional term for

exchange contributions (K_{ex}).

Characterization of TAX1BP1 UBZ1+2 and the individual domains by SAXS

The structural parameters of UBZ1+2 calculated from the experimental SAXS curves (Fig. 7A) are shown in Table 3. The estimated molecular weight agrees well with the one predicted from the sequence and observed by gel filtration (approximately 8.3 kDa). Both the crystallographic and NMR determined structures explain very well the SAXS experimental data with the best fit obtained using the models from NMR. *Ab initio* shape of UBZ1+2, reconstructed from the experimental SAXS data using the program DAMMIF⁴² reproduces very well the flat and elongated shape of the tandem and is consistent with the crystallographic and NMR structural data (Fig. 7B). The scattering curves from the individual domains are also very consistent with the CRY SOL⁴³ calculated theoretical curves, derived directly from experimentally determined structures (Fig. 7C and E, Table 3). In both cases the calculated curves can be fitted to the data with χ^2 values below 1.5. Figure 7D and F show the *ab initio* reconstructed shapes of both individual domains superimposed onto the atomic coordinates. As in the case of the tandem, the comparison shows a clear agreement with the atomic coordinates. These observations confirm that both domains are folded when bound to Zn^{2+} , in the conformation compatible for target binding.

DISCUSSION

NMR spectroscopy and X-ray crystallography are complementary techniques widely used for the determination of protein structures at atomic level^{44,45}. In general,

independent structure determination by NMR and X-ray is useful to validate protein structures, and they can also be combined to obtain a more accurate model^{46,47,48}.

The 3D structures of the tandem UBZ1+2 in the ubiquitin-adaptor protein TAX1BP1 obtained by the independent use of X-ray diffraction (Fig. 2 and 3) and NMR spectroscopy (Fig. 5) described in this paper, show that both UBZ domains present a compact structure comprised of two left-handed $\beta\beta\alpha$ -units, common in the family of classical zinc fingers, with two Cys₂His₂ motifs coordinating the Zn²⁺ atoms, crucial to maintain the 3D structure of the zinc finger domains. The packing between the β -sheet and the α -helix, forms a hydrophobic core (Fig. 8) that places the conserved Cys and His residues toward the interior of the domain in a position to coordinate the Zn²⁺ atom. Several residues (Leu12, Leu15, Phe17, Tyr21, Phe26, and Val30 for UBZ1; and Met39, Phe44, Tyr48, Phe53 and Val57 for the UBZ2 domain) serve to shield the zinc ions from solvent and stabilize the metal-interaction. It is worth to notice that the hydrophobic core is further stabilized by an interaction between the aromatic side chains of the Phe17 and Phe44 with His29 and His56 in the UBZ1 and UBZ2 domains, respectively. The formation of such aromatic anchor is very common in classical zinc fingers⁴⁹. In most zinc finger domains this aromatic residue is located three residues after the second Cys that coordinates the zinc ion (C-X₂-C-X₃-F-X₅- Φ -X₂-H). This position is called “consensus”, like for example in the classical zinc fingers from Friend Of GATA, FOG-F3⁵⁰. However, in some unusual cases the aromatic residue appears two residues after the second Cys (C-X₂-C-X₂-Y-), and is known as “aromatic swap”, like for the NEMO zinc finger⁴⁹. In both UBZ domains of TAX1BP1 the aromatic residues Phe17 and Phe44 are in the “consensus” position (Fig. 8).

Previous studies have shown that mutation of these Phe residues severely diminished the binding of the UBZ1+2 tandem to ubiquitin¹⁸. These phenylalanines act like anchors of the tetrahedral coordination of the Zn-atoms, which helps to stabilize the structure of the tandem (Fig. 8). In the same way, mutations of the conserved Cys residues of the UBZ domains of TAX1BP1 impair the ability of TAX1BP1 to mediate the inhibition of NF- κ B signaling⁸. Based on the structure of UBZ1+2 we can rationalize that effect due to the crucial role of these cysteines in the coordination of the Zn-atoms, necessary to maintain the 3D structure of both UBZ domains. In this sense, the lack of the aromatic anchor and/or of a Cys directly involved in Zn-coordination give rise to a loss of the correct native fold of the tandem, resulting in incompetent recognition of natural targets, which explains the defect in TAX1BP1 function^{8,18}.

Closer inspection of the ensemble of 11 X-ray structures (Fig. 3) shows diversity in the β -sheet region within the UBZ1 domain. Indeed, variations of the PHI angle in the X-ray ensemble in this region of the protein are substantial (Fig. 9A), and also deviate from those in the NMR ensemble. Interestingly, we observe lower intensity of several signals of UBZ1 residues in the ¹H-¹⁵N HSQC spectrum corresponding to the residues that show high variety in PHI angles in the X-ray structures (Fig. 9B). Such lowering of intensity is indicative of extensive exchange between different conformers in agreement with the conformational heterogeneity observed in the ensemble of X-ray and NMR structures. Moreover, conformational exchange within this region of the UBZ1 domain is confirmed by the elevated R2 values (Fig. 6C) and the increased K_{ex} values resulting from the model free analysis (Fig. 6F) for residues comprising the

region.

Both X-ray and NMR structures of TAX1BP1 UBZ1+2 are of high quality (Table 1 and Table 2, respectively) and the structures of the individual domains within the tandem are closely similar with backbone RMSD of 1.14 Å for UBZ1 (residues 8-34) and only 0.91 Å for UBZ2 (residues 35-62). However, superposition of the full UBZ1+2 (residues 8-62) results in a 2.30 Å backbone RMSD between the NMR and X-ray structures. The substantial increase in difference is related to the altered relative orientation of the two UBZ domains towards each other as shown in Figure 10. The hinge region involves the loop connecting the β -strand with the helix in UBZ2. Also the helix in UBZ2 is longer in the X-ray structure and its C-terminal end is bent likely as result of the tight packing. The quantitative analysis of the relaxation parameters confirms the flexibility of the loop region in the UBZ1+2 tandem where smaller hetNOE and S^2 values are found in comparison with the secondary elements in UBZ1+2 (Fig. 6). Moreover, the ^{15}N R2/R1 ratio provides a good measure of the global tumbling rate of the N-H vector, and the values of the calculated order parameters are pointing to flexibility of the C-terminal end of the helix in the UBZ2 domain, in agreement with the NMR structures. Despite the 11 independent molecules present in the crystallographic asymmetric unit, the flexibility between the two UBZ domains only became apparent when comparing the X-Ray and NMR structures indicating that both techniques are complementary to fully understand the molecular dynamic interactions involving these domains.

One of the main functions of TAX1BP1 is to act as a negative regulator of NF- κ B. Indeed, previous studies shows that TAX1BP1 knockout mice develop age-

dependent inflammatory cardiac valvulitis, dermatitis and hypersensitivity to TNF and IL-1¹⁸. Shembade and co-workers proposed a model in which the phosphorylation of TAX1BP1 induces the exposition of its UBZ domains and PPXY motifs allowing the inducible recruitment of Itch, RNF11 and A20, giving rise to the formation of the A20-ubiquitin editing complex and the recruitment of polyubiquitin substrates¹⁹. Indeed, it is known that TAX1BP1 requires the UBZ domains and associated PPXY motifs to terminate cytokine-mediated NF-kappaB signalling^{8,18}. The UBZ domains are involved in binding of polyubiquitinated RIP1 and TRAF6 while the PPXY motifs are needed for the binding to the E3 ligase Itch^{8,18}. However, till now the mechanisms involved in the interaction of TAX1BP1 with its different targets remains unknown.

The flexibility shown between the two UBZ domains might provide some degree of adaptation to accommodate the recognition of cellular partners in tandem such as polyubiquitins or tandem WW domains. Taking into account that other UBZ domains interact with ubiquitin through an α -helix³² and that the PPXY motifs are located in the loop connecting the β -strand with the α -helix in UBZ1 and UBZ2, it would be possible that the tandem of UBZ1+2 could interact with ubiquitin and the WW domains of Itch simultaneously. Thus, identification and characterization of the ubiquitin- and WW-binding interface on the UBZ domains of TAX1BP1 will be necessary to determine possible competition between WW domains and ubiquitin. Moreover, it would also be important to determine how many molecules of ubiquitin or which and how many of the four WW domains present in Itch can bind UBZ1+2 at the same time. Additionally it is of interest to know if the tandem would selectively bind different types of polyubiquitin chains besides the identified Lys63-polyubiquitin substrates^{17,18}. These results would be crucial to understand what are the mechanisms that control the assembly of the A20 ubiquitin-editing complex and the binding of

polyubiquitin substrates.

Small linear peptides containing the PPXY motif of other PPXY-containing proteins have been used for studying interactions with WW domains (PDB entries 2JO9, 2KQ0 and 1JMQ), probably because these motifs seem to appear in very flexible and/or unstructured regions. However, our UBZ1+2 structure shows that the PPXY motifs of UBZ1 and UBZ2 of TAX1BP1 are structured and flanked between other structural regions, diminishing the flexibility of the motifs. These structural and dynamical differences might have important consequences for the recognition of WW domains, especially in tandem, implying an additional level of specificity in the formation of complexes involving the UBZ1+2 tandem of TAX1BP1.

Till present, only two structures of UBZ domains can be found in the PDB, apart from the UBZ1+2 domains of the human TAX1BP1 that we report in this work. The first one corresponds to the UBZ domain of the human DNA pol eta³² (PDB entry 2I5O) a classical Cys₂His₂ UBZ domain, like the UBZ domains of TAX1BP1; the second one is a UBZ domain of the human WRNIP1 (Werner helicase-interacting protein 1) (PDB entry 3VHS)³³ that is involved in DNA replication⁵¹, and presents a variation of the classical Cys₂His₂ motif, characterized by a Cys₂HisCys motif. Other UBZ domains like the ones in ubiquitin ligase RAD18 and polymerase kappa are also Cys₂HisCys domains⁵¹. DNA pol eta, DNA pol kappa, RAD18 and WRNIP1 share a common biochemical function, besides that they are able to interact with ubiquitin, all of them bind DNA and participate in post-replication DNA repair⁵². This is one of the main differences with TAX1BP1, for which no DNA binding has been reported⁵³.

Given the degree of structural similarity between the three reported structures and knowing their functional differences, it is interesting to identify if these members of the UBZ family of proteins are evolutionary related. In order to answer this question, the program jPRED3⁵⁴ was used to search sequences similar to the ones of UBZ1 and UBZ2 domains of TAX1BP1, UBZ domain of pol eta and UBZ domain of WRNIP1, where all the alignments were filtered out for redundancy at 75% sequence identity⁵⁴. The output was used to perform a multiple sequence alignment with the program Geneious 6.1.4⁵⁵. Figure 11 shows the alignment of all the sequences corresponding to different organisms.

As we can see, UBZ1 and UBZ2 domains of TAX1BP1 are four residues longer than the other two types of UBZ domains (DNA pol eta and WRNIP1) in the region corresponding to the loop that links the second β -sheet with the α -helix (Fig.11). We have shown by comparison of the X-ray and NMR structures that this loop in the UBZ2 domain acts like a hinge in the tandem UBZ1+2, determining the relative orientation of one domain with respect to the other, likely contributing to enhance the tandem elasticity necessary for target recognition. In this sense, the existence of longer loops in UBZ domains of TAX1BP1 could be related not only with the fact that they are forming a tandem while DNA pol eta and WRNIP1 only contain a single UBZ domain, but also with their function, since the UBZ domains of TAX1BP1 are able to interact with the WW domains of Itch through the PPXY motifs that are located in these loops. In general Cys and His residues are very well conserved in the three different proteins among the different species, which highlight their essential role in maintaining the 3D structure of the domains by Zn^{2+} coordination. It is worth to notice, that the Ala656 (at the C-terminal part of the α -helix of DNA pol eta) crucial

for the binding of UBZ domain of DNA pol eta with ubiquitin³², is always conserved across the organisms within the family DNA pol eta but not for the rest of UBZ domains. For TAX1BP1, the UBZ2 domain contains an Asp in the equivalent position, whereas the shorter α -helix present in the UBZ1 does not even contain the consensus sequence centered around the Ala^{32,49}. The UBZ domain of WRNIP1 presents an α -helix whose length is comparable to TAX1BP1 UBZ1 and, similarly, does not contain this consensus region. However, it contains a conserved Asp residue at position 37 in the middle of the helix that is crucial for binding to ubiquitin, as a single point mutation of Asp37 into an Ala completely abolished ubiquitin-binding⁵². Despite that the conserved Ala of DNA pol eta is not present in any of the other UBZ domains, they are still able to bind ubiquitin, showing that the consensus sequence proposed by the sequence alignment of DNA pol eta and other zinc finger domains like NEMO and IUM^{32,49} is not a general feature to explain the binding of ubiquitin to UBD domains containing an α -helix.

Other conserved residues are the prolines positioned after the first conserved Cys and in the fifth position after the second conserved Cys; this last Pro corresponds to the first Pro of the PPXY sequence exclusively conserved in the UBZ1 and UBZ2 of TAX1BP1 but not in the UBZ domains of the DNA pol eta or WRNIP1. This Pro is always located at the beginning of the loop that connects the β -sheet with the α -helix, and its conservation is another characteristic structural feature of this family of domains. In UBZ1+2 domains of TAX1BP1 and the UBZ domain of DNA pol eta, a Val or a Met after the first consensus His is also conserved, and an aromatic residue after the second conserved His (a Trp in the UBZ1 domains and a Phe in the UBZ2 of TAX1BP1 and the UBZ of DNA pol eta) which is missing in UBZ of WRNIP1 that

contains a conserved Leu in that position. These aromatic residues in UBZ domains of TAX1BP1 and DNA pol eta are not oriented to the hydrophobic core of the zinc fingers but on the exposed surface of the α -helix. For the UBZ domain of DNA pol eta, this Phe has been reported to be involved in direct binding to ubiquitin³².

Based on the sequence alignment and using the program Geneious 6.1.4⁵⁵ we can create an unrooted phylogenetic tree (see Fig.12). This diagram represents the relationship among organisms, where the length of the branches is based on genetic distances, which correspond with the number of mutations between species since their divergence. As we can see in Figure 12, despite their structural and sequence similarity, UBZ domains of TAX1BP1, DNA pol eta and WRNIP1 cluster in three different branches.

The divergent pattern followed by the UBZ domains of these three proteins can be related with their distinct functions within the cell: DNA pol eta is implicated in DNA repair by translesion synthesis, WRNIP1 is implicated in DNA replication and repair, and TAX1BP1 is implicated in the regulation of transcription factor as NF-kappaB and JNK, not directly involved in DNA binding⁵³. The phylogenetic differences based on sequence alignment might point to the existence of distinct binding mechanisms within these UBZ domains to its common partner, ubiquitin, opening the door for new studies oriented to deeply characterize this versatile and quite unknown family of domains.

In conclusion, the work described in this paper is presenting for the first time a structure of a tandem of two UBZ domains. This structural work is the basis for future

studies to understand how these zinc finger domains recognize different targets such as ubiquitin and the WW domain of Itch, providing insight in the signaling pathways in which TAX1BP1 is involved. This may ultimately lead to the design of specific drugs that might be used in treatments of inflammatory diseases.

MATERIALS AND METHODS

Protein expression and purification

UBZ1+2 encompassing residues 725 to 789 from human TAX1BP1 was PCR amplified using specific primers (5'-GCG **CATATG** GAT GTT CAC AAG AAG TGT CCC-3' (Forward) and 5'-GC **CTCGAG** CTA GTC AAA ATT TAG AAC ATT C-3' (Reverse) and a previously reported construct as template¹⁸. Amplicon was cleaved (*NdeI/XhoI*, shown in bold in the primer sequence) and inserted into a modified version of pET28a containing a specific PreScission Protease cleavage site to facilitate tag removal. Construct was sequence verified. The plasmid pET28a containing the human UBZ1+2₇₂₅₋₇₈₉ gene covalently linked to a N-terminal 6xHis tag was overexpressed in *Escherichia coli* Rosetta (DE3)/pLysS strain (Novagen). Cells were grown in 2xTY media and induced by 1mM IPTG. Cells were harvested after 4 hours at 37 °C and resuspended in lysis buffer (50 mM TRIS-HCl pH 7.5, 0.5 M NaCl, 0.5 % Tween-20 and 0.2 mM AEBSF (4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride; Sigma) as protease inhibitor. After cell lysis by sonication and removal of cell debris by centrifugation (30 minutes at 20000xg). Supernatant was loaded on a 5 mL Nickel Chelating column, washed with 50 mM TRIS-HCl pH 7.5, 0.2 M NaCl and 10 mM Imidazole and eluted with 13 column volumes on a linear

gradient of imidazole from 0 to 500 mM. Samples were collected and digested with PreScission (GE Healthacare) protease (at 1:1000 (w/w) protease:protein concentration) in 50 mM Tris pH 7.5 and 0.2 M NaCl, at room temperature during 90 minutes. The sample was concentrated and loaded on a Superdex 75 16/60 equilibrated with 20 mM Tris pH 7.5 and 0.2 M NaCl. Fractions containing TAX1BP1 UBZ1+2 were pooled, concentrated and flash frozen in liquid nitrogen until use.

To produce ^{15}N - and ^{13}C , ^{15}N -labeled TAX1BP1-UBZ1+2₇₂₅₋₇₈₉ His-tag₆, the plasmid pET28a was overexpressed in *E. coli* BL21 (DE3) strain (Novagen). Cells were grown in 2xTY medium four times the volume of the M9 minimal medium at 37 °C till an OD₆₀₀ of 0.7. After two washing steps, cell-pellets were resuspended and incubated for 1h in M9 medium at 22 °C. Protein expression was induced with 0.2 mM IPTG at 20 °C during 24h. Cells were resuspended after centrifugation at 4600xg for 20 min in 50 mM TRIS-HCl pH7.5, 0.2M NaCl, 10 mM Imidazole, 50 μM ZnCl₂, 1 mM TCEP and 0.5 % Tween-20 (v/v), and broken in a French press. After ultracentrifugation at 48000xg during 30 min, the supernatant was treated following the same protocol as for unlabeled UBZ1+2. ^{15}N and ^{13}C , ^{15}N -labelled proteins were purified from cells grown in M9 minimal medium containing $^{15}\text{NH}_4\text{Cl}$ and ($^{13}\text{C}^6$)D-glucose (Spectra Stable Isotopes and Cortecnet) as the sole sources of nitrogen and carbon, respectively. Sample concentrations were determined spectrophotometrically using an extinction coefficient of 8480 M⁻¹cm⁻¹ determined using the ProtParam tool (Expasy).

Isolated UBZ1 and UBZ2 domains were purchased from ChinaPeptides (Shanghai, China). Both of them were N-terminal acetylated and C-terminal amidated. Before the experiments, the proper amount of lyophilized UBZ1 or UBZ2 was weighted on an

analytical balance (precision: ± 0.1 mg) and dissolved in the appropriate buffer. In order to ensure the proper folding of the proteins, 1 molar equivalent of TCEP and 1.3 molar equivalents of ZnCl_2 were added to the sample to provide a reducing environment and induce the folding of the domains, respectively. The proper folding of the isolated UBZ1 and UBZ2 domains was checked by circular dichroism (CD) spectroscopy. When necessary, UBZ1 and UBZ2 domains were dialyzed using mini dialysis membranes with a 500 Da cut-off (GE Healthcare).

Crystallization, Crystallographic Data Collection, Structure Determination, and Refinement of UBZ1+2

The protein was concentrated at 60 mg/mL and crystallized by the hanging drop method in 1 M LiSO_4 and 0.1 M Tris pH 8.5 at 22 °C. Crystals were cryo-cooled in Paratone-N for data collection at the ESRF (Grenoble), beamline ID14-1. Data reduction was performed with MOSFLM and SCALA as implemented in CCP4⁵⁶ up to a resolution of 2.59 Å.

The structure was solved by Single-wavelength anomalous diffraction (SAD). Experimental phases were obtained using SHARP/AutoSHARP⁵⁷. Initial phases using 22 sites (over 78% occupancy) for Zn^{2+} were used to calculate maps and to build initial models using SOLVE/RESOLVE as implemented in PHENIX^{58,59,60}. Eleven molecules were built in the asymmetric unit. Restrained NCS and TLS were used for refinement in Refmac5⁵⁶. Structures were validated using the tool *Validate* of the program *Cool*^{61,62} and Molprobit⁶³. Table 1 summarizes the crystallographic data collection and refinement analysis of UBZ1+2.

Nuclear Magnetic Resonance (NMR) spectroscopy

NMR experiments were run on a 0.7 mM ^{13}C , ^{15}N -labelled UBZ1+2 sample in 50 mM sodium phosphate buffer at pH 6.0, 1 mM TCEP, 9 % D_2O . All NMR spectra were acquired at 25 °C on a Varian NMR Direct-Drive Systems 600 or 800 MHz spectrometer, the latter one equipped with a salt-tolerance cold probe.

2D heteronuclear single quantum correlation (^1H , ^{15}N -HSQC and ^1H , ^{13}C -HSQC), 3D ^{15}N - and ^{13}C -NOESY-HSQC (mixing times of 125 ms) and triple resonance spectra CBCA(CO)NH, HNCACB, HNCO, HBHA(CO)NH and CCH-TOCSY⁶⁴ and aromatic (HB)CB(CGCD)HD and (HB)CB(CGCDCE)HE⁶⁵ were recorded on the double labelled sample.

All NMR data were processed using NMRPipe⁶⁶ and analyzed using CcpNmr⁶⁷. Semi-automatic assignment of N, NH, $\text{H}\alpha$, $\text{H}\beta$, CO, $\text{C}\alpha$, and $\text{C}\beta$ were obtained from the identification of intra- and inter-residue connectivities in HNCACB, CBCA(CO)NH, HBHA(CO)HN and HNCO spectra at the ^1H , ^{15}N frequencies of every peak in the ^1H , ^{15}N -HSQC spectrum. Assignments were extended to the side chain signals using correlations within CCH-TOCSY, (HB)CB(CGCD)HD and (HB)CB(CGCDCE)HE spectra. Remaining aromatic ^1H and ^{13}C frequencies were assigned from the ^1H , ^{13}C -HSQC and ^{13}C -NOESY-HSQC spectra. Side-chain NH_2 frequencies of glutamines and asparagines were assigned from HNCACB, CBCA(CO)NH and 3D ^{15}N -NOESY-HSQC spectra. All ^1H , ^{13}C and ^{15}N resonances were verified from 3D ^{15}N - and ^{13}C -NOESY-HSQC spectra.

Determination of the solution structure of UBZ1+2 by NMR

NOE cross peaks were obtained by manual peak picking in each strip of the 800 MHz 3D ^{15}N - and ^{13}C -NOESY-HSQC spectra taken at the assigned ^1H , ^{15}N and ^1H , ^{13}C frequencies, respectively, and by automatic peak picking of the 800 MHz 2D NOESY in H_2O followed by removal of diagonal peaks and peaks arising from artifacts (e.g. residual water). NOEs were assigned using the automated NOE assignment procedure of CYANA version 2.1^{68,69,70}.

Phi and psi torsion angle restraints were included based on analysis of ^1HN , ^{15}N , ^1Ha , ^{13}Ca , ^{13}CO and ^{13}Cb chemical shifts using the program Talos+⁷¹.

The standard protocol was used with seven cycles of combined automated NOE assignment and structure calculation of 100 conformers in each cycle. Restraints that were unambiguously assigned were used for a final structure calculation run using CNS⁷² in explicit solvent using the RECOORD protocol⁷³. The twenty lowest energy structures were used for the final analysis. Input data and structure calculation statistics are summarized in Table 2. PROCHECK-NMR⁷⁴ and WHATIF⁷⁵ were used to analyze the quality of the structures. MOLMOL⁷⁶ and PYMOL (<http://www.pymol.org/>) were used for further analysis and visualization.

Backbone dynamics from ^{15}N relaxation data

We have measured the relaxation parameters ^{15}N R1, R2 and ^1H - ^{15}N steady-state NOE of UBZ1+2 at 600 MHz and 25 °C. Relaxation values were obtained from series of 2D experiments with coherence selection achieved by pulse field gradients at 600 MHz using the experiments described previously⁷⁷ on the ^{15}N , ^{13}C -labelled UBZ1+2 sample. The ^1H - ^{15}N heteronuclear NOEs were determined from the ratio of peak intensities ($I_{\text{on}}/I_{\text{off}}$) with and without the saturation of the amide protons for 3 s. ^{15}N

R1 and ^{15}N R2 relaxation rates were measured from spectra with different relaxation delays: 100, 200, 300, 400, 500, 700, 900, 1200 and 1500 ms for R1 and 10, 30, 50, 70, 90, 110, 150, 190 and 230 ms for R2. Relaxation parameters and their corresponding errors were extracted with the program NMRview⁷⁸. For data analysis and interpretation, the residues that show spectral overlap were not considered (Asp5, Val6, His7, Lys8, Lys9, Cys13, Leu15, Arg55, His56, Thr59 and Val65).

TENSOR2³⁸ was used to analyze the experimental relaxation data using the model free approach^{40,41}.

Small Angle X-ray Scattering (SAXS)

SAXS data were collected at the SWING, beamline, Synchrotron Soleil, France, using the standard beamline setup in SEC (size exclusion) mode. Samples of the UBZ1+2 tandem and the isolated domains UBZ1 and UBZ2 were prepared at 1 mM concentration in 50 mM HEPES, 100 mM NaCl, 1 mM TCEP, pH 7.4. The data were analyzed with the ATSAS suite⁷⁹. For the SEC we used a Shodex KW402.5-4F column coupled to an HPLC system, in front of the SAXS data collection capillary, to separate the excess non-complexed material and thus remove this source of background. The samples was passed at a flow rate of 0.2 mL/min and the data collected at 10 °C. R_g values were obtained from the *Guinier* approximation and the I_0 by extrapolation to $q = 0$, as implemented in the ATSAS suite⁷⁹. All the *ab initio* envelopes were reconstructed from the experimental SAXS data using the program DAMMIF⁴² from the ATSAS suite⁷⁹.

ACCESSION NUMBERS

The X-ray structural coordinates have been deposited in the Protein Data Bank (PDB) with accession number 4BMJ.

The ^1H , ^{15}N and ^{13}C assigned resonances have been deposited in the BioMagResBank (<http://www.bmrb.wisc.edu/>) under accession number 19201.

The NMR structural coordinates and experimentally derived restraints have been deposited in the PDB with accession number 2M7Q.

ACKNOWLEDGEMENTS

We thank Yann Sterckx for his contribution in sequence alignment and development of the phylogenetic tree, and Alex Volkov for critical reading of the manuscript and productive scientific discussion. We thank Ivan Dikic for providing the TAX1BP1 plasmid. M.A.C. is supported by an FPU grant from the Ministerio de Educación y Cultura de España (MEC). L.B. and A.G.-P. are FWO post-doc fellows. Work in NvN laboratory is supported by the VIB and the Flanders Hercules Foundation. Work in JB laboratory was supported by Ministerio de Ciencia e Innovación (SAF2009-10667, SAF2012-31405) and Generalitat Valenciana ACOMP/2012/024. We thank support for data collection of ID14-1 staff at the ESRF.

REFERENCES

1. Hayden MS, Ghosh S. (2004) Signaling to NF-kappaB. *Genes Dev.* 15;18(18):2195-224.

2. Vallabhapurapu S, Karin M. (2009) Regulation and function of NF-kappaB transcription factors in the immune system. *Annu Rev Immunol.* 27:693-733.
3. Beg AA, Sha WC, Bronson RT, Baltimore D. (1995) Constitutive NF-kappa B activation, enhanced granulopoiesis, and neonatal lethality in I kappa B alpha-deficient mice. *Genes Dev.* 9(22):2736-46.
4. Schulze-Osthoff K, Ferrari D, Riehemann K, Wesselborg S. (1997) Regulation of NF-kappa B activation by MAP kinase cascades. *Immunobiology.* 198(1-3):35-49.
5. Karin M, Ben-Neriah Y. (2000) Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. *Annu Rev Immunol.* 18:621-63.
6. Greten FR, Eckmann L, Greten TF, Park JM, Li ZW, Egan LJ, Kagnoff MF, Karin M. (2004) IKKbeta links inflammation and tumorigenesis in a mouse model of colitis-associated cancer. *Cell.* 118(3):285-96.
7. Karin M. (2006) Nuclear factor-kappaB in cancer development and progression. *Nature.* 441(7092):431-6.
8. Shembade N, Harhaj NS, Parvatiyar K, Copeland NG, Jenkins NA, Matesic LE, Harhaj EW. (2008) The E3 ligase Itch negatively regulates inflammatory signaling pathways by controlling the function of the ubiquitin-editing enzyme A20. *Nat Immunol.* 9(3):254-62.

9. Shembade N, Parvatiyar K, Harhaj NS, Harhaj EW. (2009) The ubiquitin-editing enzyme A20 requires RNF11 to downregulate NF-kappaB signalling. *EMBO J.* 28(5):513-22.
10. Wertz IE, O'Rourke KM, Zhou H, Eby M, Aravind L, Seshagiri S, Wu P, Wiesmann C, Baker R, Boone DL, Ma A, Koonin EV, Dixit VM. (2004) De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-kappaB signalling. *Nature.* 430(7000):694-9.
11. Verstrepen L, Verhelst K, van Loo G, Carpentier I, Ley SC, Beyaert R. (2010) Expression, biological activities and mechanisms of action of A20 (TNFAIP3). *Biochem Pharmacol.* 80(12):2009-20.
12. Lee EG, Boone DL, Chai S, Libby SL, Chien M, Lodolce JP, Ma A. (2000) Failure to regulate TNF-induced NF-kappaB and cell death responses in A20-deficient mice. *Science.* 289(5488):2350-4.
13. Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W, Shifrin N, Petri MA, Kamboh MI, Manzi S, Seldin MF, Gregersen PK, Behrens TW, Ma A, Kwok PY, Criswell LA. (2008) Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. *Nat Genet.* 40(9):1062-4.
14. Dieguez-Gonzalez R, Calaza M, Perez-Pampin E, Balsa A, Blanco FJ, Cañete

- JD, Caliz R, Carreño L, de la Serna AR, Fernandez-Gutierrez B, Ortiz AM, Herrero-Beaumont G, Pablos JL, Narvaez J, Navarro F, Marengo JL, Gomez-Reino JJ, Gonzalez A. (2009) Analysis of TNFAIP3, a feedback inhibitor of nuclear factor-kappaB and the neighbor intergenic 6q23 region in rheumatoid arthritis susceptibility. *Arthritis Res Ther.* 11(2):R42.
15. Fung EY, Smyth DJ, Howson JM, Cooper JD, Walker NM, Stevens H, Wicker LS, Todd JA. (2009) Analysis of 17 autoimmune disease-associated variants in type 1 diabetes identifies 6q23/TNFAIP3 as a susceptibility locus. *Genes Immun.* 10(2):188-91.
16. De Valck D, Jin DY, Heyninck K, Van de Craen M, Contreras R, Fiers W, Jeang KT, Beyaert R. (1999) The zinc finger protein A20 interacts with a novel anti-apoptotic protein which is cleaved by specific caspases. *Oncogene.* 18(29):4182-90.
17. Shembade N, Harhaj NS, Liebl DJ, Harhaj EW. (2007) Essential role for TAX1BP1 in the termination of TNF-alpha-, IL-1- and LPS-mediated NF-kappaB and JNK signaling. *EMBO J.* 26(17):3910-22.
18. Iha H, Peloponese JM, Verstrepen L, Zapart G, Ikeda F, Smith CD, Starost MF, Yedavalli V, Heyninck K, Dikic I, Beyaert R, Jeang KT. (2008) Inflammatory cardiac valvulitis in TAX1BP1-deficient mice through selective NF-kappaB activation. *EMBO J.* 27(4):629-41.

19. Shembade N, Pujari R, Harhaj NS, Abbott DW, Harhaj EW. (2011) The kinase IKK α inhibits activation of the transcription factor NF- κ B by phosphorylating the regulatory molecule TAX1BP1. *Nat Immunol.* 12(9):834-43.
20. Ballard DW, Böhnlein E, Lowenthal JW, Wano Y, Franza BR, Greene WC. (1988) HTLV-I tax induces cellular proteins that activate the kappa B element in the IL-2 receptor alpha gene. *Science.* 241(4873):1652-5.
21. Crenon I, Béraud C, Simard P, Montagne J, Veschambre P, Jalinot P. (1993) The transcriptionally active factors mediating the effect of the HTLV-I Tax transactivator on the IL-2R alpha kappa B enhancer include the product of the c-rel proto-oncogene. *Oncogene.* 8(4):867-75.
22. Shembade N, Ma A, Harhaj EW. (2010) Inhibition of NF-kappaB signaling by A20 through disruption of ubiquitin enzyme complexes. *Science.* 327(5969):1135-9.
23. Chin KT, Chun AC, Ching YP, Jeang KT, Jin DY. (2007) Human T-cell leukemia virus oncoprotein tax represses nuclear receptor-dependent transcription by targeting coactivator TAX1BP1. *Cancer Res.* 67(3):1072-81.
24. Ulrich M, Seeber S, Becker CM, Enz R. (2007) Tax1-binding protein 1 is expressed in the retina and interacts with the GABA(C) receptor rho1 subunit. *Biochem J.* 401(2):429-36.

25. Newman AC, Scholefield CL, Kemp AJ, Newman M, McIver EG, Kamal A, Wilkinson S. (2012) TBK1 kinase addiction in lung cancer cells is mediated via autophagy of Tax1bp1/Ndp52 and non-canonical NF- κ B signalling. *PLoS One*. 7(11):e50672.
26. Gurung R, Tan A, Ooms LM, McGrath MJ, Huysmans RD, Munday AD, Prescott M, Whisstock JC, Mitchell CA. (2003) Identification of a novel domain in two mammalian inositol-polyphosphate 5-phosphatases that mediates membrane ruffle localization. The inositol 5-phosphatase skip localizes to the endoplasmic reticulum and translocates to membrane ruffles following epidermal growth factor stimulation. *J Biol Chem*. 278(13):11376-85.
27. Ling L, Goeddel DV. T6BP, a TRAF6-interacting protein involved in IL-1 signaling. *Proc Natl Acad Sci U S A*. 2000 Aug 15;97(17):9567-72.
28. Verstrepen L, Verhelst K, Carpentier I, Beyaert R. (2011) TAX1BP1, a ubiquitin-binding adaptor protein in innate immunity and beyond. *Trends Biochem Sci*. 36(7):347-54.
29. Sudol M, Chen HI, Bougeret C, Einbond A, Bork P. (1995) Characterization of a novel protein-binding module--the WW domain. *FEBS Lett*. 369(1):67-71.
30. Pelzer C, Thome M. (2011) IKK α takes control of canonical NF- κ B activation. *Nat Immunol*. 12(9):815-6.

31. Morriswood B, Ryzhakov G, Puri C, Arden SD, Roberts R, Dendrou C, Kendrick-Jones J, Buss F. (2007) T6BP and NDP52 are myosin VI binding partners with potential roles in cytokine signalling and cell adhesion. *J Cell Sci.* 120(Pt 15):2574-85.
32. Bomar MG, Pai MT, Tzeng SR, Li SS, Zhou P. (2007) Structure of the ubiquitin-binding zinc finger domain of human DNA Y-polymerase eta. *EMBO Rep.* 8(3):247-51.
33. Suzuki N, Wakatsuki S, Kawasaki M. (2010) Crystal structure of GFP-Wrip1 UBZ domain fusion protein in complex with ubiquitin. PDB code 3VHT.
34. Berg, JM. (1988) Proposed structure for the zinc-binding domains from transcription factor TFIIIA and related proteins. *Proc. Natl. Acad. Sci.* 85 99–102.
35. Gibson TJ, Postma JP, Brown RS, Argos P. (1988) A model for the tertiary structure of the 28 residue DNA-binding motif ('zinc finger') common to many eukaryotic transcriptional regulatory proteins. *Protein Eng.* 2(3):209-18.
36. Kay LE, Torchia DA, Bax A. (1989) Backbone dynamics of proteins as studied by ¹⁵N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry.* 28(23):8972-9.

37. Blake-Hall J, Walker O, Fushman D. (2004) Characterization of the overall rotational diffusion of a protein from ¹⁵N relaxation measurements and hydrodynamic calculations. *Methods Mol Biol.* 278:139-60.
38. Dosset P, Hus JC, Blackledge M, Marion D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *Journal of biomolecular NMR* 16 (1):23-28.
39. Tjandra N, Feller SE, Pastor RW, Bax A. (1995) Rotational diffusion anisotropy of human ubiquitin from ¹⁵N relaxation. *J. Am. Chem. Soc.* 117, 12562–12566.
40. Lipari, G. and Szabo, A. (1982a). Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules I. Theory and range of validity. *J. Am. Chem. Soc.*, 104(17), 4546-4559.
41. Lipari, G. and Szabo, A. (1982b). Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules II. Analysis of experimental results. *J. Am. Chem. Soc.*, 104(17), 4559-4570.
42. Svergun DI. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing, *Biophysical journal.* 23, 2707-12.
43. Svergun DI, Barberato C, Koch MHJ. (1995) CRY SOL - a program to

evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Cryst.* 28, 768–773.

44. Synder DA, Chen Y, Denissova NG, Acton T, Aramini JM, Ciano M, Karlin R, Liu J, Manor P, Rajan PA, Rossi P, Swapna GV, Xiao R, Rost B, Hunt J, Montelione GT. (2005) Comparisons of NMR spectral quality and success in crystallization demonstrate that NMR and X-ray crystallography are complementary methods for small protein structure determination. *J Am Chem Soc.* 127(47):16505-11.
45. Yee AA, Savchenko A, Ignachenko A, Lukin J, Xu X, Skarina T, Evdokimova E, Liu CS, Semesi A, Guido V, Edwards AM, Arrowsmith CH. (2005) NMR and X-ray crystallography, complementary tools in structural proteomics of small proteins. *J Am Chem Soc.* 127(47):16512-7.
46. Shaanan B, Gronenborn AM, Cohen GH, Gilliland GL, Veerapandian B, Davies DR, Clore GM. (1992) Combining experimental information from crystal and solution studies: joint X-ray and NMR refinement. *Science.* 257(5072):961-4.
47. Schiffer CA, Huber R, Wüthrich K, van Gunsteren WF. (1994) Simultaneous refinement of the structure of BPTI against NMR data measured in solution and X-ray diffraction data measured in single crystals. *J Mol Biol.* 241(4):588-99.

48. Hoffman DW, Cameron CS, Davies C, White SW, Ramakrishnan V. (1996) Ribosomal protein L9: a structure determination by the combined use of X-ray crystallography and NMR spectroscopy. *J Mol Biol.* 264(5):1058-71.
49. Cordier F, Vinolo E, Véron M, Delepierre M, Agou F. (2008) Solution structure of NEMO zinc finger and impact of an anhidrotic ectodermal dysplasia with immunodeficiency-related point mutation. *J Mol Biol.* 11;377(5):1419-32.
50. Simpson RJ, Yi Lee SH, Bartle N, Sum EY, Visvader JE, Matthews JM, Mackay JP, Crossley M. (2004) A classic zinc finger from friend of GATA mediates an interaction with the coiled-coil of transforming acidic coiled-coil 3. *J Biol Chem.* 279(38):39789-97.
51. Crosetto N, Bienko M, Hibbert RG, Perica T, Ambrogio C, Kensche T, Hofmann K, Sixma TK, Dikic I. (2008) Human Wrnip1 is localized in replication factories in a ubiquitin-binding zinc finger-dependent manner. *J Biol Chem.* 283(50):35173-85.
52. Bish RA, Myers MP. (2007) Werner helicase-interacting protein 1 binds polyubiquitin via its zinc finger domain. *J Biol Chem.* 282(32):23184-93.
53. Hofmann K. (2009) Ubiquitin-binding domains and their role in the DNA damage response. *DNA Repair (Amst).* 8(4):544-56.

54. Cole C, Barber JD, Barton GJ. (2008) The Jpred 3 secondary structure prediction server. *Nucleic Acids Res.* 36(Web Server issue):W197-201.
55. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 28(12):1647-9.
56. CCP4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50: 760-763.
57. Vonrhein C, Blanc E, Roversi P, Bricogne G. (2007) Automated structure solution with autoSHARP. *Methods Mol Biol.* 364:215-30.
58. Terwilliger, T.C. and Berendzen J. (1999) Automated MAD and MIR structure solution. *Acta Crystallographica D* 55, 849-861.
59. Terwilliger, T. C. (2000) Maximum likelihood density modification. *Acta Cryst.* D56, 965-972.
60. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. (2010) PHENIX: a comprehensive Python-based system for

macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* 66(Pt 2):213-21

61. Emsley P, Lohkamp, B, Scott WG and Cowtan, K. (2010) Features and Development of Coot. *Acta Crystallographica D66*, 486-501.
62. Emsley P, Cowtan K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60: 2126-2132.
63. Chen et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica D66*:12-21.
64. Sattler M, Schleucher J, Griesinger C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog Nucl Magn Reson Spectrosc* 34:93–158.
65. Yamazaki T, Forman-Kay JD, Kay LE. (1993) Two-dimensional NMR experiments for correlating ¹³C β and ¹H δ /e chemical shifts of aromatic residues in ¹³C-labeled proteins via scalar couplings. *J Am Chem Soc* 115:11054–11055.
66. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293.

67. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59:687–696.
68. Guntert P, Mumenthaler C, Wuthrich K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol* 273 (1):283-298.
69. Guntert P. (1997) Calculating protein structures from NMR data. *Methods Mol Biol* 60:157-194.
70. Herrmann T, Guntert P, Wuthrich K. (2002) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *Journal of biomolecular NMR* 24 (3):171-189.
71. Chen Y, Delaglio F, Cornilescu G, Bax A. (2009) TALOS plus: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44:213–223
72. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr.* 54(Pt 5):905-21.

73. Nederveen AJ, Doreleijers JF, Vranken W, Miller Z, Spronk CA, Nabuurs SB, Guntert P, Livny M, Markley JL, Nilges M, Ulrich EL, Kaptein R, Bonvin AM. (2005) RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins*. 59(4):662-672.
74. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8: 477-486.
75. Vriend G. (1990) WHAT IF: a molecular modeling and drug design program. *J Mol Graph*. 8(1):52-6, 29.
76. Koradi R, Billeter M, Wuthrich K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph*. 14(1): 51-5, 29-32.
77. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ¹⁵N NMR relaxation. *Biochemistry*. 33 (19):5984-6003.
78. Johnson, BA. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol*. 278, 313-352.

79. Konarev PV, Petoukhov MV, Volkov VV, Svergun DI. (2006) ATSAS 2.1, a program package for small-angle scattering data analysis. *J Appl Crystallogr* 39:277–286.

TABLES

Table 1. Crystallographic data collection and refinement analysis of UBZ1+2. Numbers in parenthesis indicate the highest resolution shell statistics.

Table 2. Structural statistics over the twenty lowest-energy water-refined NMR structures of UBZ1+2.

Table 3. Experimental and model based parameters obtained from the analysis of the small angle X-ray scattering data of the UBZ1+2 tandem and the individual domains.

FIGURES

Figure 1. Schematic domain structure of TAX1BP1. The numbers indicate the amino acids that delimit the structural domains. The N-terminus contains a SKICH domain, SKIP (skeletal muscle and kidney enriched inositol phosphatase) carboxyl homology domain, the central part contains three coil-coil (CC1, CC2 and CC3) structures and a region (O) which is responsible for homodimerization, and the C-terminal extreme contains the two zinc fingers domains UBZ1 (red) and UBZ2

(green). The sequence of the UBZ1+2 tandem is indicated below, including the four residues (GPHM) that remain after removal of the His-tag. Numbers in blue indicate the residue numbering used in this work, and numbers in black correspond to the numbering of the equivalent residues within the complete TAX1BP1 protein. The conserved PPXY motif, and the cysteine and histidine residues that coordinate the zinc atoms are highlighted in bold.

Figure 2. Cartoon representation of the tight packing of the 11 molecules (A to K) of UBZ1+2 in the asymmetric unit. Two different orientations are used. **A.** Using the C-axis of the unit cell to show the parallel disposition of every molecule, and **B.** Using the A-axis of the unit cell, showing the clockwise sense of rotation of every molecule with respect to the other molecules. The letters A to K are placed at the N-terminal extreme of every molecule.

Figure 3. Cartoon representation of the crystal structure of UBZ1+2 of TAX1BP1. The eleven structures in the unit cell were superimposed on the backbone of residues 8-62. UBZ1 domain is colored light grey, UBZ2 in dark grey. The side-chains of the two Cys₂His₂ motifs coordinating the two zinc ions (spheres) are shown in sticks. Inset: Tetrahedral coordination of the zinc atom that corresponds to the UBZ1 domain of chain C. The electronic density map localized around the area at a level of 1.5 σ is shown. The approximate distance between the zinc atoms and the Cys-S γ and His-N ϵ 2 atoms is 2.3 Å and 1.97 Å, respectively.

Figure 4. Assigned 800 MHz ¹H-¹⁵N-HSQC spectra of UBZ1+2 of TAX1BP1 at pH6.0, 25°C. All backbone ¹H-¹⁵N pairs are indicated by their corresponding number

in the amino acid sequence. The H ϵ 1 proton of the Trp34 is indicated. Side-chain NH₂ resonances of Asn and Gln residues are connected by a solid line.

Figure 5. Cartoon representation of the ensemble of 20 water-refined NMR structures superimposed on the backbone of residues 8-62. The UBZ1 and UBZ2 domains are colored in green and red, respectively. The side-chains of the two Cys₂His₂ motifs coordinating the two zinc ions (grey spheres) are shown in sticks. Orientation as in Figure 3.

Figure 6. Backbone dynamics of UBZ1+2 of TAX1BP1 measured at 600 MHz and 25 °C. Values for residues in UBZ1 and UBZ2 are colored in red and green, respectively. **A.** ¹H-¹⁵N steady-state NOEs (hetNOE), **B.** ¹⁵N R1, **C.** ¹⁵N R2, **D.** R2 over R1 ratios (R2/R1), **E.** Order parameter (S²), and **F.** Chemical exchange term (K_{ex}). S² and K_{ex} obtained from the TENSOR2 analysis using the model-free approach to describe the rotational diffusion tumbling of UBZ1+2 by an axially symmetric tensor with an anisotropy of D_{par}/D_{perp} = 1.36 ± 0.1. Secondary structure elements are drawn above A and D.

Figure 7. Shape reconstruction of the UBZ1+2 tandem and the individual domains by SAXS. **A.** Experimental scattering curve of UBZ1+2 (grey circles) and theoretical curve (solid line) calculated from the atomic coordinates of the NMR structure. **B.** DAMMIF *ab-initio* reconstructed envelope of the tandem superimposed onto the atomic coordinates. The UBZ1 domain is colored in blue and the UBZ2 domain in red. **C.** Experimental scattering curve of UBZ1 (grey circles) and theoretical curve (solid line) calculated from the rigid body model of UBZ1. **D.**

DAMMIF *ab-initio* reconstructed envelope of UBZ1 superimposed onto the pseudo-atomic coordinates. **E.** Experimental scattering curve of UBZ2 (grey circles) and theoretical curve (solid line) calculated from the rigid body model of UBZ2. **F.** DAMMIF *ab-initio* reconstructed envelope of UBZ2 superimposed onto the pseudo-atomic coordinates. In panels D and F, β -strands are colored in yellow and α -helices in red.

Figure 8. Cartoon representation of the lowest energy NMR structure of UBZ1+2. Side chains of residues that form the hydrophobic core are shown as sticks (Leu12, Leu15, Phe17, Tyr21, Phe26, and Val30 for UBZ1 in red; and Met39, Phe44, Tyr48, Phe53 and Val57 for the UBZ2 domain in green). The conserved Cys and His coordinating the Zn^{2+} atoms (spheres) are highlighted in yellow sticks. The representation shows how residues Phe17 and Phe44 (in cyan) act like aromatic anchors at the zinc coordination sites, both in the consensus position. Orientation as in Figure 5.

Figure 9. Comparison between the solution and crystal structures of UBZ1+2 of TAX1BP1. **A.** Variation of the PHI torsion angles of residue 9-62 of UBZ1+2 calculated from the X-ray (black) and NMR (red) ensembles; the error bars correspond to the standard deviation for every angle. **B.** Representation of the intensity of 1H - ^{15}N NMR signals of every amino acid of UBZ1+2 relative to that of C-terminal residue 69 in the HSQC spectrum shown in figure 4. Secondary structure elements are drawn above.

Figure 10. Structural differences between solution and crystal structures of

UBZ1+2 of TAX1BP1. Cartoon representation of X-ray molecule E in the asymmetric unit as shown in Figure 2. UBZ1 in light grey, UBZ2 in dark grey. The representative NMR structure was superimposed on the UBZ1 domain of X-ray molecule E. UBZ1 in red, UBZ2 in green. The different orientation of the UBZ2 domain relative to the UBZ1 domain in the two structures is indicated by an arrow.

Figure 11. Multiple sequence alignment of UBZ domains. Alignment was performed with the program Geneious 6.1.4⁵⁵ for the UBZ domains of TAX1BP1 (Human UBZ1 highlighted by a red box and UBZ2 by a green one), DNA pol eta (Human UBZ highlighted by a blue box) and WRNIP1 (Human UBZ highlighted by a pink box) present in different organisms.

Figure 12. Unrooted phylogenetic tree based on the multiple sequence alignment of the UBZ domains of TAX1BP1, WRNIP1 and DNA pol eta in different organism. Color-coding of the tree as in Figure 11. For every group a cartoon representation of the human UBZ domains is shown (UBZ domain of DNA pol eta, PDB code 2I5O; UBZ domain of WRNIP1, PDB code 3VHS and UBZ1+2 tandem of TAX1BP1, PDB code 2M7Q and 4BMJ). The four extra loop residues in the UBZ domains of TAX1BP1 are represented as sticks. Zn²⁺ atoms are represented as black spheres.

TABLES

Table 1. Crystallographic data collection and refinement analysis of UBZ1+2. Numbers in parenthesis indicate the highest resolution shell statistics.

Space group	P6 ₅ 22
Unit cell dimensions (Å)	a=118.88, b=118.88, c=327.35 α=β=90°; γ=120°
Resolution range (Å) (last shell)	51.50-2.59 (2.73-2.59)
Observations (unique)	527791 (43605)
Completeness (%) (last shell)	100 (99.9)
Multiplicity	12.1 (10.4)
R _{sym} ^a	0.095 (0.793)
R _{pim} ^b	0.030(0.264)
<I/σ>	19.1 (3.1)
Non-hydrogen atoms	5887
Heteroatoms	11 Cl, 22 Zn
Water	92
R _{factor} ^c (R _{free} ^d)	0.215 (0.255)
RMSD ^e bond length (Å)	0.010
RMSD ^e angles (°)	1.346
RMSD from average for residues 8-62 (Å) ^e	
Backbone N, CA, C'	0.16 ± 0.04
Heavy atoms	0.28 ± 0.06
Ramachandran plot ^f	
Most favored regions (%)	90.69
Additional allowed regions (%)	8.85
Generously allowed regions (%)	0
Disallowed regions (%)	0.46

^aR_{sym} is the unweighted R value on I between symmetry mates.

^bR_{pim} = $\{\sum hkl [I/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|\} / \sum hkl \sum_i I_i(hkl)$, where $I_i(hkl)$ are the observed intensities, $\langle I(hkl) \rangle$ are the average intensities and N is the multiplicity of the reflection hkl
^c<I/σ> is the average of the relation between the intensity of the diffraction and the background

^cR_{factor} = $\sum_{hkl} [|F_{obs}(hkl) - F_{calc}(hkl)|] / \sum_{hkl} F_{obs}(hkl)$, where $F_{obs}(hkl)$ and $F_{calc}(hkl)$ are the structure factors observed and calculated, respectively.

^dR_{free} is the cross-validation R factor for 5% of reflections against which the model was not refined.

^eCoordinate precision is given as the pair-wise Cartesian coordinate Root Mean Square Deviations from the average structure over the ensemble.

^fValues obtained from the *Validate* analysis of the program *Coot* over all residues.

Table 2. Structural statistics over the twenty lowest-energy water-refined NMR structures of UBZ1+2.

Experimental restraints	
Distance Restraints ^a	1141
Intra-residue (i-j = 0)	305
Sequential (i-j = 1)	370
Medium range (2 ≤ i-j ≤ 4)	295
Long range (i-j ≥ 5)	171
Dihedral restraints	
Torsion angle (phi/psi)	40/40
Restraints statistics	
NOE violations >0.5 Å	0.2 ± 0.4
Dihedral violations >5°	0
CNS energies (Kcal/mol)	
E _{total}	-2393.8 ± 25.7
E _{vdw}	-251.9 ± 12.4
E _{elec}	-2787.2 ± 38.9
RMSD from average for residues 8-62 (Å) ^b	
Backbone N, CA, C'	0.87 ± 0.20
Heavy atoms	1.19 ± 0.17
Ramachandran plot ^c	
Most favoured regions (%)	92.3
Additional allowed regions (%)	7.4
Generously allowed regions (%)	0.3
Disallowed regions (%)	0

^aRestraint statistics reported for all NOEs. ^bCoordinate precision is given as the pair-wise Cartesian coordinate Root Mean Square Deviations from the average structure over the ensemble. ^cValues obtained from the PROCHECK-NMR analysis over all residues including the highly flexible residues.

Table 3. Experimental and model based parameters obtained from the analysis of the small angle X-ray scattering data of the UBZ1+2 tandem and the individual domains.

Species	Experimental Molecular Weight SAXS (kDa)	Theoretical Molecular Weight (kDa)	Rg (Å) (exp/model)	Dmax(Å)	χ^2
UBZ ₁₊₂	8.1	8.3	19.16/19.14	66.0	0.76
UBZ ₁	2.8	3.7	11.78/12.2	40.3	1.12
UBZ ₂	2.6	4.3	11.54/11.04	40.9	1.29

Figure 1

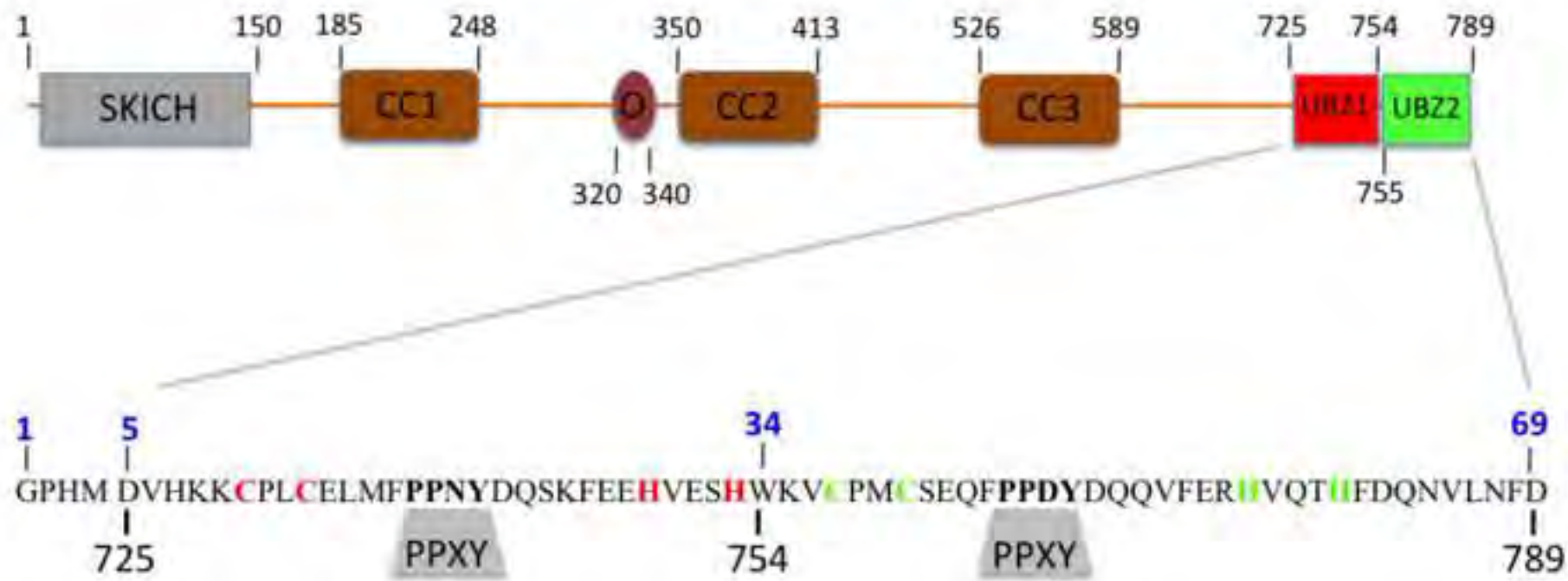


Figure 2

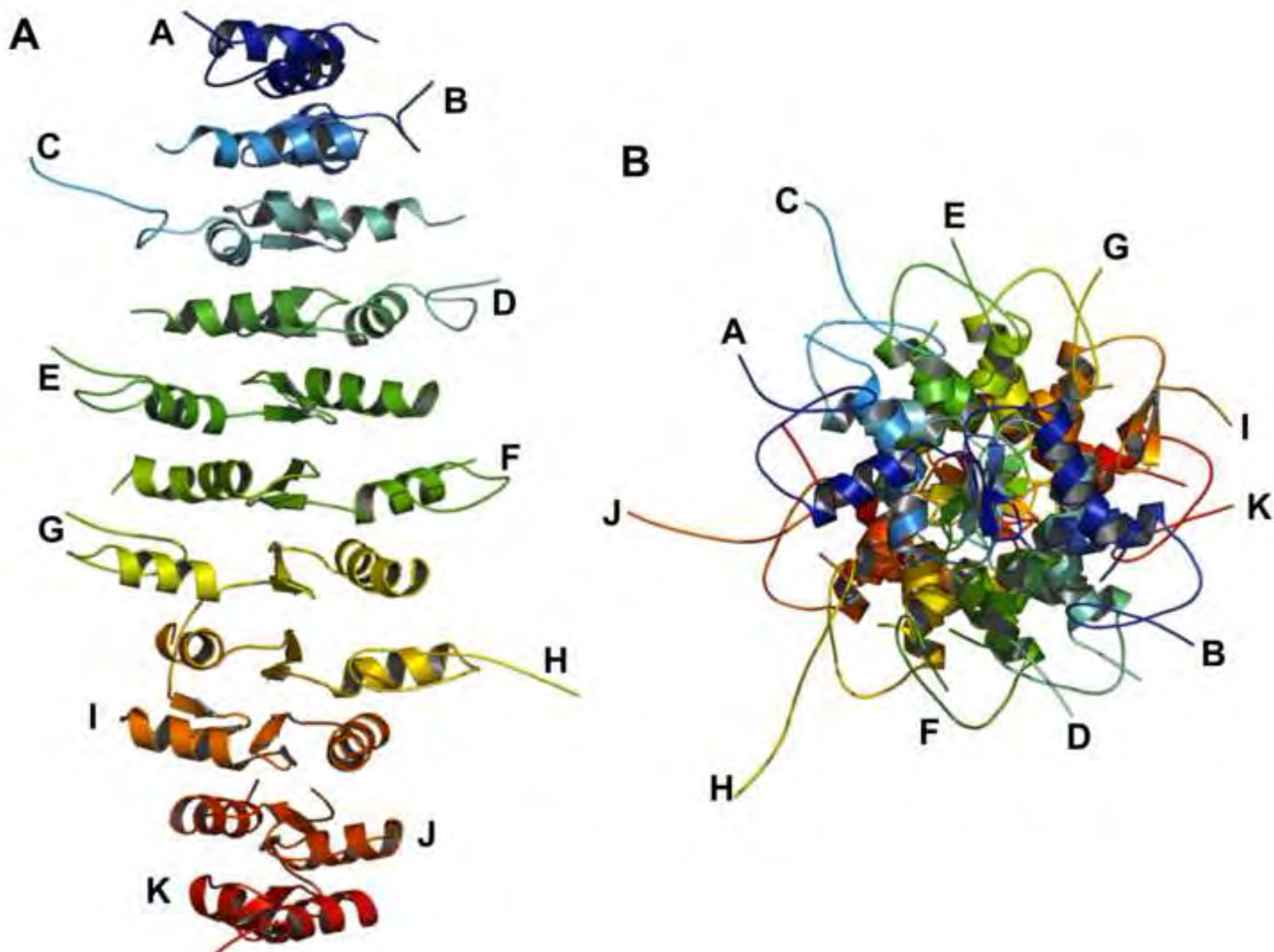


Figure 3

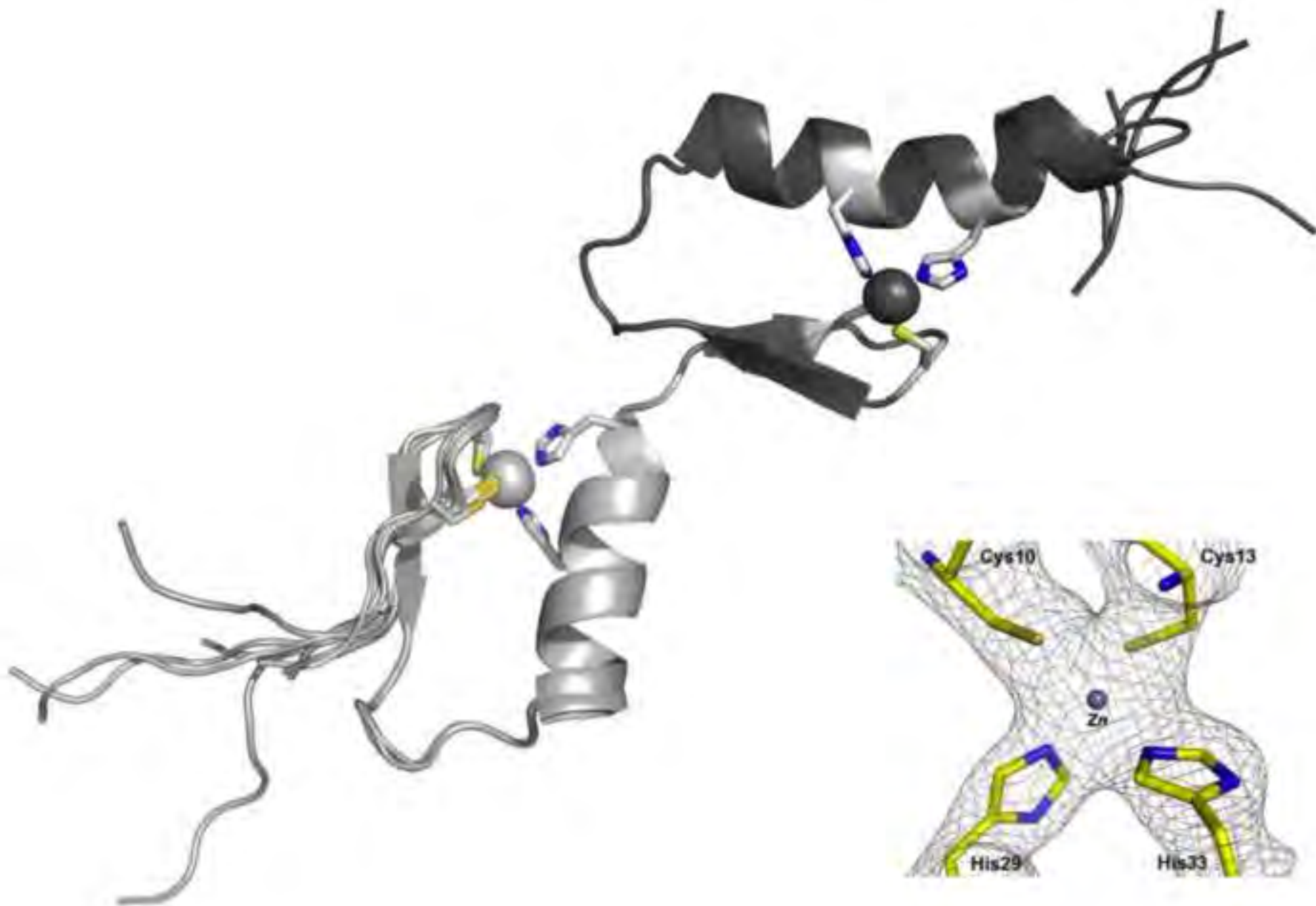


Figure 4

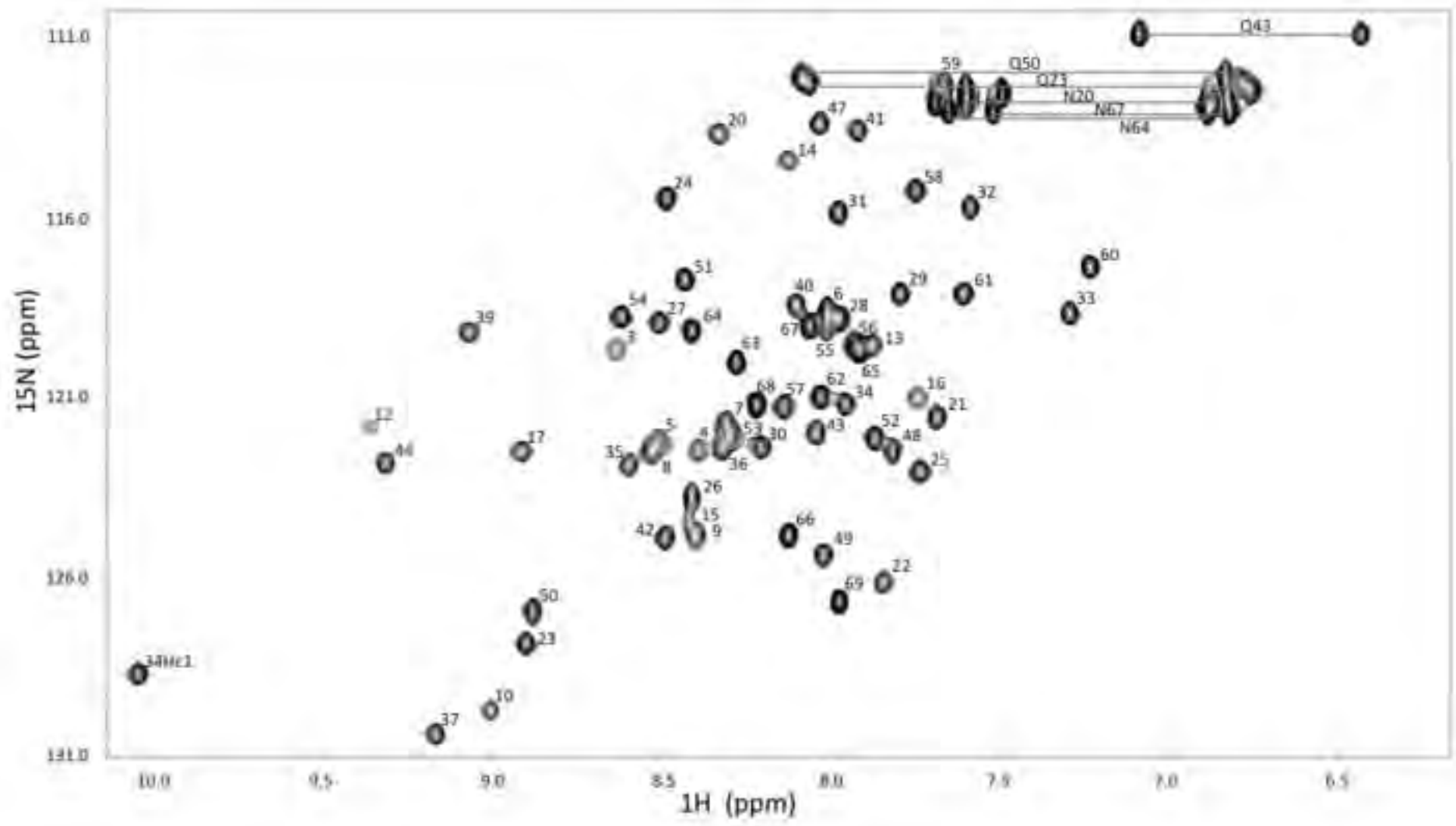


Figure 5



Figure 6

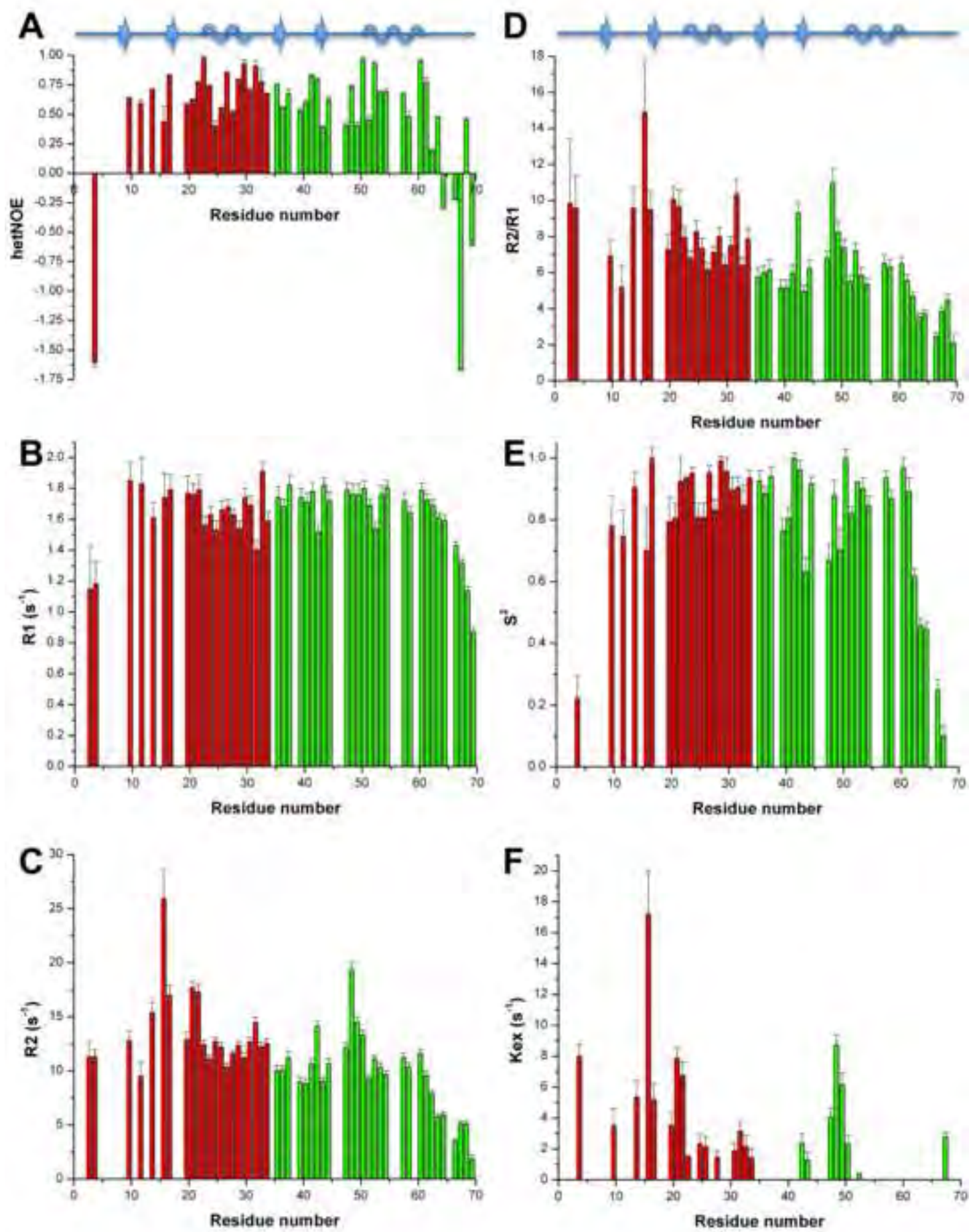


Figure 7

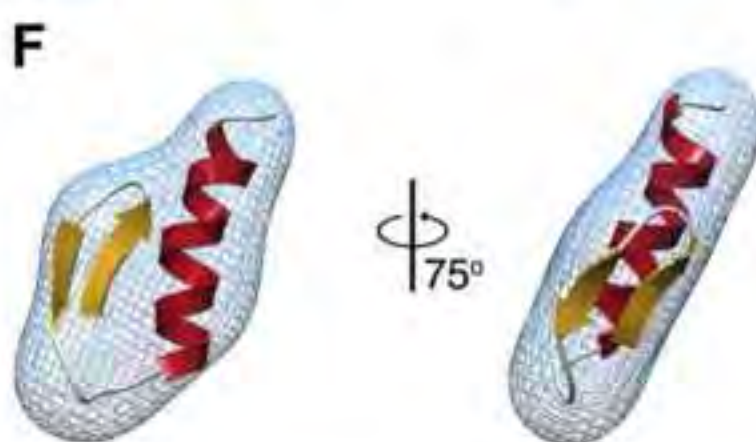
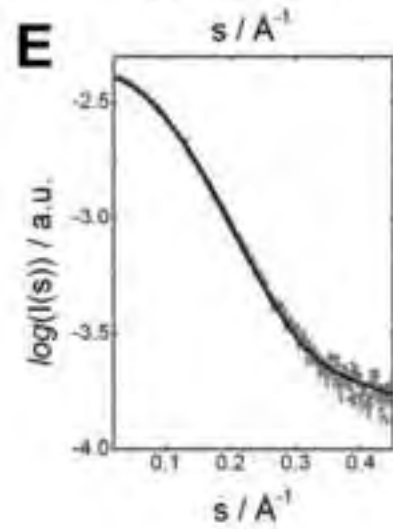
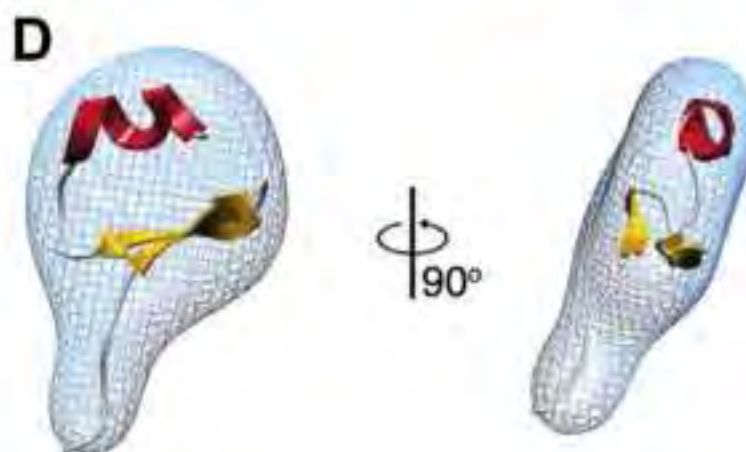
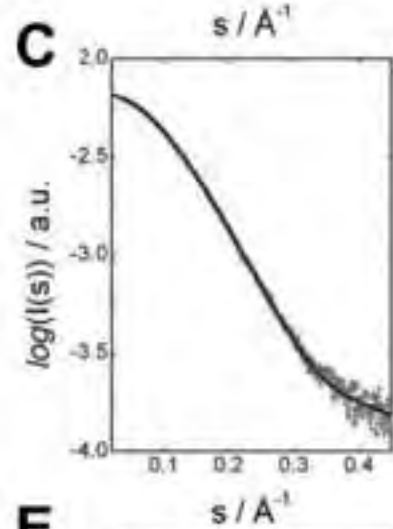
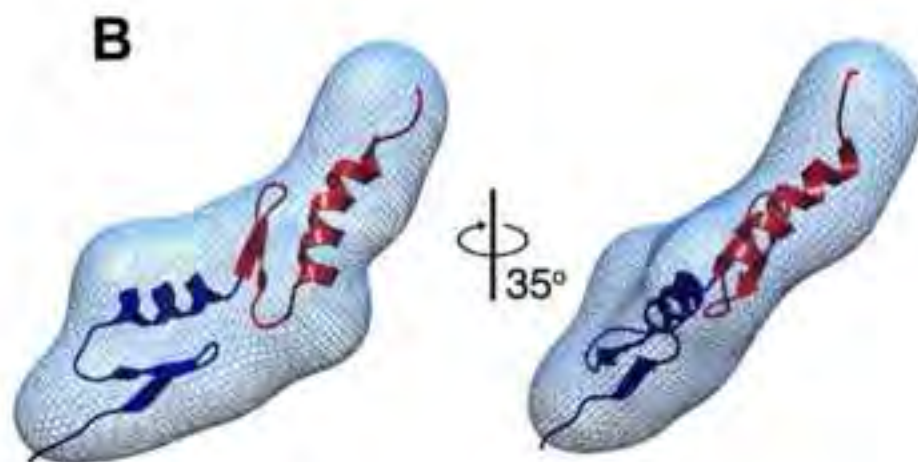
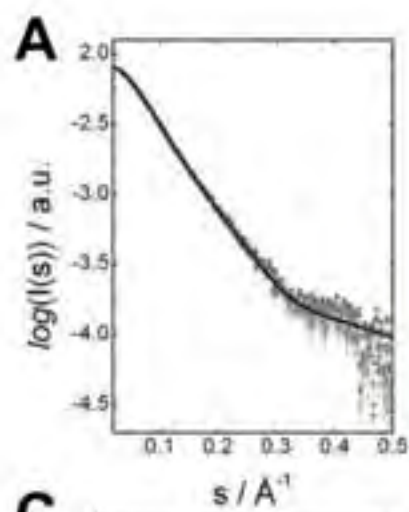


Figure 8

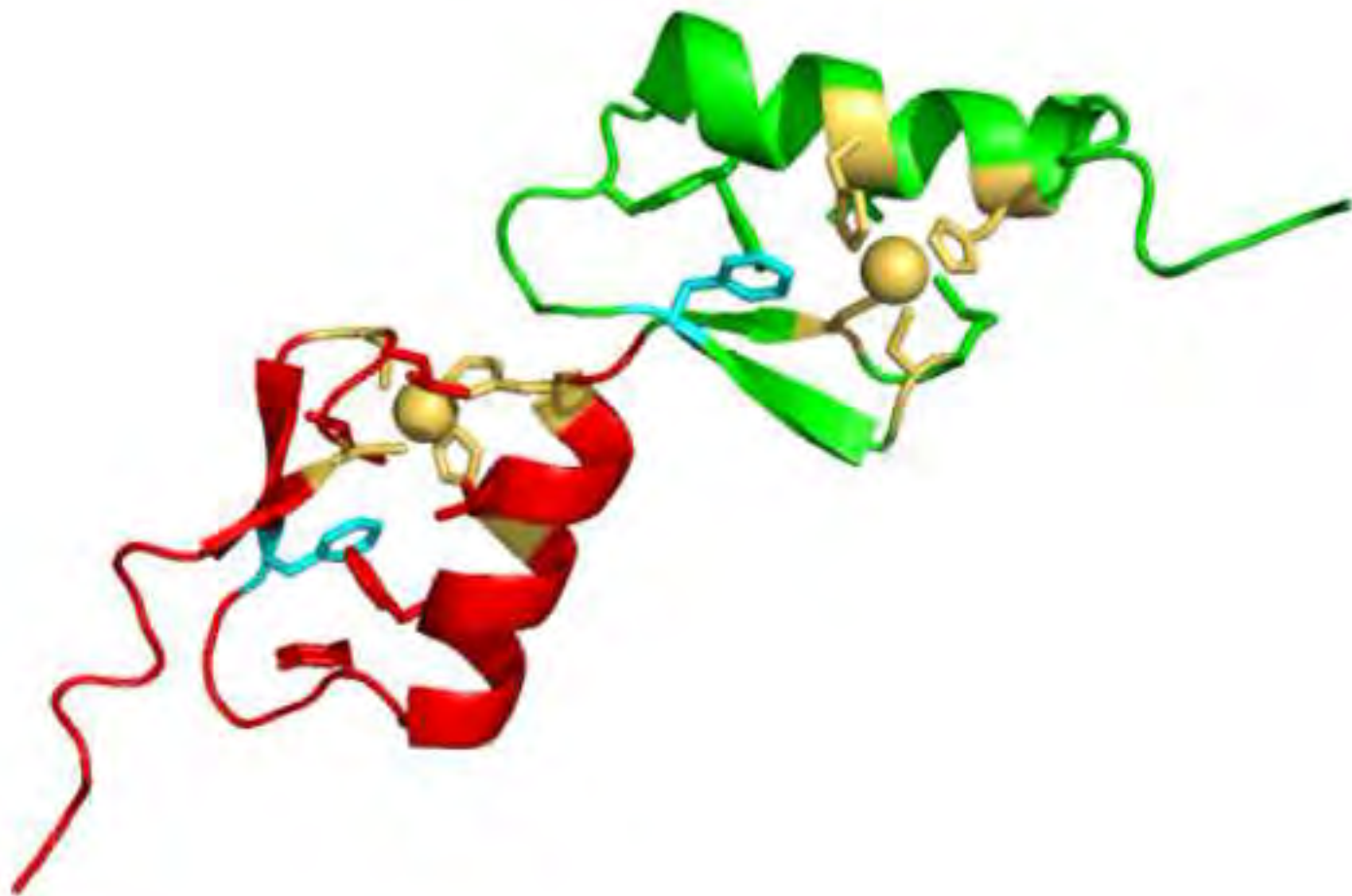


Figure 9

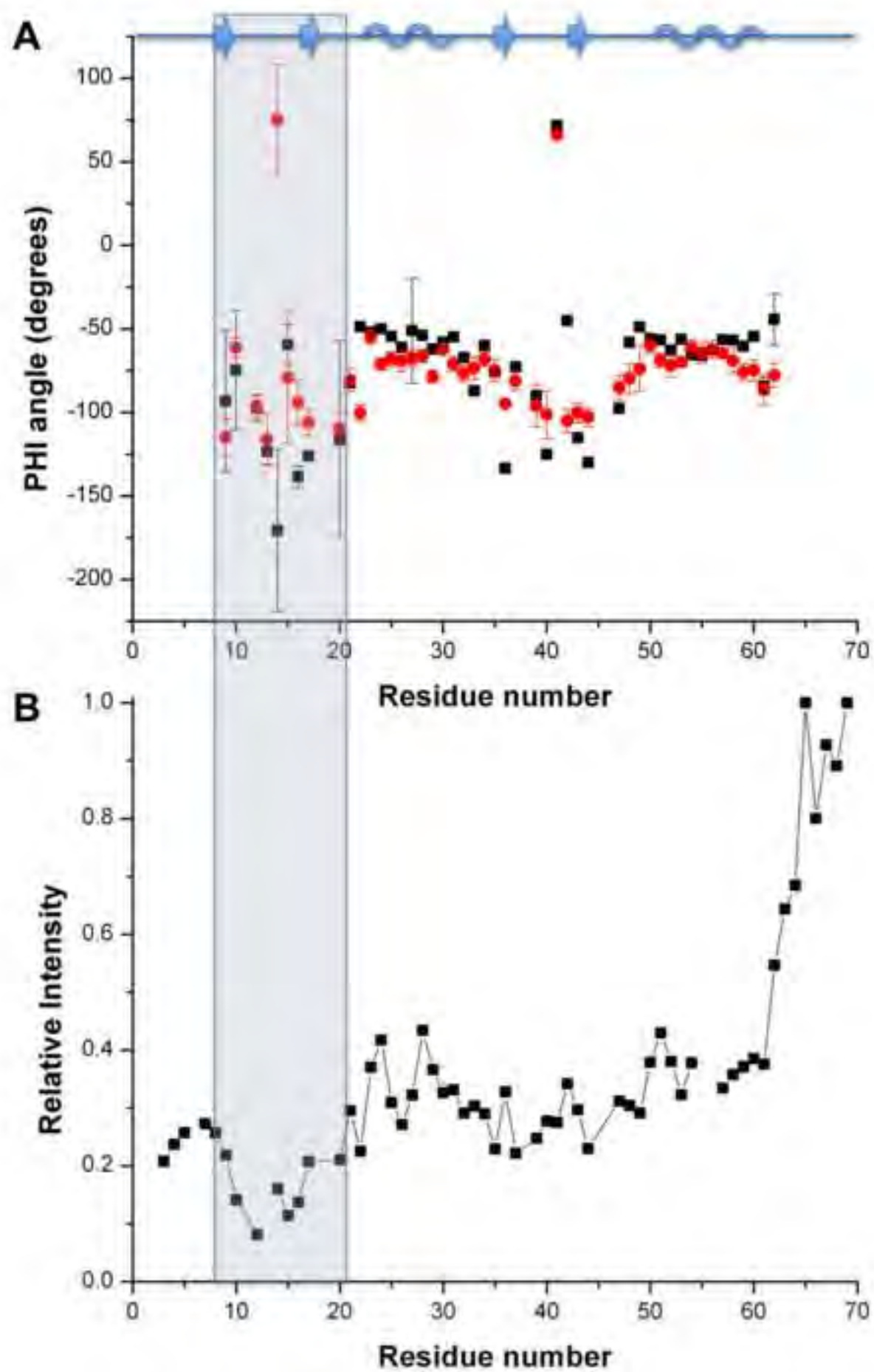


Figure 10



Figure 11

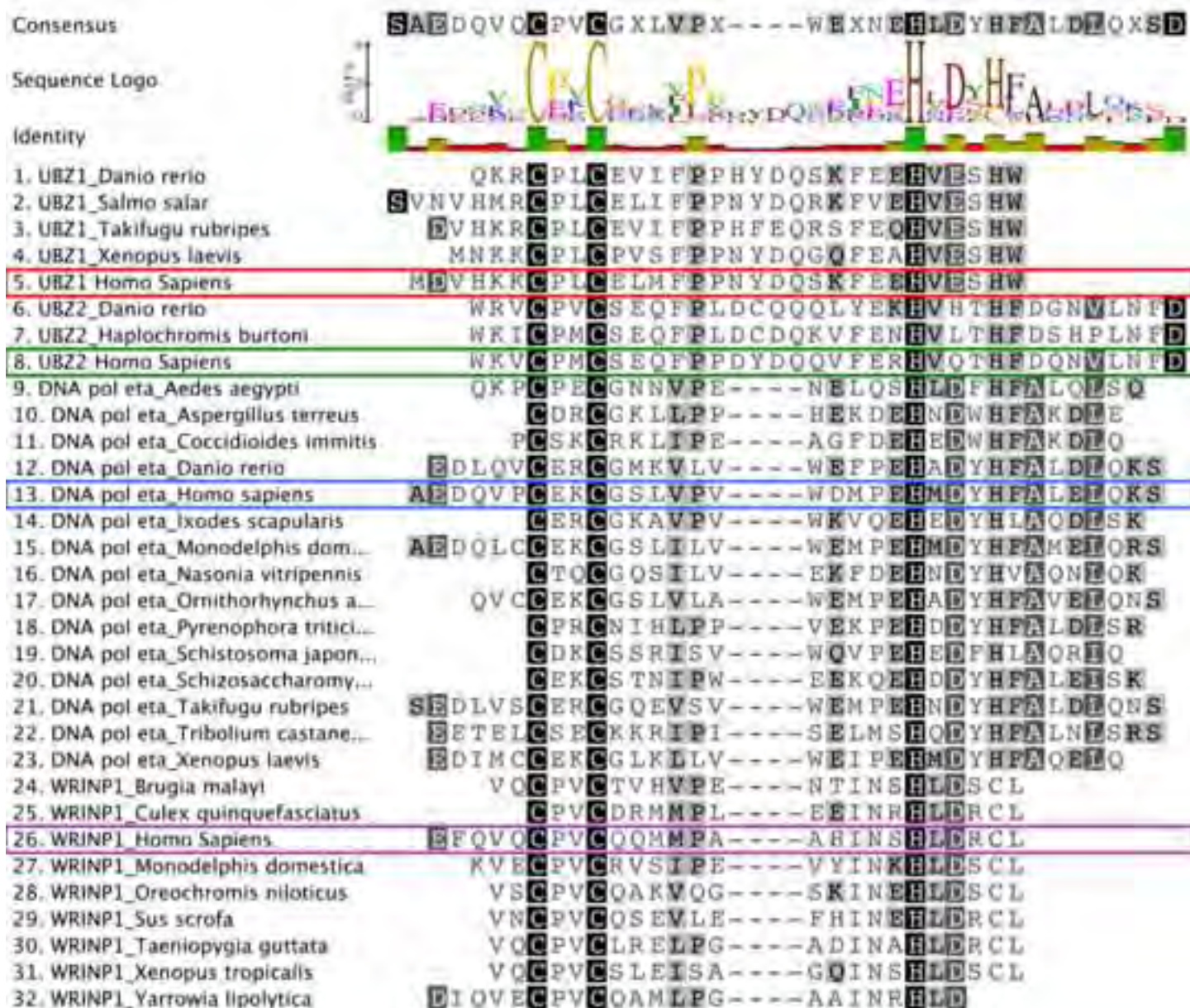


Figure 12

