Protein J (2012) 31:275–284
DOI 10.1007/s10930-012-9401-1

Crystal Structure of *Human* ASB9-2 and Substrate-Recognition of CKB

Xiangwei Fei · Xing Gu · Shilong Fan · Zhenxing Yang · Fan Li · Cheng Zhang · Weimin Gong · Yumin Mao · Chaoneng Ji

Published online: 15 March 2012

Abstract *Human* ankyrin repeat and suppressor of cytokine signaling box protein 9 (hASB9) is a specific substrate-recognition subunit of an elongin C-cullin-SOCS box E3 ubiquitin ligase complex. It recognizes its substrate, brain type creatine kinase (CKB), using the ankyrin repeat domain; and facilitates the polyubiquitination of CKB to mediate proteasomal degradation through the SOCS box domain. HASB9-2 is an isoform of hASB9 that contains one ankyrin repeat domain. In this study, the crystal structure of hASB9-2 is shown at 2.2-Å resolution using molecular replacement. Overall, hASB9-2 forms a slightly curved arch with a characteristic L-shaped cross-section. Amino acid substitution analysis based on docking experiments revealed that His103 and Phe107 in hASB9-2 are essential for binding to CKB. Analysis of truncation mutants demonstrated that the first six ankyrin repeats along with the N-terminal region of hASB9-2 contribute to the interaction with CKB.

The crystal structure reported in this paper has been submitted to the PDB with ID code 3D9H.

Electronic supplementary material The online version of this article (doi:10.1007/s10930-012-9401-1) contains supplementary material, which is available to authorized users.

X. Fei · X. Gu · Z. Yang · F. Li · Y. Mao · C. Ji (✉)
State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai 200433, People's Republic of China
e-mail: chnji@fudan.edu.cn

S. Fan · C. Zhang · W. Gong
National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China

Keywords Ankyrin repeat and suppressor of cytokine signaling box protein 9 · Ankyrin repeat · Brain type creatine kinase · Crystal structure · Glutathione S-transferase pull-down

Abbreviations

hASB9	*Human* ankyrin repeat and suppressor of cytokine signaling box protein 9
ASBs	Ankyrin repeat and suppressor of cytokine signaling box proteins
SOCS	Suppressor of cytokine signaling
ECS	Elongin C-cullin-SOCS box
CKB	Brain type creatine kinase
CK	Creatine kinase
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
PDB	Protein Data Bank
GST	Glutathione S-transferase
PCR	Polymerase chain reaction
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline

1 Introduction

The ankyrin repeat and suppressor of cytokine signaling box proteins (ASBs) are composed of an N-terminal ankyrin repeat domain with different numbers of ankyrin repeats, and a C-terminal SOCS box domain [11, 12]. They are members of the suppressor of cytokine signaling (SOCS) box protein super-family. These ASB proteins recognize their specific substrate through ankyrin repeat domain. Meanwhile they can indirectly associate with a complex that contains elongin B, cullin-5 and Rbx1, via

binding with elongin C using SOCS box domain; and then constitute an elongin C-cullin-SOCS box (ECS) protein complex [3, 10]. The ECS complex is an E3 ubiquitin ligase, which along with an ubiquitin-activating enzyme (E1) and an ubiquitin-conjugating enzyme (E2), advantages the polyubiquitination of substrate and consequential proteasomal mediated degradation.

In humans, there are eighteen ASB proteins (ASB1 to ASB18). HASB9 is involved in substrate recognition in the ECS E3 ubiquitin ligase complex by interacting with brain type creatine kinase (CKB) through its ankyrin repeat domain [13]. The formation of hASB9/CKB heterodimer can increase polyubiquitination of cellular CKB, and facilitate its proteasomal mediated degradation, when the SOCS box domain of hASB9 is intact [4, 13].

HASB9-2, a truncated splice variant of hASB9, comprises its substrate-recognition domain that consists of six ankyrin repeats [13]. The 33-residue ankyrin repeat is a widespread motif that participates in protein interactions. This motif consists of two anti-parallel α-helices separated by loops. Neighboring repeats heap up continuously into a bundle, giving a steady surface for protein interactions [20, 21]. Several crystal structures of ankyrin repeat proteins have been determined, and the mechanism by which ankyrin repeats interact with their binding partners varies considerably. For example, 53BP2 interacts with p53 through the β-hairpin of the ankyrin repeat; IκBα interacts with p65 through the helices in the ankyrin groove; and INK4 interacts with CDK6 via the loops between the helix pairs [20]. In general, most binding partners interact with ankyrin repeats via the β-hairpin and/or the ankyrin groove surface [17]. The groove surface has the broadest range of curvature variation and appears to be capable of interacting with a wide variety of protein secondary structural elements, including loops (GABPb), β-sheets (INK4), helices (IκBα) and extended strands (D34); however, how ankyrin repeats of the hASB9-2 interact with CKB is unclear [16].

Creatine kinase (CK), an evolutionarily conserved enzyme, is an important cytosolic enzyme in cell energy metabolism [24]. It regulates the ATP level of the tissues that require a great quantity of energy, through reversibly catalyzes the phosphorylation of creatine, and thus is related to brain and muscle function [25]. Inhibition of CK activity and a disruption in creatine metabolism are found in a number of neurological and muscular diseases [4]. In addition, the CK system is correlated to tumor growth due to its role in regulating ATP production. Therefore, molecules that disrupt this system may have some effect upon tumor growth or progression [4].

CKB is the brain-specific cytosolic enzyme of CK. It is negatively regulated by the tumor suppressor gene p53 and positively regulated by the oncogene E1a [9, 26]. Many growth factors and hormones stimulate CKB activity and expression [23, 26]. CKB is over-expressed in a great variety of solid tumors and tumor cell lines and has been used as a prognostic marker of cancer and metastasis [25].

Here, we present the crystal structure of hASB9-2, an isoform of hASB9. The structure and glutathione S-transferase (GST) pull-down experiments reveal the mechanism by which the ankyrin repeats of hASB9-2 interact with CKB.

2 Materials and Methods

2.1 Protein Expression, Purification, Crystallization and Data Collection

The cloning, expression, purification, crystallization and data collection of hASB9-2 are performed as previously described [7].

2.2 Structure Solution and Refinement

The structure of hASB9-2 was obtained by molecular replacement using the program Phaser [22]. The truncated ankyrin repeats 1–6 of *human* ankyrinR structure (Protein Data Bank (PDB) code 1N11) [18], which has 31% identity to the target structure, was used as the search model. Refinement was carried out by using REFMAC to 2.2 Å [18]. The statistics of the structure refinement are summarized in Table 1. Stereo-chemical quality of the final model was determined by PROCHECK [15].

2.3 Protein Docking Analysis

GRAMM-X Protein Docking (http://vakser.bioinformatics.ku.edu/resources/gramm/grammx) was used to model the complex of hASB9-2 and *bovine* retinal creatine kinase (CK, PDB 1G0 W), which shares 97% sequence homology with CKB. To determine residues involved in the binding interface, we analyzed residues in hASB9-2 within 3 Å of CK.

2.4 Preparation of Expression Plasmids for Glutathione S-transferase (GST) Pull-Down Assay

A *Nhe* I site was added to the 5′ terminus and an *Xho* I site was added to the 3′ terminus of the CKB gene using polymerase chain reaction (PCR). The gene was inserted into the corresponding sites of a pET-28b vector (Novagen, USA), which contains a His tag. An *Eco*R I site was added to the 5′ terminus and an *Xho* I site was added to the 3′ terminus of the CKB (as well as hASB9-2) gene using PCR. The gene was inserted into the corresponding sites of a pGEX-4T-1 vector (Novagen, USA), which contains a GST tag.

Table 1 Statistics for data collection and refinement

Protein name	hASB9-2
Space group	$P4_332$
Cell parameters, Å	a = 129.25, b = 129.25
	c = 129.25
	$\alpha = \beta = \gamma = 90°$
Resolution range, Å	30.0–2.2
Observed reflections	Total: 709,934
	Unique: 18,315
Completeness (%)[a]	99.94 (99.94)
Redundancy	13.4 (13.5)
R_{merge} (%)[b]	6.1 (48.9)
I/σ (I)	44.3 (6.5)
R_{factor}/R_{free}, %[c]	19.6/22.0
Mosaicity	0.266 (0.266)
R.m.s.d. from ideal geometry:[d]	
Bond lengths, Å	0.015
Bond angles	1.369
Average B factor, Å2	38.45
Ramachandran plot	
Most favored regions (%)	88.1
Additionally allowed (%)	11.9

[a] Data for high-resolution bins are in parentheses

[b] $R_{merge} = \Sigma|I_i - I_m|/\Sigma I_i$, where I_i is the intensity of the measured reflection and I_m is the mean intensity of all symmetry-related reflections

[c] $R_{free} = \Sigma_{SUB}||F_{obs}| \_ |F_{calc}||/\Sigma_T|F_{obs}|$, where T is a test data set of about 10% of the total reflections randomly chosen and set aside prior to refinement

[d] R.m.s.d., root-mean-square deviations

Point mutants of hASB9-2 at residues Ile45 (I45E, I45R), Cys77 (C77A), Leu78 (L78E, L78R), Gly79 (G79E, G79R), Trp102 (W102E, W102R), His103 (His103), Thr104 (T104E, T104R), Leu106 (L106E, L106R), Phe107 (F107E, F107R), Cys110 (Cys110) and Val111 (V111E, V111R) as well as three double mutations I45E/V111E, I45E/H103V and V111E/H103V were introduced by PCR-based site-directed mutagenesis using an appropriate set of primers (shown in the supplementary Table 1). PCR products were subcloned into a pET-28b vector, resulting in His-tagged fusion proteins. All mutants were confirmed by sequencing analysis.

The following ankyrin repeat truncated mutants of hASB9-2 (Fig. 4d): ankyrin repeat 1–2 (1–2), ankyrin repeat 1–3 (1–3), ankyrin repeat 1–4 (1–4), ankyrin repeat 1–5 (1–5), ankyrin repeat 1–6 (1–6), ankyrin repeat 1-C-terminus (1-C), N-terminus-ankyrin repeat 2 (N-2), N-terminus-ankyrin repeat 3 (N-3), N-terminus-ankyrin repeat 4 (N-4), N-terminus-ankyrin repeat 5 (N-5), N-terminus-ankyrin repeat 6 (N-6), ankyrin repeat 3 (3), ankyrin repeat 2–3 (2–3) and ankyrin repeat 2–4 (2–4), were produced from the parental plasmids using an appropriate set of primers (shown in the supplementary Table 2). DNA fragments were subcloned into a pGEX-4T-1 vector in frame with the GST sequence. All mutants were confirmed by sequencing analysis.

2.5 Expression and Purification of Fusion Proteins

His-tagged proteins (hASB9-2 mutants as well as wild-type CKB) were essentially prepared as described [7]. GST-tagged proteins (truncated mutants and wild-type hASB9-2, as well as wild-type CKB) were expressed in *Escherichia coli* Rosetta (DE3) pLyS cells and performed as recommended by the manufacturer. The purified proteins were authenticated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and quantified by the Bradford method (Pierce) with *bovine* serum albumin (BSA) as a standard.

2.6 GST Pull-Down Assays

To test possible binding sites on hASB9-2, full-length GST-CKB in bacterial cell lysates was bound to 1 mL of Glutathione Sepharose 4B (Amersham Biosciences) beads. The beads were rotated for 2 h at 16 °C and washed four times with 1× phosphate buffered saline (PBS) lysis buffer (containing 137 mM NaCl, 2.7 mM KCl, 10 mM sodium phosphate dibasic, 2 mM potassium phosphate monobasic and a pH of 7.4). Beads containing 30 μg of protein were incubated with 15 μg of purified His-tagged hASB9-2 (wild-type and mutants) and rotated for 4 h at 4 °C. After washing the beads eight times with 1× PBS lysis buffer, the proteins were eluted by boiling in 1× SDS-PAGE sample buffer. The eluates were separated by SDS-PAGE on a 12% gel and analyzed using western blotting with an anti-His antibody (Proteintech Group). The presence of GST fusion proteins were verified with an anti-GST antibody (Proteintech Group).

To identify which region of hASB9-2 was essential for binding to CKB, 15 μg of GST-fusion fragments of hASB9-2 were immobilized onto Glutathione-Sepharose 4B beads and incubated with 15 μg of CKB for 4 h at 16 °C in 1× PBS lysis buffer. The beads were lightly washed eight times with 1× PBS lysis buffer. The bound proteins were eluted with 1× SDS sample buffer and separated by SDS-PAGE on a 12% gel. In some experiments, the proteins were transferred to a nitrocellulose membrane, probed with an anti-His antibody or anti-GST antibody and detected by enhanced chemiluminescence (ECL; Pierce).

3 Results

3.1 Structure Overview

The crystal structure of hASB9-2 was resolved at 2.2-Å resolution using molecular replacement. The final refined

model consisted of residues 19–253 of the hASB9-2 with no observable electron density for residues 1–18 and 254–262. The N-terminal stretch of residues Phe19 to Ser26 in two crystallographic symmetry-related molecules were related by a crystallographic twofold axis and their densities overlapped. These data suggested that the high symmetry space group made crystallographic intermolecular contacts through the N-terminal residues.

As shown in Fig. 1a, the overall architecture of hASB9-2 consists of an elongated array of anti-parallel α-helices, consistent with an ankyrin repeat structure. It consists of fourteen α-helices, which compose six previously identified ankyrin repeats (α1–α12, residues 35–227) and two additional anti-parallel α-helices (α13 and α14) at the C-terminus (residues 228–253) [13]. In each ankyrin repeat, the helices are grouped in anti-parallel mode, and then connected with a loop region that is at an angle of approximately 90° to the helices, forming a characteristic L-shaped cross-section. Moreover, the C-terminal region clearly adopts this ankyrin fold. The overall stack exhibits a slightly curved arch structure, which appears to be the result of the different helix lengths and the inter-helix packing interactions between ankyrin repeats, producing a groove along the long axis of the protein.

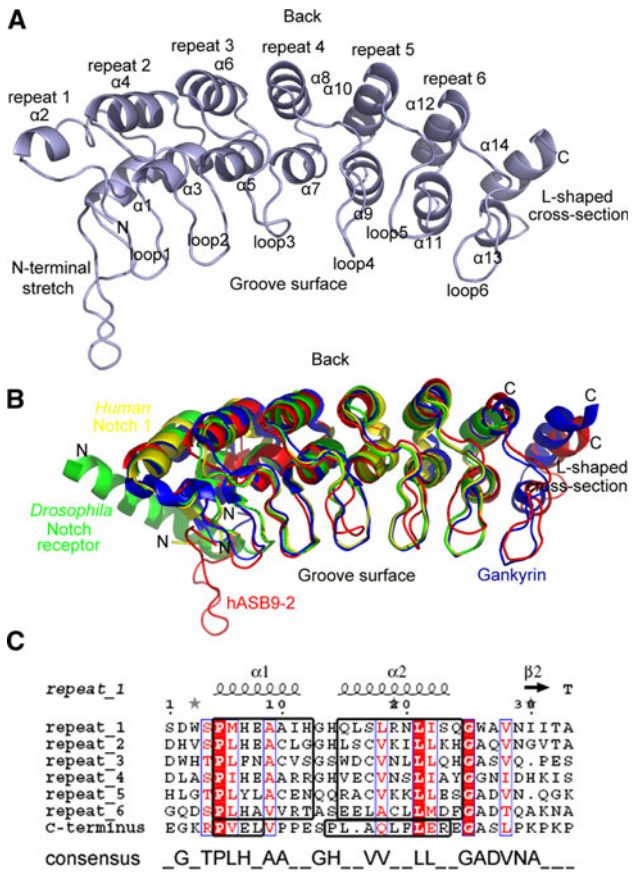

Fig. 1 Structure of hASB9-2. **a** Cartoon representation of the hASB9-2 structure. This figure is generated using PyMOL. **b** Cartoon diagram showing the superimposition of hASB9-2 (*red*, PDB code 3D9H) with ankyrin repeat domains from *human* gankyrin (*blue*, 1QYM), *human* Notch 1 (*yellow*, 1YYH) and *Drosophila* Notch receptor (chain A, *green*, PDB 1OT8). The structures are superposed primarily using the Cα atoms from the core ankyrin repeat domain. This figure is generated using PyMOL. **c** Sequence alignment by Clustal W [19] of the six ankyrin repeats and the C-terminus of hASB9-2. The secondary structure of ankyrin repeat 1 is indicated *above the alignment*. *White characters on a red background* represent strictly conserved residues, while well-conserved residues are drawn in *red* and framed in *blue*. The ankyrin repeat consensus sequence is showed *below* [5, 14, 16, 20]. Two α-helices of ankyrin repeats are marked in *black rectangles*. This figure is prepared with ESPript [8] (Color figure online)

3.2 Molecular Modeling

To determine how hASB9-2 interacts with its substrate CKB, we modeled the interaction of hASB9-2 and CK, which shares 97% sequence homology with CKB, using GRAMM-X Protein Docking. To some extent this was not the optimal way. The docking might be constrained, for there had no evidence that hASB9-2 interacted with CK in solution, the CKB and CK structures probably had some partly differences, and the putative docked complexes might not fit in with the real binding even if they could be proved between hASB9-2 and CK. But during the docking analysis, *human* CKB structure has not been deposited in PDB, so we used CK instead of CKB. It helped us to initiate our study on the binding modes between hASB9-2 and CKB. Consequently, we got ten docked models (Fig. 2). The result stated that different regions of hASB9-2 were used as the binding interface to contact with different complementary patches on CK. To verify which model was credible, experimental studies on certain residues were in urgent needs.

Therefore, we analyzed residues in all the ten models as follows:

First, we analyzed the residues in hASB9-2 involved in the binding interface; residues within 3 Å of CK were considered important for binding. These residues are listed in Table 2 for the ten models. The residues involved in hASB9-2 within 3 Å of CK primarily localized to the ankyrin repeats with the exception of the C-terminal ones in seven of the models. In the other three models, the binding interface was somewhat different. In model 6, residues involved in the interaction localized from ankyrin repeat 3 to the C-terminus of hASB9-2. In model 3 and 7, residues involved in the interaction were observed throughout the 3D structure of hASB9-2.

Then, we conducted a statistical analysis of these residues. Residues Leu78 and Gly79 appeared most frequently (Leu78 in model 3, 4, 5, 7 and 9 and Gly79 in model 3, 4, 7, 9 and 10); therefore, we conclude that hydrophobic residues Leu78 and Gly79 are probably involved in the

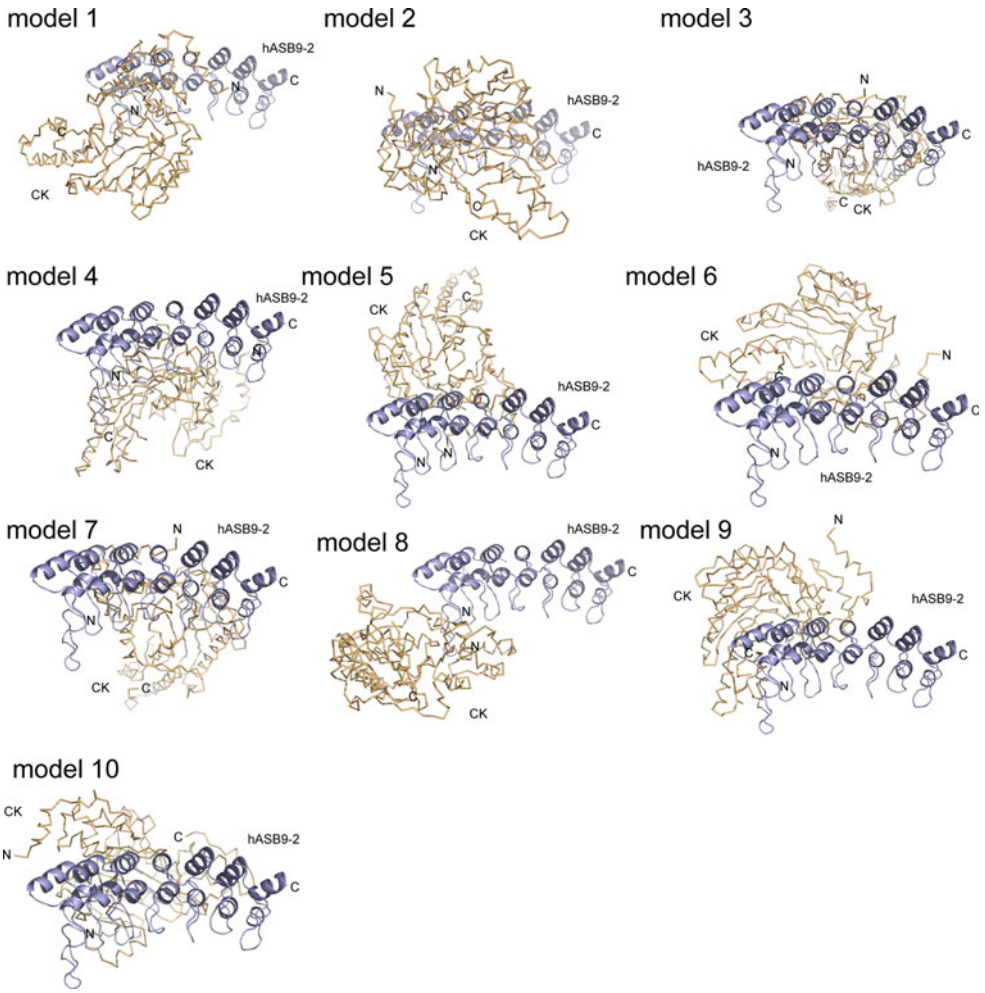

Fig. 2 Representation of ten docking models. HASB9-2 is shown in cartoon representation in *light blue*, whereas CK is shown in ribbon representation in *light orange*. This figure is performed using PyMOL (Color figure online)

interaction, and therefore the binding might be hydrophobic. Consequently, we carried out further investigations focused on Leu78, Gly79 and residues nearby.

3.3 Site-directed Mutagenesis Experiments of hASB9-2

In the hASB9-2 structure, hydrophobic residues Leu78 and Gly79 are located on the groove surface of α3 (ankyrin repeat 2). We performed a sequence alignment of ankyrin repeat domains from *human* ASB family members (Fig. 3), and compared residues located surround Leu78 and Gly79. We got two poorly conserved cysteines (residues 77 and 110), two highly conserved hydrophobic residues (Thr104 and Leu106) and four specific hydrophobic residues (Ile45, Trp102, Phe107 and Val111). However, many polar residues are unique to hASB9-2, such as His103. We hypothesized that these residues might participate in the interaction.

To determine whether these residues were essential for the interaction between hASB9-2 and CKB, we analyzed nineteen single mutations and three double mutations of hASB9-2. To interrupt the supposed hydrophobic interaction, we mutated all hydrophobic residues (Ile45, Leu78, Gly79, Trp102, Thr104, Leu106, Phe107 and Val111) to positively charged (arginine) and negatively charged (glutamic acid) residues. Similarly, polar residue His103 was mutated to a hydrophobic residue (valine), and Cys77 and Cys110 were mutated to a small, hydrophobic residue (alanine). Finally, we made three double mutations (I45E/V111E, I45E/H103V and V111E/H103V). These residues are at the edge of the tested region and therefore a single mutation at these sites may not be expected to dramatically affect binding.

His-tagged hASB9-2 fusion genes were transformed into Rosetta (DE3) *E. coli* cells for expression. Mutants T104E, T104R, L106E and L106R localized to inclusion bodies, which consist of insoluble aggregates of the expressed protein. Therefore, conserved residues Thr104 and Leu106 were likely important for the proper folding of hASB9-2. The other mutants and wild-type hASB9-2 were further investigated using GST pull-down assays with GST-CKB or GST as a control.

I45R, L78R, G79E, F107E, H103V and I45E/V111E mutants abolished the interaction between the hASB9-2 and CKB, whereas F107R, C110A, V111E, I45E/H103V

Table 2 The binding sites determined from models using GRAMM-X protein docking analysis

Model	Binding sites (corresponding region in hASB9-2)
Model 1	Phe19, Pro20, Gly21, Ala61, Asn63, His91, Gly92, Gln94, Val95, Asn96, Gly97, Gly125, Glu131 (N-terminal to ankyrin repeat 3)
Model 2	Phe19, Pro20, Ile22, Ile56, Gly59, His91, Gln94, Leu122, Gln123, Ser127, Tyr156 (N-terminal to ankyrin repeat 4)
Model 3	His46, **Leu78**, **Gly79**[a], Gly80, His81, Leu82, Ser112, Gly113, Trp115, Arg144, His146, Val147, Gln179, Ala181, Ala210, Leu238, Val239, Pro240 (ankyrin repeat 1 to C-terminal)
Model 4	Ser26, Leu29, Gly31, Asp32, Ile45[b], **Leu78**, **Gly79**, Asp101, His103, Leu134, His166, Lys197 (N-terminal to ankyrin repeat 5)
Model 5	Ile45, Leu50, Glu75, **Leu78**, Ser114, Trp115, Arg143, Arg144, His146, Asn177, Gln179, Ala181 (ankyrin repeat 1–5)
Model 6	Val111, Ser112, Arg144, His146, Glu176, Gln178, Gln179, Ala181, Thr209, Ala210, Glu212, Pro244 (ankyrin repeat 3 to C-terminal)
Model 7	Glu42, Ile45, His46, **Leu78**, **Gly79**, Gly80, Val111, Ser112, Arg143, Arg144, Asn177, Gln179, Ala181, Ala210, Leu238, Pro240 (ankyrin repeat 1 to C-terminal)
Model 8	Phe19, Pro20, Gly21, Ile22, Arg23, Leu25, Ser26, Pro28, Leu29, Asp68 (N-terminal and ankyrin repeat 1)
Model 9	Asp32, Pro39, Ile45, His46, His48, Ser51, Arg53, Leu55, Ser57, Trp60, **Leu78**, **Gly79**, Leu82, Ser83, Ser112, Trp115, Asp116 (ankyrin repeat 1–3)
Model 10	His46, **Gly79**, Gly80, His81, Gly113, Ser114, Trp115, Asp116, Gly145, His146, Val147 (ankyrin repeat 1–4)

[a] Leu78 and Gly79 are marked in bold

[b] Other residues analyzed for site-directed mutagenesis are underlined

and V111E/H103V weakly interacted with CKB (Fig. 4a–c). V111R was presumed to strengthen the interaction, considering the loading of V111R was much lower than wild-type hASB9-2 and the other mutants (Fig. 4c).

Mutations of Phe107 to either a positive or a negative residue abolished or decreased the binding between hASB9-2 and CKB. H103V, I45E/H103V and V111E/H103V showed weak binding to CKB. Based on these results, Phe107 and His103 in ankyrin repeat 3 might be essential but not sufficient to the interaction. Furthermore, His103 is in a connecting loop between the helix pairs, and Phe107 is on the groove surface. Therefore, we presume that the groove surface and the connecting loops between the helix pairs of hASB9-2 compose the suitable binding interface for CKB recognition.

3.4 Truncated Mutagenesis Experiments of hASB9-2

Ankyrin repeat is a common protein/protein interaction motif with well-defined structure. We were eager to know whether all ankyrin repeats facilitated the binding to CKB. To further investigate the sufficient binding interface of hASB9-2, we prepared a series of truncated mutants with different number of ankyrin repeats. Based on ankyrin repeat 3, the truncated mutants were as follows: 1–2, 1–3, 1–4, 1–5, 1–6, 1-C, N-2, N-3, N-4, N-5, N-6, 3, 2–3 and 2–4 (Fig. 4d).

The ability of these GST-tagged truncated mutants to bind to CKB was investigated using a GST pull-down assay. The N-6 truncated mutant retained binding, the N-4 truncated mutant detected very weak fluorescence and the other truncated mutants abolished binding, while the wild type hASB9-2 was shown separately (Fig. 4e). The detected weak fluorescence of N-4 truncated mutant is thought to be inconformity with the full-length hASB9-2, because there have some tertiary structure changes we have not tested, or there have some discrepancy between the real binding state and this result. So, it seemed like the only construct that bound besides wild type is N-6. Therefore, both ankyrin repeats 1–6 and the N-terminus of hASB9-2 were essential for binding to CKB.

We conclude that the groove surface of ankyrin repeats 1–6 and the N-terminus of hASB9-2 compose the interface with CKB. Furthermore, His103 and Phe107 are essential but not sufficient for binding and are likely part of the binding sites.

4 Discussions

4.1 α13 and α14 Composed the Seventh Ankyrin Repeat

The overall structure of hASB9-2 showed that the C-terminal anti-parallel α-helix pair (α13 and α14) was ankyrin repeat-like. To analyze whether it was indeed the seventh ankyrin repeat, we performed a structural alignment of hASB9-2 with the ankyrin repeat domain of *Human* gankyrin [14], *Human* Notch 1 [6] and *Drosophila* Notch receptor [29], which are composed of seven ankyrin repeats. Superposition of the four structures produced an RMSD of 1.3 Å over 138 Cα atoms between hASB9-2 and *human* gankyrin, 2.0 Å over 133 Cα atoms between hASB9-2 and *human* Notch 1, and 1.5 Å over 139 Cα

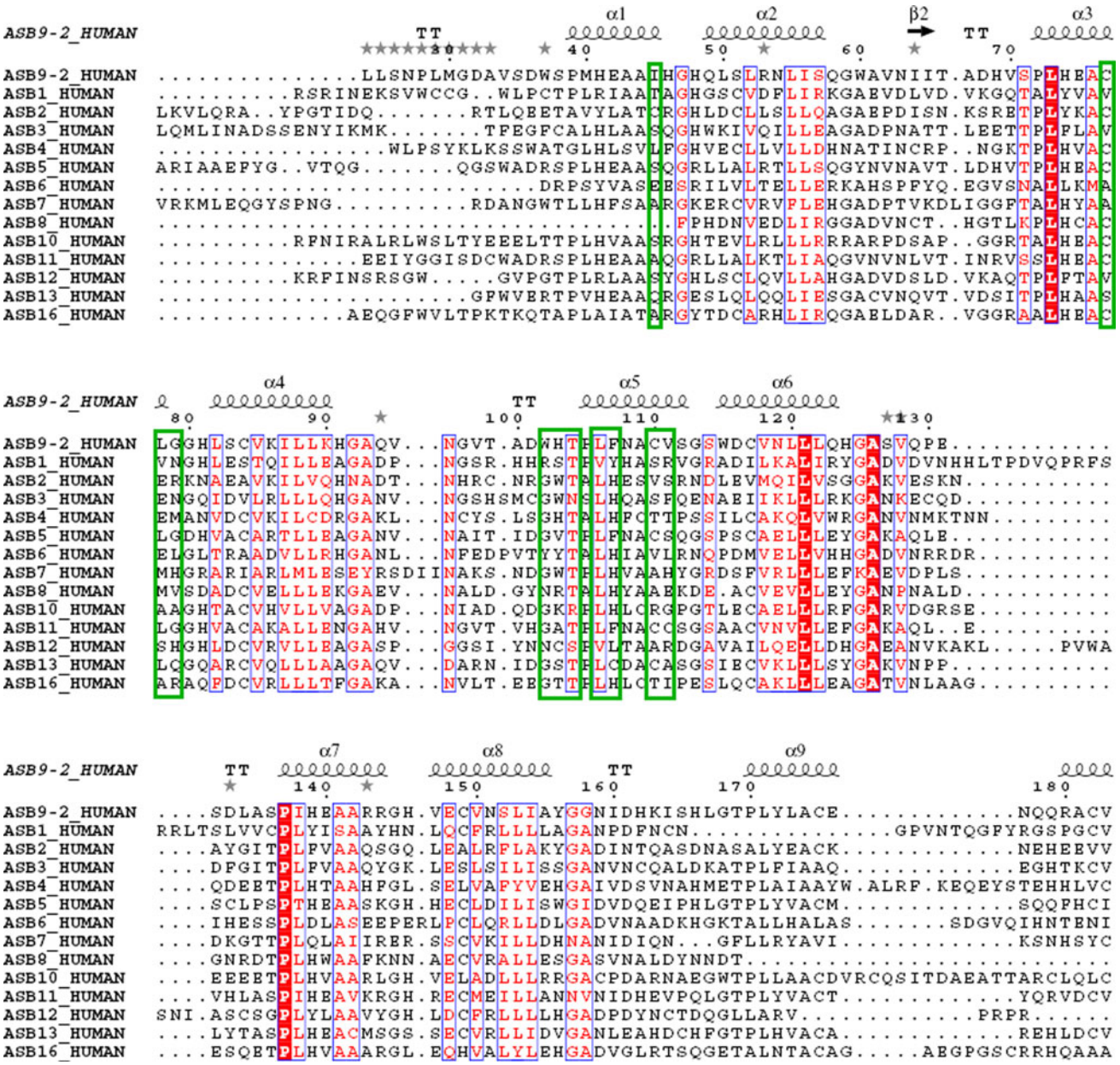

Fig. 3 Sequence alignment of the ankyrin repeat domains from *human* ASB family based on the structure of hASB9-2. α-helices are shown *above* the sequences as determined using Clustal W [19]. Absolutely conserved residues are colored *red* and positions with >50% conservation are marked in *red* and framed in *blue*. The numbering refers to the ankyrin repeat domain construct studied here; residue 40 corresponds to residue 22 of the full length hASB9-2. This figure was made using the program ESPript [8]. In this figure, residues analyzed for site-directed mutagenesis were marked in *green squares*. Only the result of the N-terminus and ankyrin repeat 1–4 portion was exhibited here; the whole one was provided in the supplementary Fig. 1. GenBank accession numbers are as follows. *Human*-ASB9-2 (NM_024087), *human*-ASB1 (NM_001040445), *human*-ASB2 (NM_016150), *human*-ASB3 (NM_016115), *human*-ASB4 (NM_016116), *human*-ASB5 (NM_080874), *human*-ASB6 (NM_017873), *human*-ASB7 (NM_024708), *human*-ASB8 (NM_024095), *human*-ASB10 (NM_080871), *human*-ASB11 (NM_080873), *human*-ASB12 (NM_130388), *human*-ASB13 (NM_024701), *human*-ASB16 (NM_080863) (Color figure online)

atoms between hASB9-2 and *Drosophila* Notch receptor. The result revealed that the overall structure of hASB9-2 was highly homologous with ankyrin repeat domains from these proteins (Fig. 1b). Thus, at the structural level, the ankyrin repeat probably pervades hASB9-2 in its entirety.

Many "signature" residues are included in ankyrin repeat that compose its consensus sequence and define its shape [5, 14]. Among the ankyrin consensus sequence, Thr4, Pro5, Gly13, Gly25 and Ala26 determine the boundaries of the helices; Ala9 and Ala30 help orient the

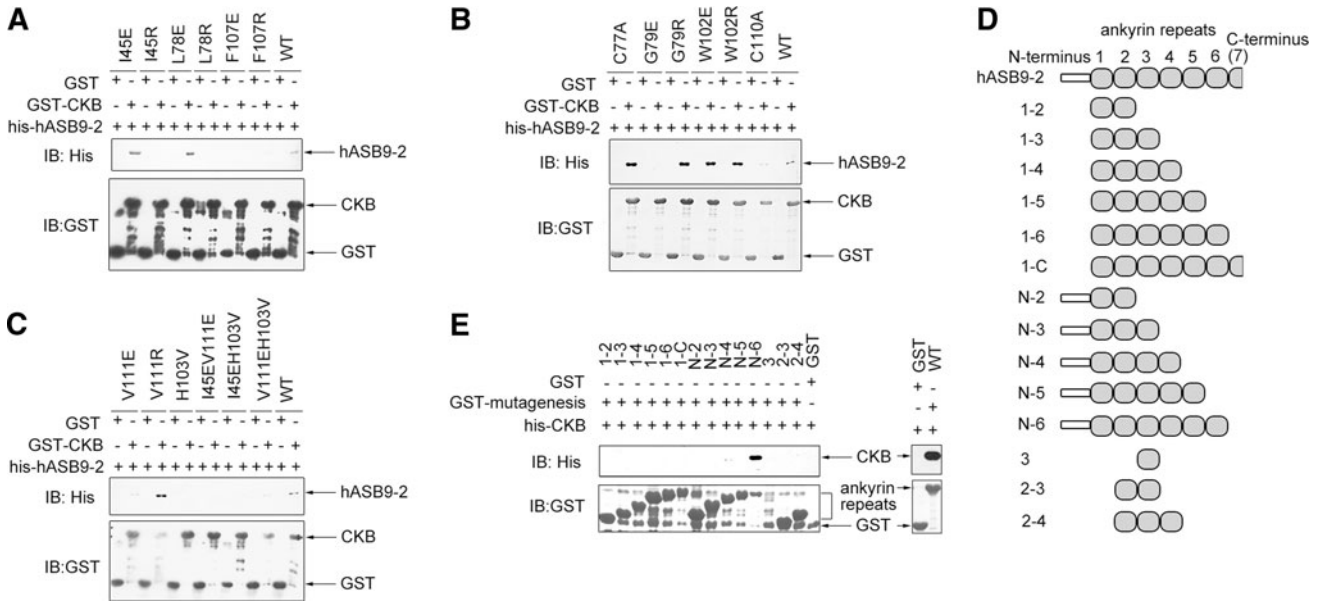

Fig. 4 Interaction between CKB and wild-type, mutated or truncated hASB9-2 using GST pull-down assays. **a–c** Wild-type and hASB9-2 mutants were incubated with GST-CKB(−) or GST-CKB(+) proteins immobilized on Sepharose beads. Associated proteins were detected with an anti-His antibody. DAB tetrahydrochloride staining with an anti-GST antibody demonstrated the amounts of GST and GST-CKB fusion proteins. **d** Schematic of truncated hASB9-2. **e** Purified His-CKB was mixed with GST-tagged wild-type or truncated hASB9-2 immobilized on Glutathione-Sepharose beads; the coprecipitated CKB was separated by SDS-PAGE and detected using western blotting with an anti-his antibody. GST was used as a negative control

two helices relative to each other; while Ala10 and valine at positions 17 and 18 control the curvature of the repeat stack [5, 16, 20]. To further confirm our inference, we made a sequence alignment of the C-terminus with six previously identified ankyrin repeats in hASB9-2. It exhibited low similarity to the ankyrin repeat consensus sequence (Fig. 1c). Pro5, Leu21 and Gly25 were strictly conserved among all seven anti-parallel α-helix pairs; residues 4, 6, 9, 18, 22 and 28 were well-conserved, whereas others were less-conserved even among ankyrin repeats 1–6. Furthermore, the length of α13 and α14 also differed from that of the helices in ankyrin repeats 1–6 (Fig. 1c). Each inner α-helix in the six ankyrin repeats in hASB9-2 was eight residues long (positions 5–12), whereas the outer α-helices were ten residues long (positions 15–24). However, α13 was only four residues long (Pro235-Leu238) whereas α14 was nine residues long (Pro244-Arg252). Previous studies have demonstrated that even though the terminal ankyrin repeats adopt the ankyrin fold, they often deviate from the established consensus sequence and are frequently truncated [17]. For example, the C-terminal ankyrin repeats from *Human* Notch 1, gankyrin and *Drosophila* Notch receptor significantly deviate these properties [6, 14, 27, 29]. Furthermore, a set of investigations on *Drosophila* Notch receptor have demonstrated that the capping "seventh" repeat is important for structural integrity and stability [27, 28]. The sequence differences may be a result of its role as a "capping" ankyrin repeat which interacts with solvent on one face [28]. Therefore, α13 and α14 might form a putative ankyrin repeat.

In contrast, the N-terminal ankyrin repeat from *Human* Notch 1 and *Drosophila* Notch receptor do not adopt a regular ankyrin fold (Fig. 1b). The repeats are disordered, when the sequences are conformed to the ankyrin repeat consensus sequence [6, 27, 29].

Say on the whole, the terminal ankyrin elements are able to tolerate sequence and structure variations, because they use only one face to contact with their neighbors, when the opposite face is exposed. Therefore we predict that α13 and α14 form the seventh ankyrin repeat of hASB9-2.

4.2 The Possible Binding Interface on hASB9-2

The site-directed and truncated mutagenesis experiments of hASB9-2 reveals that the groove surface of ankyrin repeats 1–6 and the N-terminus of hASB9-2 compose the interface with CKB. Among them, Phe107 and His103 compose a hot spot for binding and are likely part of the binding site.

The groove surface seems credible to be the specific binding interface for several reasons. First, previous studies have indicated that most binding partners interact with ankyrin repeats between the loop and/or the ankyrin groove surface [17, 20]. Second, the groove surface contains many specific residues within the *human* ASB family, including Phe107 (Fig. 3). They are more prone to give specificity to the recognition. Finally, this surface has the broadest range

of curvature variation (Fig. 1a). Meanwhile, ankyrin repeat stacks of different proteins have unique curvatures that adapt to associate best with specific substrate [16].

We infer that the curvature of hASB9-2 is necessary to substrate recognition. Previous studies have indicated that the geometric parameters of the curvature are determined by the relative side chain volumes of residues at positions 10, 17 and 18 of adjacent repeats [16]. Contrasting the three residues in each ankyrin repeat of hASB9-2 with the ankyrin consensus sequence, we noticed that amino acid substitutions were observed at all three residues in ankyrin repeats 6 and 7, positions 17 in ankyrin repeat 4, and two residues in the other repeats (Fig. 1c). These contributed to the unique curvature of hASB9-2, thereby conduced to its specific substrate recognition.

Furthermore, the N-terminal loop also has some effect. First, the N-terminal residues possess high specificity within *human* ASB family (Fig. 3). Second, crystallographic intermolecular contacts were observed through the N-terminal residues. It implies that the structure of N-terminal loop may be different between the substrate bound state and monomeric state in solution. Thus, it probably acts as a structural anchor.

Thence, we analyzed the contacted surface of hASB9-2 in all ten docking models, to find out which putative docked complex have the highest similarity with the real binding state in solutions. We obtained the model as follows. First, as listed in Table 2, residues on the N-terminal stretch of hASB9-2 in models 1, 2, 4 and 8 were involved in the interaction. So we picked out these models. Then, the groove surface of hASB9-2 in model 4 composed part of the binding surface. In models 1 and 2, hASB9-2 used loops between adjacent repeats, while only the N-terminus contacted with CK in model 8. So model 4 was favored. Finally, His103 in model 4 was involved in the binding sites. Therefore, we predict that hASB9-2 interacts with CKB in a similar manner as model 4 from the docking results (Fig. 5).

To perfectly procure our conclusion, we would better validate both the site-directed and truncated mutants for folding and stability, as well as the binding affinity. In this manuscript, we authenticated the binding of hASB9-2 mutants to CKB using GST pull-down assay, for we thought this experiment could differentiate whether the mutants could bind to CKB or not, even without the value of affinity. We thought the false folding mutants would result in the formation of inclusion bodies, which consisted of insoluble aggregates of the expressed protein. For example Thr104 and Lue106 mutants localized to inclusion bodies. So in the GST pull-down assay, we supposed the tertiary structures of mutants have not changed a lot. In addition, similar ankyrin repeat truncations were performed to study the binding of asb3 to its substrate [2]; moreover, whole-repeat truncations as well as point substitutions were

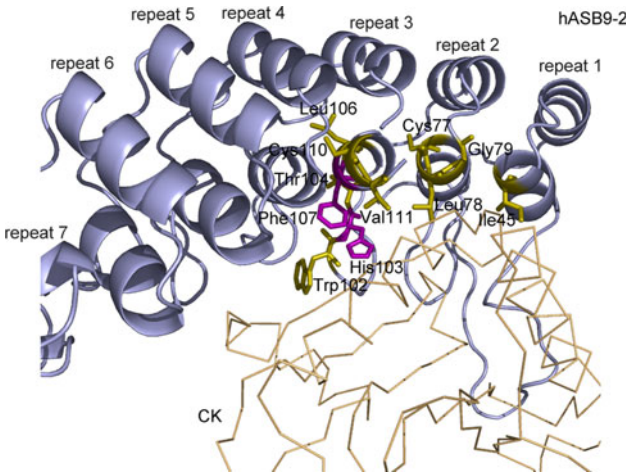

Fig. 5 Representation of hASB9-2 bound to CK. HASB9-2 is shown in cartoon representation in *light blue* whereas CK is shown in ribbon representation in *light orange*. All the residues used for amino acid substitution analysis are shown as *sticks*. Among them, residues His103 and Phe107 which were essential for binding are colored in *purple*, while others are colored in *olive*. This figure is generated using PyMOL (Color figure online)

used to map the Deltex binding surface on the ankyrin domain of *Drosophila* Notch receptor [1]. Significantly, researchers had used sedimentation velocity analytical ultracentrifugation to characterize the Notch/Deltex heterodimer and measured its affinity [1]. It gives us a direction to perfect our research and, to measure whether N-4 truncated mutant still folded and bound.

Understanding the interaction between hASB9-2 and CKB will help us to better understand the mechanism by which hASB9 recognizes CKB and targets it to be degraded through the ubiquitin–proteasome system. Here, we showed the first crystal structure of the ASB protein and present a model how hASB9-2 can bind to CKB based on in vitro pull-down assays. The binding model requires further verification by either co-crystallization of hASB9-2 and CKB or mutational analysis throughout hASB9-2 and counterpart residues in CKB.

Acknowledgments This work was supported by the National Basic Research Program of China (973 Program, 2007CB914304 & 2009CB825505), the National Natural Science Foundation of China (30770427), the New Century Excellent Talents in University (NCET-06-0356), the Shanghai Leading Academic Discipline Project (B111) and the National Talent Training Fund in Basic Research of China (No. J0630643).

References

1. Allgood AG, Barrick D (2011) J Mol Biol 414(2):243–259
2. Chung AS, Guan YJ, Yuan ZL, Albina JE, Chin YE (2005) Mol Cell Biol 25(11):4716–4726

3. De Sepulveda P, Ilangumaran S, Rottapel R (2000) J Biol Chem 275:14005–14008
4. Debrincat MA, Zhang JG, Willson TA, Silke J, Connolly LM, Simpson RJ, Alexander WS, Nicola NA, Kile BT, Hilton DJ (2007) J Biol Chem 282:4728–4737
5. Edwards MS, Sternberg JE, Thornton JM (1987) Protein Eng 1:173–181
6. Ehebauer MT, Chirgadze DY, Hayward P, Martinez Arias A, Blundell TL (2005) Biochem J 392(Pt 1):13–20
7. Fei X, Zhang Y, Gu X, Qiu R, Mao Y, Ji C (2009) Protein Pept Lett 16:333–335
8. Gouet P, Courcelle E, Stuart DI, Metoz F (1999) Bioinformatics 15:305–308
9. Kaddurah-Daouk R, Lillie JW, Daouk GH, Green MR, Kingston R, Schimmel P (1990) Mol Cell Biol 10:1476–1483
10. Kamizono S, Hanada T, Yasukawa H, Minoguchi S, Kato R, Minoguchi M, Hattori K, Hatakeyama S, Yada M, Morita S, Kitamura T, Kato H, Nakayama K, Yoshimura A (2001) J Biol Chem 276:12530–12538
11. Kile BT, Viney EM, Willson TA, Brodnicki TC, Cancilla MR, Herlihy AS, Croker BA, Baca M, Nicola NA, Hilton DJ, Alexander WS (2000) Gene 258:31–41
12. Kohroki J, Nishiyama T, Nakamura T, Masuho Y (2005) FEBS Lett 579:6796–6802
13. Kwon S, Kim D, Rhee J, Park J, Kim D, Kim D, Lee Y, Kwon H (2010) BMC Biol 8:23
14. Manjasetty BA, Quedenau C, Sievert V, Büssow K, Niesen F, Delbrück H, Heinemann U (2004) Proteins 55(1):214–217
15. McArthur MW, Laskowski RA, Moss DS, Thornton JM (1993) J Appl Cryst 26:283–291
16. Michaely P, Tomchick DR, Machius M, Anderson RG (2002) EMBO J 21:6387–6396
17. Mosavi LK, Cammett TJ, Desrosiers DC, Peng Z (2004) Protein Sci 13:1435–1448
18. Murshudov GN, Vagin AA, Dodson EJ (1997) Acta Crystallogr D 53:240–255
19. Notredame C, Higgins DG, Heringa J (2000) J Mol Biol 302:205–217
20. Sedgwick SG, Smerdon SJ (1999) Trends Biochem Sci 24:311–316
21. Stebbins CE, Kaelin WG, Pavletich NP (1999) Science 284:455–461
22. Storoni LC, McCoy AJ, Read RJ (2004) Acta Crystallogr D Biol Crystallogr 60:432–438
23. Wallimann T, Hemmer W (1994) Mol Cell Biochem 133: 193–220
24. Wallimann T, Wyss M, Brdiczka D, Nicolay K, Eppenberger HM (1992) Biochem J 281:21–40
25. Wyss M, Kaddurah-Daouk R (2000) Physiol Rev 80:1107–1213
26. Zhao J, Schmieg FI, Simmons DT, Molloy GR (1994) Mol Cell Biol 14:8483–8492
27. Zweifel ME, Barrick D (2001) Biochemistry 40(48):14344–14356
28. Zweifel ME, Barrick D (2001) Biochemistry 40(48):14357–14367
29. Zweifel ME, Leahy DJ, Hughson FM, Barrick D (2003) Protein Sci 12(11):2622–2632