HHS Public Access

Author manuscript

Proteins. Author manuscript; available in PMC 2021 December 01.

Published in final edited form as:

Proteins. 2020 December ; 88(12): 1701–1711. doi:10.1002/prot.25987.

Crystal structure of a dimerization domain of Drosophila Caprin renders further support for a conserved homodimeric structure in the Caprin protein family

Jiang Zhu[a], **Xia Zhou**[a], **Xiaolan Huang**[b], **Zhihua Du**[a]

[a]Department of Chemistry and Biochemistry, Southern Illinois University at Carbondale, IL 62901, USA

[b]Department of Computer Science, Southern Illinois University at Carbondale, IL 62901, USA

Abstract

Caprin (cytoplasmic activation/proliferation-associated protein) proteins assume diverse functions in many important biological processes, including synaptic plasticity, stress response, innate immune response and cellular proliferation. The Caprin family members are characterized by the presence of a highly conserved homologous region (HR1) at the N-terminus and RGG boxes at the C-terminus. We had previously determined the crystal structures of human Caprin-1 and Caprin-2 fragments corresponding to the C-terminal 2/3 of HR1. Both fragments adopt homodimeric structures. Based on sequence conservation, we speculated that all Caprin proteins should have similar homodimeric structures. Here we report the crystal structure of a fragment (residues 187–309) of Drosophila Melanogaster Caprin (dCaprin). The dCaprin fragment adopts an all α-helical fold which self-associates to form a homodimer. The overall dCaprin homodimeric structure is similar to the Caprin-1 and Caprin-2 homodimeric structures. Most of the amino acids residues mediating homodimerization in the three structures are conserved among all Caprin family members. These structural and sequence data suggest that homodimerization through a conserved dimerization domain is a common structural feature of the Caprin protein family. The dimeric structures may also involve in interaction with Caprin partners. Dimer formation creates a V-shape concave surface that may serve as a protein binding groove. The concave surfaces in Caprin-1, Caproin-2, and dCaprin should have different and specific binding partners due to the large difference in electrostatic potentials. We propose the existence of a multi-functional domain in Caprin proteins, which not only mediate homodimerization but also involve in interaction with specific Caprin partners.

Keywords

dCaprin; Caprin-1; Caprin-2; G3BP1; FMRP; RNA stress granule

Correspondence to Zhihua Du: zdu@chem.siu.edu. Tel: 618-453-7186; Fax: 618-453-6408.

Introduction

Caprin-1, Caprin-2 and dCaprin are three prototypic members of the Caprin (cytoplasmic activation/proliferation-associated protein) family. Caprin-1 and Caprin-2 are human proteins that share two highly conserved homologous regions (HR1 and HR2) (Grill et al., 2004; Shiina, Shinkura, & Tokunaga, 2005). A less conserved E-rich region is present shortly C-terminal to the HR1 region (Figure 1). The other Caprin protein, dCaprin, is from Drosophila Melanogaster. It contains an HR1 region but lacks the HR2 and E-rich regions. All Caprin proteins also contain RGG boxes within the HR2 region and/or at the C-terminus. RGG boxes are characteristic of RNA-binding proteins. Compared the HR1 regions of dCaprin and Caprin-1, they share 31% identity and 50% similarity (29% identity and 52% similarity to Caprin-2).

Caprin-1 is the most intensively studied Caprin protein in the family. Caprin-1 performs various functions. Caprin-1 plays a critical role in normal cellular proliferation (Grill et al., 2004; Wang, David, & Schrader, 2005). The expression level of Caprin-1 is high in proliferating T or B lymphocytes and hemopoietic progenitors in the thymus and spleen. Caprin-1 acts as a positive regulator of cell proliferation at the G1/S transition of cell cycle (Grill et al., 2004; Wang et al., 2005). Caprin-1 and another RNA-binding protein, Ras GTPase activating protein SH3 domain-binding protein 1 (G3BP1), form a stable complex in cytoplasmic RNA granules and selectively binds mRNAs of cyclin D2 and c-Myc (Chiles, 2004; Kaczmarek, Hyland, Watt, Rosenberg, & Baserga, 1985; Pardee, 1989; Solomon et al., 2007). Cyclin D2 and c-Myc are positive regulators of cell cycle G1/S phase transition.

Caprin-1 is highly expressed in brain and plays important roles in brain development. In dendrites of hippocampal and neocortical pyramidal neurons, Caprin-1 localizes in the neuronal RNA granules, which also contain G3BP1 and some ribosome components. The RNA granules bind to mRNAs of key proteins for synaptic plasticity and suppress their local translation. Synaptic stimulation dissociates Caprin-1 from the neuronal RNA granules, releasing the translational suppression of the specific mRNAs (Shiina et al., 2005).

Caprin-1 is involved in interferon (IFN)-mediated antiviral innate immune response. A protein complex, consisting of Caprin-1, G3BP1, and G3BP2, is necessary for the translation of mRNAs of antiviral IFN-stimulated genes (ISGs) (Bidet, Dadlani, & Garcia-Blanco, 2014).

Caprin-1 also plays an important role in cellular stress response. Mammalian cytoplasmic RNA stress granules (SGs) form in response to stress. Caprin-1, G3BP1, and another RNA-binding protein Fragile X mental retardation protein (FMRP) are SG markers (Solomon et al., 2007). Caprin-1 containing SGs are involved in the pathogenesis of several human diseases, such as viral infection by Japanese encephalitis virus (JEV), (H. Katoh et al., 2013), tumorigenesis and metastasis in osteosarcoma (OS), hearing loss due to mutations in the Pou4f3 gene (Collin et al., 2008; Pauw et al., 2008), and Huntington disease (HD) (Ratovitski et al., 2012).

Caprin-2 has distinct functions compared to Caprin-1. Like Caprin-1, Caprin-2 is highly expressed in brain. However, Caprin-2 localizes in a different type of neuronal RNA

granules which do not contain G3BP1 and ribosome components. Suppression of Caprin-2 by siRNA knockdown reduces spine density in neurons (Shiina, Yamaguchi, & Tokunaga, 2010). Caprin-2 also promotes the Wnt signalling pathway (Ding et al., 2008) through interaction with the low-density lipoprotein receptor-related proteins 5 & 6 (LRP5/6, co-receptors for Frizzled). Caprin-2 is also involved in body's response to osmotic stress by enhancing the abundance and poly-A tail length of the mRNA of arginine vasopressin (AVP). AVP is a protein hormone that regulates water and salt levels in the whole body (Konopacka, Greenwood, Loh, Paton, & Murphy, 2015).

Only a few dCaprin relating studies are available (Baumgartner, Stocker, & Hafen, 2013; Papoulas et al., 2010). In early embryos, dCaprin associates with dFMRP to modulate the timing of the mid-blastula transition (MBT). dCaprin and dFMRP collaborate to control cell cycle by regulating translation of specific mRNAs. Translation of maternal Cyclin B mRNA is repressed and translation of zygotic frühstart mRNA is activated (Papoulas et al., 2010). In the ovaries, dCaprin is required for the maintenance of follicle stem cell population and proper encapsulation of developing germ cells. Loss of dCaprin leads to misregulation of Cyclin B and alters the dynamics of the cell cycle. In addition, dFMRP works synergistically with dCaprin in these processes. dCaprin is also involved in growth control of eye and wing epithelial tissues. Lingerer (Lig), an ubiquitin-associated (UBA) domain-containing protein, interacts with the RNA-binding proteins dCaprin, dFMRP, and Rasputin (the Drosophila ortholog of human G3BP1). Via these interactions, Lig inhibits cell proliferation in epithelial tissues (Baumgartner et al., 2013).

Albeit the biological and physiopathological importance of the Caprin proteins, their functional mechanisms remain largely elusive. To help gaining a better understanding of these mechanisms, we had previously determined the crystal structures of a protein fragment within the HR1 regions of Caprin-1 (Wu, Zhu, Huang, & Du, 2015, 2016) and Capring-2 (Wu, Zhu, Huang, Zhou, & Du, 2018). The structural results revealed the existence of a dimerization domain in Caprin-1 and Caprin-2 (residues 132–251 in Caprin-1, residues 199–329 in Caprin-2). Each of the dimerization domains adopts a similar novel all α-helical protein fold that self-associates to form a homodimer. While the overall dimeric structures of Caprin-1 and Caprin-2 are similar, the molecular surface properties of the structures are quite different, suggesting the structures may interact with different protein partners. Indeed, the Caprin-1 region (residues 231–245) required for interaction with FMRP (El Fatimy et al., 2012) is within the α5 helix (residues 230–245) of the Caprin-1 dimeric structure. One side of the α5 helix is fully exposed, readily accessible for protein-protein interaction. The Caprin-1 sequence (residues 352–380) required for interaction with G3BP1 (Solomon et al., 2007) is separated from the dimerization domain by ~100 residues. The dimeric Caprin-1 structure should be able to function as a scaffold for the assembly of a multi-component RNP complex containing Caprin-1, FMRP, G3BP1, their specific mRNAs, and other proteins.

The dimeric structures of Caprin-1 and Caprin-2 in our previous studies identified the residues that mediate the homodimerization of the domains. Most of these residues are conserved in the corresponding dCaprin sequence (Figure 1A), suggesting that dCaprin might also contain a dimerization domain within the HR1 region. Similar to Caprin-1,

dCaprin is known to associate with dFMRP and Rasputin (Baumgartner et al., 2013). It is possible that the Caprin dimeric structures in Caprin-1 and dCaprin, as well as the ways they interact with FMRP/dFMRP and G3BP1/ Rasputin are evolutionarily conserved. To this end, it becomes very critical to gain structural knowledge about the dCaprin protein. In this paper, we present the crystal structure of a dCaprin fragment (residues 187–309) corresponding to the fragments containing the dimerization domains in Caprin-1 and Caprin-2 in our previous studies.

Materials and methods

Protein sample preparation and crystallization

A synthetic DNA with E. coli preferred codons encoding a fragment of dCaprin (residues 187–309), was made by custom gene synthesis (GenScript). The DNA fragment was amplified from the plasmid by PCR using Phusion DNA polymerase. The PCR product was purified by agarose gel electrophoresis and subsequently processed by T4 DNA polymerase in the presence of 2.5 mM dATP. The processed DNA was inserted into an in-house developed cloning vector using ligation independent cloning (LIC) method. The LIC vector contains DNA sequences encoding the Halo-tag (Los et al., 2008), His-tag, and an octapeptide recognition sequence for human rhinovirus (HRV) 3C protease, preceding a specific LIC cloning sequence. Therefore the dCaprin fragment was expressed as a fusion protein containing N-terminal Halo and His tags, followed by a HRV 3C protease recognition motif, and the target protein.

The recombinant plasmid was transformed into NiCo21(DE3) E. coli cells (New England Biolabs). Cell culture was grown in LB media until it reached an OD_{600} of 0.8–1. IPTG (1 mM final concentration) was added to the culture to induce protein expression. After induction, the culture was allowed to grow for 15–18 hours at 12°C before harvest. The overexpressed proteins were purified by cobalt affinity resin. After elution from the resin with a buffer containing 200 mM imidazole, 25 mM Tris (pH=7.0), 200 mM NaCl, the fusion protein was processed by HRV 3C protease in a dialysis tubing (10 ml of protein solution against 5 liters of dialysis buffer containing 25 mM Tris, pH=7.0, 200 mM NaCl) at 4°C overnight. The cleaved tags were separated from the target protein by a reverse IMAC (immobilized-metal affinity chromatography) process with cobalt affinity resin. The purified proteins were concentrated to a concentration of ~10 mg/ml in a buffer containing 25 mM Tris (pH 7.0), 200 mM NaCl. For Se-Met labelled protein, the bacterial cells were grown in M9 minimal culture. When the cultures reached an OD_{600} of 0.6–0.8, six amino acids leucine, isoleucine, lysine, phenylalanine, threonine, and valine were added to the culture at a final concentration of 50–100 mg/L. After half hour growth at 37°C, L-seleno-methionine was added to the cell cultures at a final concentration of 50 mg/L. The subsequent steps of protein expression and purification were identical to those for the native protein.

Crystallization trials were carried out using ten sets of in-house prepared screening solutions each containing 96 different conditions. The solutions use various PEGs as the precipitants. Other variables include buffers, pH values, salts, and cryoprotectants. The crystallization trials were set up by using 96-well format plates. Multiple crystallization conditions were identified after 3–4 days incubation of the plates at 22°C. After optimizing the

crystallization conditions, diffracting crystals were obtained at 22°C by sitting-drop vapour diffusion against 50 microliters of well solution containing 18% PEG2000, 0.1 M BisTris (pH6.8), 5% glycerol, using 96-well format crystallization plates.

Data collection, data processing, and structure determination

Data collection was carried out at beamline 21ID-F of LS-CAT at the Advanced Photon Source (Argonne National Laboratory). Data were processed, integrated, and scaled with the programs Mosflm and Scala in CCP4 (Battye, Kontogiannis, Johnson, Powell, & Leslie, 2011). The structure was solved by single-wavelength anomalous diffraction (SAD) using a data set collected at the peak wavelength of Se (0.97872 Å) on a single crystal containing Se-Met labelled dCaprin proteins. Building the initial structure was performed using the PHENIX package (Adams et al., 2011). Interactive model building was carried out with Coot (Emsley, Lohkamp, Scott, & Cowtan, 2010). The structure was refined using PHENIX (Afonine et al., 2012). Structure determination statistics are shown in Table 1. The figures were prepared with the program PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.). Atomic coordinates and diffraction data for the structures have been deposited in the Protein Data Bank with accession code 6BK4.

Sequence alignments of the dimerization domains of Caprin-1, Caprin-2, and dCaprin were carried out by ClustalW.

Results

dCaprin assumes a dimeric structure similar to those of Caprin-1 and Caprin-2

Several different protein constructs derived from the HR1 region of dCaprin with various N-terminal and C-terminal boundaries were cloned and tested for protein expression in E. coli bacterial host. At first, a construct corresponding to the whole HR1 region (residues 112–309, see Figure 1A) was tested. Many different expression vectors with different fusion tags (including Halo, MBP, NusA, GST, TRX, SUMO, GB1, IgG-Fc, DnaK, etc) were used for the cloning and many different E. coli host strains were tested. However, this protein construct didn’t express in any of the vector-host combinations tested. We then tested shorter protein constructs with a different N-terminus, starting at residues 130, 145, 158 or 187 (all ending at residue 309). Only the protein construct with residues 187–309 as a Halo-tag fusion protein expressed well and remained stable in solution for an extended period of time. After removing the Halo and His tags by HRV 3C protease cleavage, the target protein (containing an artificial sequence of GPSSPS at the N-terminus as a cloning artefact) remained stable in solution. In the cases of Caprin-1 and Caprin-2 we studied previously, similar scenarios were observed. These data indicate that in each of the Caprin proteins, a stable protein domain of about 120 residues exists within the C-terminal portion of the HR1 region.

To solve the structure of the dCaprin fragment independently from the previous Caprin-1 and Caprin-2 structures, SAD phasing was used for structure solution on a data set of a single crystal containing Se-Met labelled proteins, collected at the Se peak wavelength (0.9787 Å). The data collection and structure determination statistics are shown in Table 1.

The dCaprin fragment adopts an all α-helical structure with five α-helices (Figure 1). The structure self-associates to form a homodimer in a head-to-head manner. Electron densities were observed for most of the residues except for several residues at the N-terminus, residues 255–260 of chain A, and residues 278–290 of chain B. The regions with missing densities presumably are more flexible than other regions of the structure. The overall protein fold and dimeric structure assumed by the dCaprin fragments are similar to those of Caprin-1 and Caprin-2 we determined previously (Figures 2A and 2B) (Wu et al., 2016). The most noticeable difference between the dCaprin and Caprin-1 or Caprin-2 structures is that the α1 helix in the dCaprin structure is shorter (by about one and a half helical turns) than the one in the Caprin-1 or Caprin-2 structure (Figures 2A and 2B). Other than this difference, all of the secondary structural elements (helice) and loops have comparable lengths and orientations. The r.m.s.d is 1.4 Å for the superimposition of 942 common atoms between the dCaprin and the Caprin-1 dimeric structures (Figure 2A).

In the dimerization interface, 27 residues from each protomer participate in direct intermolecular interactions. These residues dispersedly locate in several regions of the primary sequence: the α1 and α4 helice, the linker between the α3 and α4 helices, and the C-terminus (Figures 1B). Formation of the homodimer buries a total of 2793 $Å^2$ of solvent accessible surface area from the two protomers. In comparison, the corresponding values for the Caprin-1 dimer and the Caprin-2 dimer are 3398 $Å^2$ and 2810 $Å^2$ respectively (Wu et al., 2016). The buried surface areas in the three Caprin protein dimers are comparable to each other and are significantly larger than the estimated minimal area of 1200 $Å^2$ required for a stable protein-protein complex (Chothia & Janin, 1975; Jones & Thornton, 1995).

The dCaprin dimer shares conserved molecular interactions in the dimerization interface with the Caprin-1 and Caprin-2 dimers

In the dimerization interface of the dCaprin dimeric structure, several factors contribute to the formation of the dimer. First of all, the surface contours of the two protomers in the dimerization interface are highly complementary to each other (Figure 3A). The two protomers may stabilize each other through this shape complementarity. In addition, various kinds of inter-molecular interactions are observed in the dimerization interface, including van der Waals interactions, hydrogen bonds, and electrostatic interactions (Figures 3B and 4).

Hydrophobic surface dominates the dimerization interface. The intermolecular hydrophobic interactions are mediated by the side-chains of many hydrophobic residues (Ile201, Val204, Leu205, Ile206, Cys213, Pro253, Phe261, Ile262, Ala265, Ala269, Phe272, Tyr273, Ile276, and Leu304), as well as the aliphatic portions of the side-chains of many polar residues (Thr194, Glu197, Thr198, Lys210, Arg202, Gln208, Thr250, Arg252, Gln266, Ser268, and Gln299) (Figures 3B and 4). Figure 3B depicts a region of the dimerization interface, with one protomer rendered in surface mode and the other promoter in ribbon mode (the side-chains of some of the dimerization mediating residues are shown in sticks inside a mesh representing the observed electron density). In this view, the molecular surface shows a large and continuous hydrophobic area defined by the side-chains of Lys200, Ile201, Val203, Leu205, Phe261, Ile262, Ala265, Tyr273, and Ile276. Most of this area participate in

hydrophobic contact with the other protomer. Almost all of the hydrophobic residues involved in homodimerization of the dCaprin protein are conserved in the Caprin-1 and Caprin-2 sequences (Figure 1B).

As many as fifteen intermolecular hydrogen bonds are also observed in the dimerization interface (Figure 4). These hydrogen bonds are mediated by the polar side-chains of Thr194, Glu197, Lys200, Arg202, Gln208, Asn209, Arg252, Gln266, Tyr273, Asn277, Gln299, and Asp305. In the cases of Caprin-1 and Caprin-2 dimeric structures, a fewer number of intermolecular hydrogen bonds are observed in the dimerization interface. Among the intermolecular hydrogen bonds, the three hydrogen bonds formed by the side-chains of the $Gln208^{A}$-$Gln208^{B}$ pair (two hydrogen bonds, A and B denote residues in protomers A and B of the homodimer) and the $Glu197^{A}$-$Lys200^{B}$ pair (1 hydrogen bond) are conserved in all three structures. Glu-197, Lys200, and Gln208 are conserved in the three Caprin sequences (Figure 1B). They are all located in the same α1 helix of the structures, spanning three helical turns (Figure 3B). These specific intermolecular interactions, compared to the less specific hydrophobic interactions, may have an important role in defining the relative spatial positioning of the two α1 helice from the two interacting protomers.

The dCaprin dimerization interface also show a couple of intermolecular electrostatic interactions, mediating by the $Glu197^{A}$ -$Lys200^{B}$ and $Arg252^{A}$-$Asp305^{B}$ pairs (Figures 3B and 4). The $Glu197^{A}$-$Lys200^{B}$ intermolecular electrostatic interaction is conserved in the Caprin-1 and Caprin-2 structures.

Overall, the detailed intermolecular interactions that mediate the homodimerization of the dCaprin structure are largely conserved in the dimeric structures of Caprin-1 and caprin-2 we determined previously.

Formation of a homodimeric structure may represent a common feature of the Caprin protein family

We have determined the crystal structures of the dimerization domains of the three prototypic members of the Caprin protein family, namely dCaprin, Caprin-1, and Caprin-2. The structures were determined independently from each other by SAD phasing method. The three structures show conserved intermolecular interactions for homodimerization. These results raise the possibility that similar homodimeric structure may exist in other members of the Caprin protein family.

To explore this possibility further, we retrieved all available full-length Caprin protein sequences from the NCBI protein sequence database and performed a multiple sequence alignment on the sequences corresponding to the dimerization domains of Caprin-1, Caprin-2, and dCaprin (Figure 5). As can be seen in Figure 5, the sequences have a high degree of overall conservation. Especially, most of the hydrophobic residues that involved in the homodimerization of dCaprin are conserved among all Caprin protein sequences. Moreover, Glu-197, Lys200, and Gln208 in dCapin, which mediate intermolecular hydrogen bonds in the central area of the dimerization interface of the dCaprin demeric structure, are also conserved in all Caprin sequences. These sequence data strongly suggest that the overall dimeric structures observed in the Caprin-1, Caprin-2, and dCaprin cases should be present

in all members of the Caprin protein family. Formation of a homodimeric structure may represent a conserved common feature of the Caprin protein family.

The dimeric structures of dCaprin, Caprin-1, and Caprin-2 have distinctly different molecular surfaces

Caprin proteins interact with other proteins for function. In the case of Caprin-1, several protein partners have been identified, including FMRP, G3BP1, and the JEV core protein. It is known that residues within the dimerization domain directly interact with FMRP (El Fatimy et al., 2012). The α5 helix in the Caprin-1 dimeric structure contains the FMRP-interacting motif (residues 231–245). The α5 helix only accounts for a small fraction of the molecular surface of the Caprin-1 dimeric structure. It is possible that the dimeric structure may also be able to use other surface areas to mediate additional protein-protein interactions. To this end, it is meaningful to compare the molecular surfaces of the three dimeric structures of dCaprin, Caprin-1, and Caprin-2. The results may provide insights into how different Caprin proteins interact with their specific protein partners through the dimeric structures.

While the sequences of the dimerization domains of Caprin-1, Caprin-2, and dCaprin are highly conserved (Figure 1), the relative compositions of negatively charged and positively charged residues within the sequences are quite different. The sequence of the Caprin-1 dimerization domain consists of significantly more negatively charged residues (26 Asp+Glu vs 17 Arg+Lys, accounting for 21% and 14% of the residues in the sequence shown in Figure 1B respectively). In contrast, the sequence of the Caprin-2 dimerization domain has equal number of negatively charged and positively charged residues (17 Asp+Glu vs 17 Arg +Lys, each accounting for 14% of the residues in the sequence shown in Figure 1B). The sequence of the dCaprin dimerization domain has slightly more negatively charged residues than positively charged residues (20 Asp+Glu vs 17 Arg+Lys, accounting for 16% vs 14% of the residues in the sequence shown in Figure 1B). Overall, the Caprin-1 dimerization domain is more negatively charged than the Caprin-2 and dCaprin dimerization domains. Interestingly, the “extra” acidic residues in the Caprin-1 structure (compared to the Caprin-2 and dCaprin structures) are all located in the V-shaped concave molecular surface of the dimeric structures (Wu et al., 2016).

It was revealed previously that there is a large and continuous negatively charged molecular surface in the V-shaped concave area of the Caprin-1 dimeric structure (Wu et al., 2016) (Figure 6). This surface is dominated by 24 negatively charged residues (12 from each protomer, Asp157, Asp161, Asp162, Glu163, Asp167, Glu180, Glu181, Asp187, Glu188, Asp194, Glu196, and Asp198). In the Caprin-2 sequence, residues at 9 out of the 12 corresponding positions have either positively charged or neutral residues (Q225 vs D157, Q229 vs D161, H231 vs E163, S248 vs E180, K249 vs E181, I255 vs D187, K256 vs E188, C262 vs D194, and N266 vs D198) (Figure 1B). In the dCaprin sequence, 7 out of the 12 corresponding positions have either positively charged or neutral residues (N212 vs D157, Q218 vs E163, N235 vs E180, T236 vs E181, K243 vs E188, Q249 vs D194, and A256 vs D198) (Figure 1B). Due to these different amino acids, the V-shaped concave molecular surfaces of the three Caprin (Caprin-1, Caprin-2 and dCaprin) dimeric structures have

different electrostatic surface properties (Figure 6). If these molecular surfaces are used in mediating protein-protein interactions, different protein partners might be expected for the three different Caprin proteins.

Discussion

Results from this study and our previously published works on human Caprin-1 and Caprin-2 (Wu et al., 2015, 2016) showed that all three prototypic members of the Caprin protein family contained a similarly structured dimerization domain. The three dimeric structures were determined independently from each other. The high resolution structures revealed the intermolecular interactions mediating the formation of the dimeric structures. Most of these interactions are conserved in the three structures. Based on sequence conservations among members of the Caprin family, formation of a dimeric structure as in the cases of Caprin-1, Caprin-2, and dCaprin should be a common structural feature of the Caprin protein family.

The HR1 region is the only conserved region among the Caprin proteins from Drosophila to human (Figure 1A). In spite of the substantial effort we put into the cloning and expression of the full HR1 regions of the three prototypic members (Caprin-1, Caprin-2, and dCaprin) of the Caprin protein family, none of the HR1 region could be expressed. Many protein constructs corresponding to different lengths within the HR1 regions were tested. In all three proteins, only the constructs identified as the dimerization domains that account for the C-terminal 60% of the HR1 region were expressed successfully. Based on these observations and our structural results on all three dimeric structures, we propose that the dimeric structure should be considered as the signature structure of the Caprin-1 protein family.

Conservation of the dimeric structures from Drosophila to human imply that they should be important for the functions of the Caprin proteins. At the present time, knowledge about the biological functions of the Caprin protein family is still limited. Studies are only available for Caprin-1, Caprin-2, and dCaprin. For many of the reported functions of these proteins, Caprin-containing macromolecular assemblies are involved, such as neuronal RNA granules and stress granules (Baumgartner et al., 2013; H. Katoh et al., 2013; Ratovitski et al., 2012; Shiina et al., 2005; Shiina & Tokunaga, 2010; Shiina et al., 2010; Solomon et al., 2007). The conserved homodimeric structures of the Caprin proteins may provide a useful scaffold for the assemblies. In the case of Caprin-1, some functions involve the interactions of Caprin-1 with G3BP1 and FMRP (such as in the stress granules) (H. Katoh et al., 2013; Solomon et al., 2007). The amino acids sequences participating in Caprin-1/G3BB1 and Caprin-1/FMRP interactions were determined (El Fatimy et al., 2012; Solomon et al., 2007). The G3BP1-interacting sequence of Caprin-1 resides in the HR2 region (residues 352–380) (Solomon et al., 2007), therefore G3BP1 (which is implicated in all reported Caprin-1 functions) should be able to interact with Caprin-1 in its dimeric form mediated by the dimerization domain. The FMRP-interacting sequence of Caprin-1 (residues 231–245) is within the dimerization domain and forms the α5 helix in the structure (El Fatimy et al., 2012). This is an amphipathic α helix in all three structures from Caprin-1, Caprin-2, and dCaprin (Figure 1B). The hydrophobic side of the helix is mainly buried in the interior of the structure, while the hydrophilic side is fully exposed in a widely accessible molecular surface, opposite to

the dimerization interface (Figures 1C and 2). This surface could interact with FMRP in such a way that would not interfere with the formation of the Caprin dimer. In short, a Caprin-1 dimer formed by the dimerization domain we identified could interact with two FMRP and two G3BP1 proteins. FMRP and G3BP1 (Vognsen, Moller, & Kristensen, 2013) both were reported to have the potential to form homodimers (Adinolfi et al., 2003; Dolzhanskaya, Merz, Aletta, & Denman, 2006). Homodimerization of the Caprin-1 bound FMRP and G3BP1 would grow the protein-protein interaction network involving Caprin-1, FMRP, and G3BP1. Each of these RNA binding proteins would bind to their specific RNA targets in the macromolecular assembly. In Drosophila, dCaprin, dFMRP and Rasputin are found in the same complex (Baumgartner et al., 2013). The amino acids sequences mediating the protein-protein interactions in the complex have not been determined. Our current study reveals the formation of a dCaprin homodimer, and the structural conservation of the homodimeric structures from Drosophila to human. It remains to be seen whether other protein-protein interactions in the dCaprin/dFMRP/Rasputin complex are similar to those in the human Caprin-1/FMRP/G3BP1 complex.

Besides mediating homodimerization of the Caprin proteins, the dimerization domain may also assume other functions. While biological data regarding interactions of Caprin proteins with their partners are limited, it is know that in the case of Caprin-1, sequence within the dimerization domain is responsible for interaction with FMRP (El Fatimy et al., 2012). Our structural results show that the FMRP-interacting sequence is presented on an accessible surface of the Caprin dimeric structure (Wu et al., 2016). Other surface areas of the Caprin dimeric structures may also provide binding sites for Caprin partners. Most noticeably, dimerization creates a large V-shape concave surface area (Figure 6), which can potentially serve as a protein binding groove. Although the overall dimeric structures of Caprin-1, Caprin-2, and dCaprin are similar to each other, the surface residues show significant degrees of variation. On the V-shape concave surfaces, the three dimeric structures show very different molecular electrostatic potentials (Figure 6). In Caprin-1, the concave surface is predominantly negatively charged, and the E-rich sequence is adjacent to this surface area (Wu et al., 2016). We therefore speculated that the Caprin-1 concave surface might interact with the Caprin-1 partner JEV core protein, which is a highly basic protein with 25 basic residues (Arg/Lys) and 4 acidic residues (Asp/Glu) out of 105 amino acids (Hiroshi Katoh et al., 2013). The concave surfaces in the other two dimeric structures (Caprin-2 and dCaprin) have significantly fewer number of acidic residues (Figure 6). These surfaces would be less likely to bind to highly basic proteins. Caprin proteins are functionally diverse. In human, Caprin-1 and Caprin-2 assume distinct and non-redundant functions. Different functions of the Caprin proteins depend on their interactions with specific protein and RNA partners. Results from our studies on Caprin-1, Caprin-2, and dCaprin suggest that the Caprin dimeric structures may also involve in binding specific Caprin partners, thus contributing to functional specificities of the Caprin proteins.

Conclusion

This paper presents the latest results from our research on the structures and functions of the Caprin protein family. We had previously determined two crystal structures of human Caprin-1 and Caprin-2 fragments and identified a homodimerization domain in the two

proteins. Sequence data suggested that similar dimeric structures might exist in all Caprin proteins. To prove this hypothesis, it is essential to determine an atomic resolution structure of a corresponding fragment from a Caprin protein that has a distant evolution relationship with the human Caprin proteins. With high resolution structures for the three prototypic members of the Caprin family (from human to Drosophila) become available now, we are able to analyze the detailed molecular interactions that mediate homodimerization and reveal that most of the interactions should be conserved among all Caprin family members. In short, results from our current and previous studies establish a conserved overally homodimeric structure as a signature feature of the Caprin protein family.

The results further suggest that the Caprin homodimeric structures also participate in specific interaction with different Caprin partners, thus contributing to the diverse functions of the Caprin proteins. In the case of Caprin1, the FMRP binding site locates on the opposite molecular surface from the homodimerization interface. Other surface areas of the Caprin homodimer could also provide binding sites for other Caprin partners. A large V-shape concave surface is formed upon homodimerization, which may serve as a protein binding groove. The concave surfaces in dCaprin, caprin1, and Caprin2 show large difference in electrostatic potentials, which may dictate binding of specific partners.

Acknowledgement

We thank staff of the LS-CAT at the Advance Photon Source (Argonne National Laboratory) for assistance with data collection. The work was supported by the start-up fund and a seed grant from Southern Illinois University Carbondale, as well as a grant from the National Institute of Health (1R15GM116062-01) to Z. Du.

Reference

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Echols N, Headd JJ, … Zwart PH (2011). The Phenix software for automated determination of macromolecular structures. Methods, 55(1), 94–106. doi:10.1016/j.ymeth.2011.07.005 [PubMed: 21821126]

Adinolfi S, Ramos A, Martin SR, Dal Piaz F, Pucci P, Bardoni B, … Pastore A (2003). The N-terminus of the fragile X mental retardation protein contains a novel domain involved in dimerization and RNA binding. Biochemistry, 42(35), 10437–10444. doi:10.1021/bi034909g [PubMed: 12950170]

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, … Adams PD (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr, 68(Pt 4), 352–367. doi:10.1107/s0907444912001308 [PubMed: 22505256]

Battye TG, Kontogiannis L, Johnson O, Powell HR, & Leslie AG (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr, 67(Pt 4), 271–281. doi:10.1107/s0907444910048675 [PubMed: 21460445]

Baumgartner R, Stocker H, & Hafen E (2013). The RNA-binding proteins FMR1, rasputin and caprin act together with the UBA protein lingerer to restrict tissue growth in Drosophila melanogaster. PLoS Genet, 9(7), e1003598. doi:10.1371/journal.pgen.1003598 [PubMed: 23874212]

Bidet K, Dadlani D, & Garcia-Blanco MA (2014). G3BP1, G3BP2 and CAPRIN1 are required for translation of interferon stimulated mRNAs and are targeted by a dengue virus non-coding RNA. PLoS Pathog, 10(7), e1004242. doi:10.1371/journal.ppat.1004242 [PubMed: 24992036]

Chiles TC (2004). Regulation and function of cyclin D2 in B lymphocyte subsets. J Immunol, 173(5), 2901–2907. [PubMed: 15322145]

Chothia C, & Janin J (1975). Principles of protein-protein recognition. Nature, 256(5520), 705–708. [PubMed: 1153006]

Collin RW, Chellappa R, Pauw RJ, Vriend G, Oostrik J, van Drunen W, … Kremer H (2008). Missense mutations in POU4F3 cause autosomal dominant hearing impairment DFNA15 and affect subcellular localization and DNA binding. Hum Mutat, 29(4), 545–554. doi:10.1002/humu.20693 [PubMed: 18228599]

Ding Y, Xi Y, Chen T, Wang JY, Tao DL, Wu ZL, … Li L (2008). Caprin-2 enhances canonical Wnt signaling through regulating LRP5/6 phosphorylation. J Cell Biol, 182(5), 865–872. doi:10.1083/jcb.200803147 [PubMed: 18762581]

Dolzhanskaya N, Merz G, Aletta JM, & Denman RB (2006). Methylation regulates the intracellular protein-protein and protein-RNA interactions of FMRP. J Cell Sci, 119(Pt 9), 1933–1946. doi:10.1242/jcs.02882 [PubMed: 16636078]

El Fatimy R, Tremblay S, Dury AY, Solomon S, De Koninck P, Schrader JW, & Khandjian EW (2012). Fragile X mental retardation protein interacts with the RNA-binding protein Caprin1 in neuronal RiboNucleoProtein complexes [corrected]. PLoS One, 7(6), e39338. doi:10.1371/journal.pone.0039338 [PubMed: 22737234]

Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr, 66(Pt 4), 486–501. doi:10.1107/s0907444910007493 [PubMed: 20383002]

Grill B, Wilson GM, Zhang KX, Wang B, Doyonnas R, Quadroni M, & Schrader JW (2004). Activation/division of lymphocytes results in increased levels of cytoplasmic activation/proliferation-associated protein-1: prototype of a new family of proteins. J Immunol, 172(4), 2389–2400. [PubMed: 14764709]

Jones S, & Thornton JM (1995). Protein-protein interactions: A review of protein dimer structures *Prog*. Biophys. Mol. Biol, 63, 31–59.

Kaczmarek L, Hyland JK, Watt R, Rosenberg M, & Baserga R (1985). Microinjected c-myc as a competence factor. Science, 228(4705), 1313–1315. [PubMed: 4001943]

Katoh H, Okamoto T, Fukuhara T, Kambara H, Morita E, Mori Y, … Matsuura Y (2013). Japanese Encephalitis Virus Core Protein Inhibits Stress Granule Formation through an Interaction with Caprin-1 and Facilitates Viral Propagation. Journal of Virology, 87(1), 489–502. doi:10.1128/JVI.02186-12 [PubMed: 23097442]

Katoh H, Okamoto T, Fukuhara T, Kambara H, Morita E, Mori Y, … Matsuura Y (2013). Japanese encephalitis virus core protein inhibits stress granule formation through an interaction with Caprin-1 and facilitates viral propagation. J Virol, 87(1), 489–502. doi:10.1128/jvi.02186-12 [PubMed: 23097442]

Konopacka A, Greenwood M, Loh SY, Paton J, & Murphy D (2015). RNA binding protein Caprin-2 is a pivotal regulator of the central osmotic defense response. Elife, 4. doi:10.7554/eLife.09656

Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, … Wood KV (2008). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. ACS Chem Biol, 3(6), 373–382. doi:10.1021/cb800025k [PubMed: 18533659]

Papoulas O, Monzo KF, Cantin GT, Ruse C, Yates JR 3rd, Ryu YH, & Sisson JC (2010). dFMRP and Caprin, translational regulators of synaptic plasticity, control the cell cycle at the Drosophila mid-blastula transition. Development, 137(24), 4201–4209. doi:10.1242/dev.055046 [PubMed: 21068064]

Pardee AB (1989). G1 events and regulation of cell proliferation. Science, 246(4930), 603–608. [PubMed: 2683075]

Pauw RJ, van Drunen FJ, Collin RW, Huygen PL, Kremer H, & Cremers CW (2008). Audiometric characteristics of a Dutch family linked to DFNA15 with a novel mutation (p.L289F) in POU4F3. Arch Otolaryngol Head Neck Surg, 134(3), 294–300. doi:10.1001/archotol.134.3.294 [PubMed: 18347256]

Ratovitski T, Chighladze E, Arbez N, Boronina T, Herbrich S, Cole RN, & Ross CA (2012). Huntingtin protein interactions altered by polyglutamine expansion as determined by quantitative proteomic analysis. Cell Cycle, 11(10), 2006–2021. doi:10.4161/cc.20423 [PubMed: 22580459]

Shiina N, Shinkura K, & Tokunaga M (2005). A novel RNA-binding protein in neuronal RNA granules: regulatory machinery for local translation. J Neurosci, 25(17), 4420–4434. doi:10.1523/jneurosci.0382-05.2005 [PubMed: 15858068]

Shiina N, & Tokunaga M (2010). RNA granule protein 140 (RNG140), a paralog of RNG105 localized to distinct RNA granules in neuronal dendrites in the adult vertebrate brain. J Biol Chem, 285(31), 24260–24269. doi:10.1074/jbc.M110.108944 [PubMed: 20516077]

Shiina N, Yamaguchi K, & Tokunaga M (2010). RNG105 deficiency impairs the dendritic localization of mRNAs for Na+/K+ ATPase subunit isoforms and leads to the degeneration of neuronal networks. J Neurosci, 30(38), 12816–12830. doi:10.1523/jneurosci.6386-09.2010 [PubMed: 20861386]

Solomon S, Xu Y, Wang B, David MD, Schubert P, Kennedy D, & Schrader JW (2007). Distinct structural features of caprin-1 mediate its interaction with G3BP-1 and its induction of phosphorylation of eukaryotic translation initiation factor 2alpha, entry to cytoplasmic stress granules, and selective interaction with a subset of mRNAs. Mol Cell Biol, 27(6), 2324–2342. doi:10.1128/mcb.02300-06 [PubMed: 17210633]

Vognsen T, Moller IR, & Kristensen O (2013). Crystal structures of the human G3BP1 NTF2-like domain visualize FxFG Nup repeat specificity. PLoS One, 8(12), e80947. doi:10.1371/journal.pone.0080947 [PubMed: 24324649]

Wang B, David MD, & Schrader JW (2005). Absence of caprin-1 results in defects in cellular proliferation. J Immunol, 175(7), 4274–4282. [PubMed: 16177067]

Wu Y, Zhu J, Huang X, & Du Z (2015). Bacterial expression and preliminary crystallographic studies of a 149-residue fragment of human Caprin-1. Acta Crystallogr F Struct Biol Commun, 71(Pt 3), 324–329. doi:10.1107/s2053230×15002642 [PubMed: 25760709]

Wu Y, Zhu J, Huang X, & Du Z (2016). Crystal structure of a dimerization domain of human Caprin-1: insights into the assembly of an evolutionarily conserved ribonucleoprotein complex consisting of Caprin-1, FMRP and G3BP1. Acta Crystallogr D Struct Biol, 72(Pt 6), 718–727. doi:10.1107/s2059798316004903 [PubMed: 27303792]

Wu Y, Zhu J, Huang X, Zhou X, & Du Z (2018). Crystal structure of a dimerization domain of human Caprin-2: similar overall dimeric fold but different molecular surface properties to that of human Caprin-1. J Biomol Struct Dyn, 1–23. doi:10.1080/07391102.2018.1532817

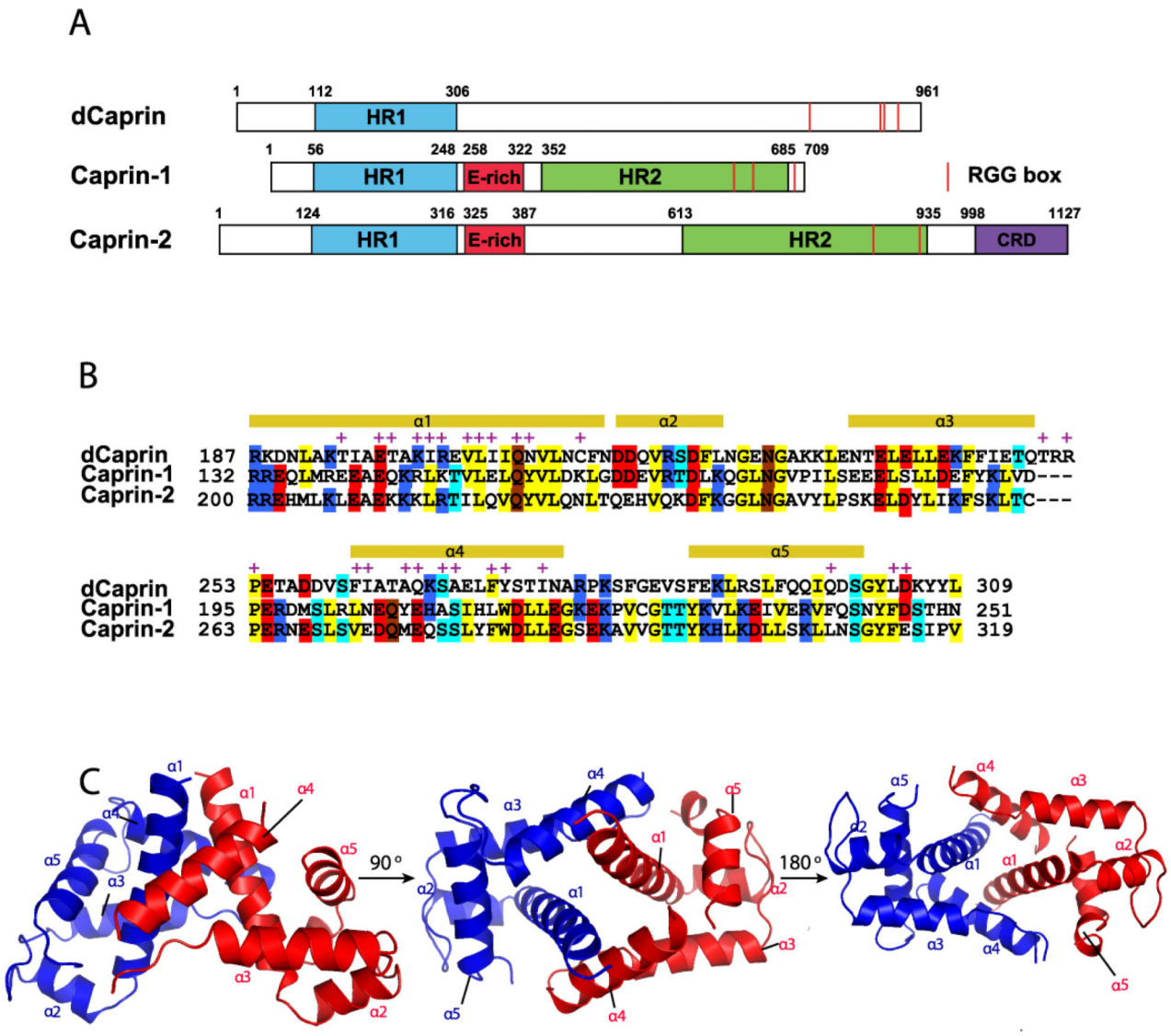

Figure 1. Caprin proteins and structure of the dimerization domain of dCaprin.

A) Schematic representation of dCaprin, and human Caprin-1 and Caprin-2. HR1 and HR2: homologous region 1 and 2. CRD: C1q-related domain. The thin red lines indicate the locations of the RGG boxes. B) Sequence of the dCaprin dimerization domain aligned with homologous sequences of human Caprin-1 and Caprin-2. The secondary structures are indicated above the dCaprin sequence. Residues involved in dCaprin dimerization are indicated with a "+" sign above the sequence. C) Structure of the dCaprin homodimer rendered in cartoon mode, viewed from three different angles. The two protomers are coloured red and blue respectively.

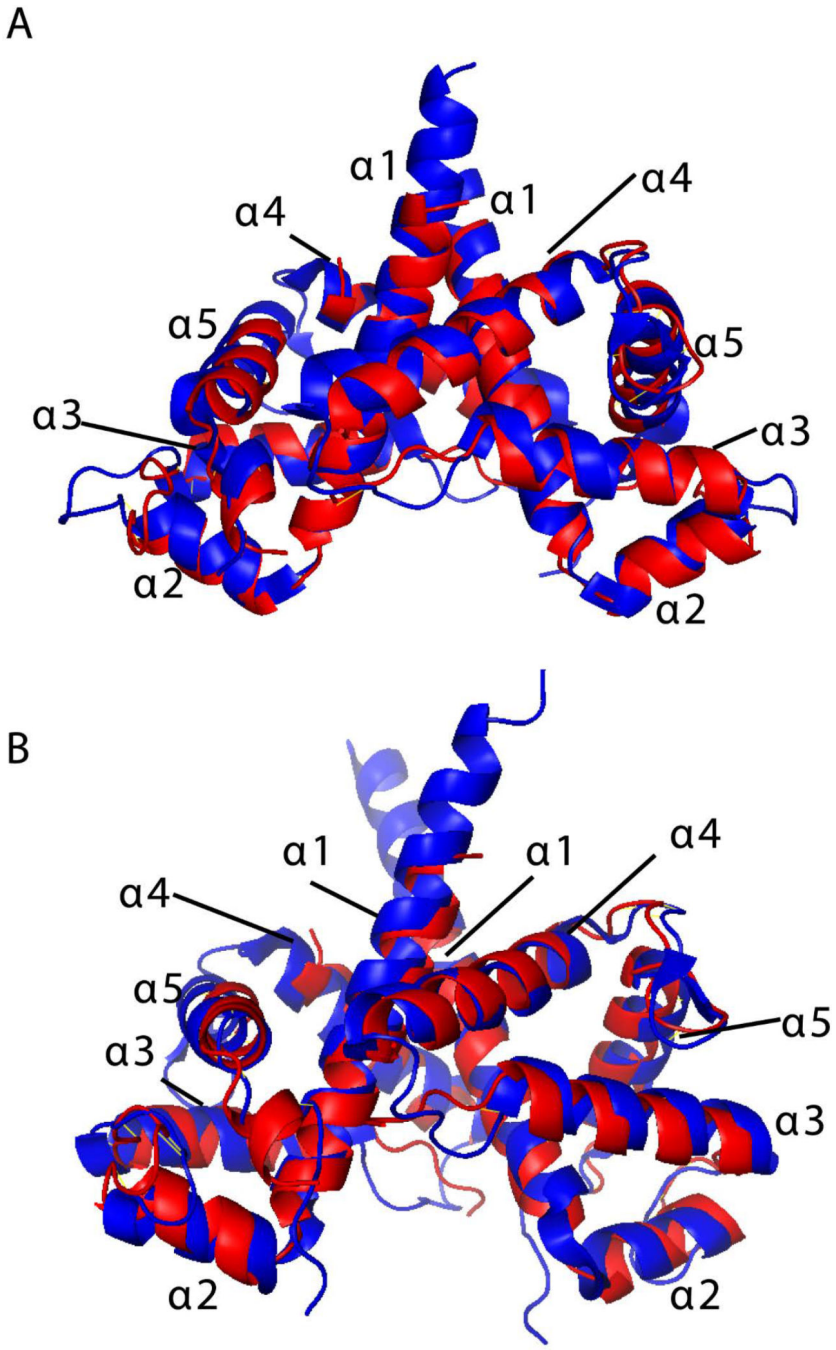

Figure 2. Superimposition of the caprin homodimeric structures.

A) Superimposition of the dCaprin (in red, PDB code 6BK4) and Caprin-1 (in blue, PDB code 4WBE) structures. B) Superimposition of the dCaprin (in red, PDB code 6BK4) and Caprin-2 (in blue, PDB code 6J97) structures.

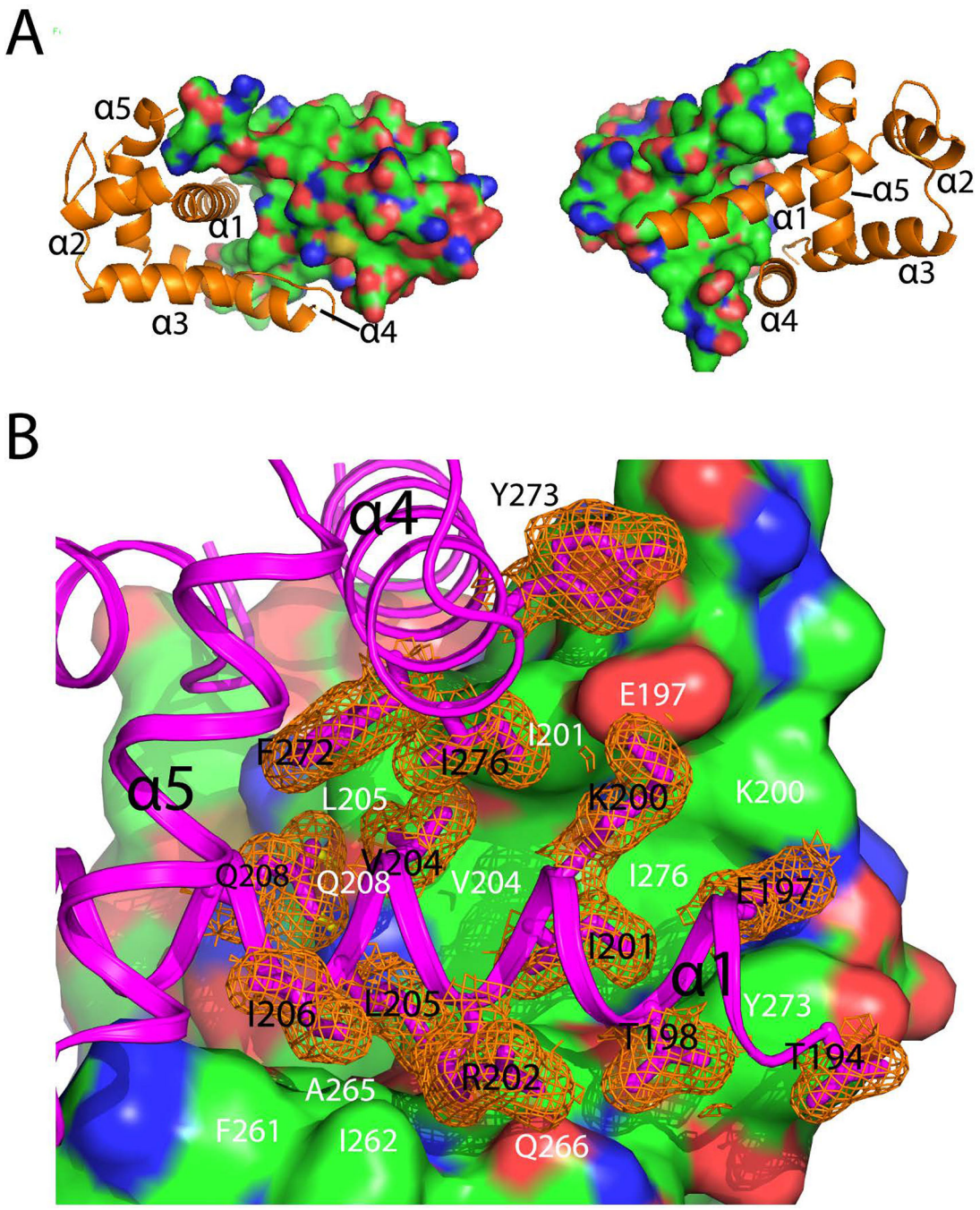

Figure 3. Molecular interactions that mediate the homodimerization of dCaprin.

A) Shape complementarity at the homodimerization interface. The two protomers are rendered in surface mode (colored by elements with C, N, O and S in green, blue, red and yellow respectively) and cartoon mode (colored in orange) respectively. B) Intermolecular hydrophobic contacts, salt bridges, and hydrogen bonds at the homodimerization interface. The side-chains for some of the key residues belonging to the protomer rendered in cartoon mode are shown in sticks. 2Fo-Fc electron density maps (mesh) contoured at 1σ are shown for these side-chains. The residues involved in the intermolecular interactions are indicated

differently for the two protomers: one letter abbreviation-residue number in black for residues belonging to the protomer rendered in cartoon mode and one letter abbreviation-residue number in white for residues belonging to the protomer rendered in surface mode. Hydrogen bonds are represented as yellow dashed lines.

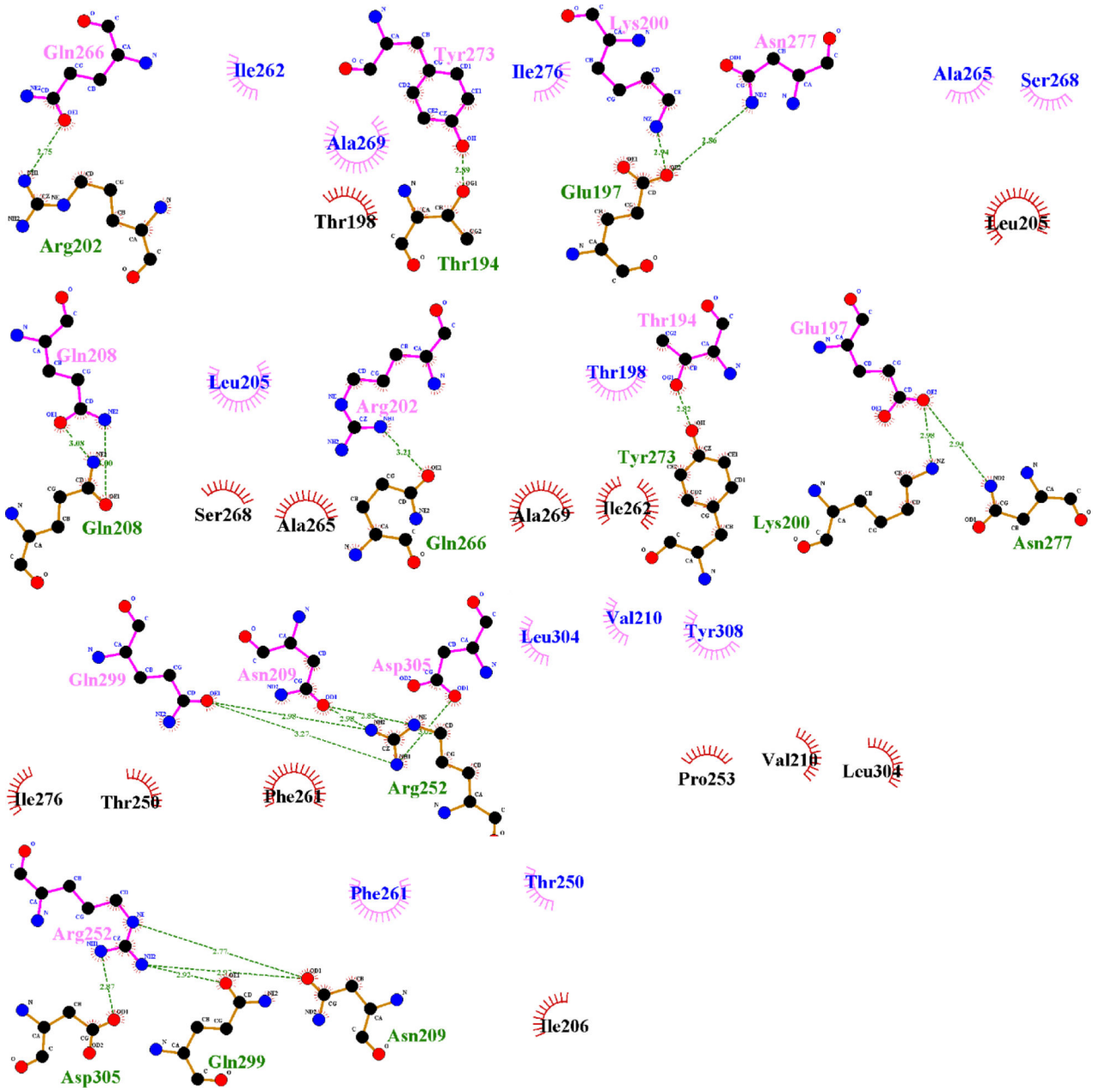

Figure 4. Ligplot presentation of molecular interactions in the dCaprin dimerization interface.
Hydrogen bonds are designated with green dashed lines (the distance between the two heavy atoms is indicated by a number in angstrom). Hydrophobic interactions are represented as starbursts.

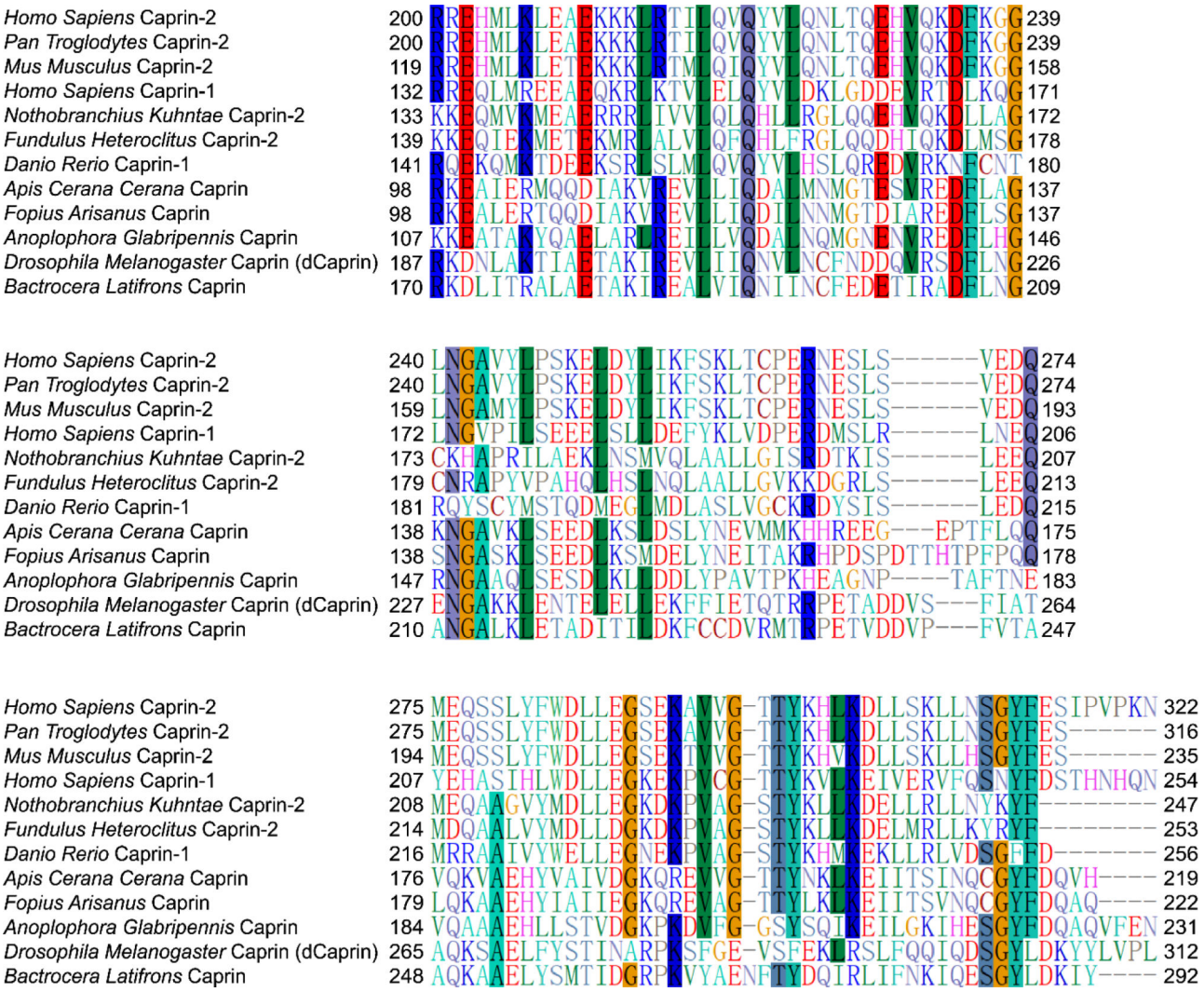

Figure 5. Alignment of the dimerization domain sequences from Caprin-1, Caprin-2, dCaprin, and corresponding sequences from diverse animal species.

Residues are color-coded based on properties. The highly conserved residues (65% or more) are highlighted by background coloring.

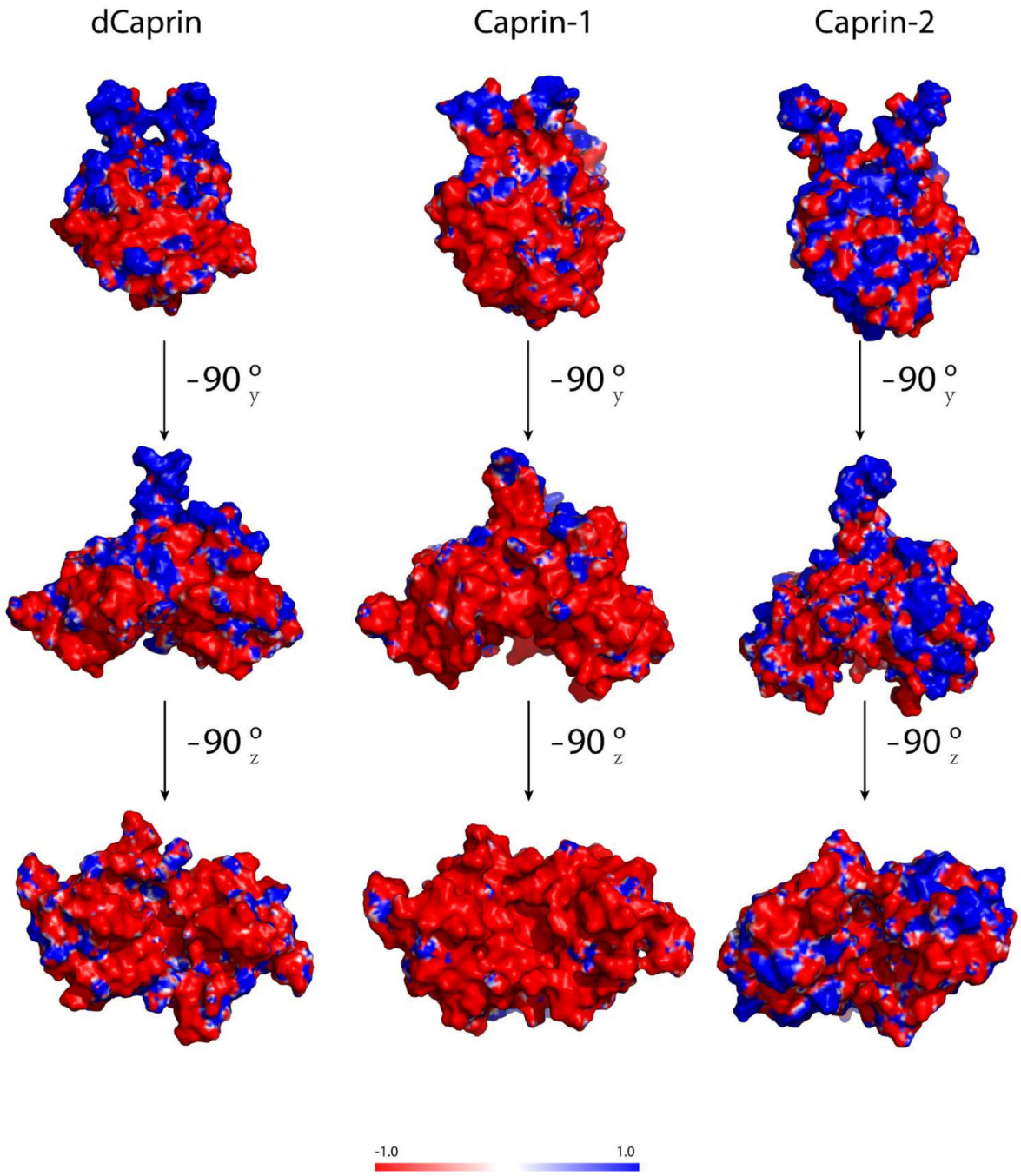

Figure 6. Electrostatic surface representation of the homodimeric structures of dCaprin, Caprin-1, and Caprin-2.

The structures from three different viewing angles are shown. Blue and red represent regions of positive and negative potentials respectively.

Table 1:

Data collection and processing statistics

	SeMet
Data Collection	
Beamline	ID-F, LS-CAT, APS
Wavelength (Å)	0.9787
Space group	$P2_12_12_1$
Cell dimensions	53.05
a,*b*,*c* (Å)	60.04
	99.40
α,*β*,*γ* (°)	90.0
	90.0
	90.0
Resolution (Å)	60.04–1.80
	(1.85–1.80)[a]
Total observations	153362
Unique reflections	30126
Completeness (%)[a]	99.8
Wilson B-factor (Å^2)	28.076
R_{merge} (%)[b]	6.1
I/σ(I)[a]	13.9
$CC_{1/2}$	0.997
Multiplicity	5.1
Refinement	
Resolution (Å)	51.39–1.80
No. reflections	3391
R_{work}/R_{free} (%)[c]	24.08/26.80
No. of atoms	
Protein	1714
Water	75
B factors	
Protein	37.7
Water	17.7
No. of protein residues	195
RMSD bonds (Å)	0.003
RMSD angles (°)	0.63
Ramachandran	
favored (%)	100
allowed (%)	0
outliers (%)	0

	SeMet
Clashscore	3.37
PDB code	6BK4

[a] Values in parentheses are for the highest-resolution shell

[b] $R_{merge} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - \langle I(hkl) \rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $\langle I(hkl) \rangle$ is the average intensity of reflection *hkl*.

[c] $R_{work} = \Sigma_{hkl} \big| |F_{obs}| - |F_{calc}| \big| / \Sigma_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors respectively. R_{free} is calculated as for R_{work} but only use a randomly selected subset of data (6%) which were excluded from refinement.