doi:10.1016/j.jmb.2009.07.069 J. Mol. Biol. (2009) 392, 1242–1252

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

# Crystal Structures of α-Crystallin Domain Dimers of αB-Crystallin and Hsp20

C. Bagnéris[1]†, O. A. Bateman[1]†, C. E. Naylor[1]†, N. Cronin[1], W. C. Boelens[2], N. H. Keep[1] and C. Slingsby[1]*

[1]*Department of Crystallography, Birkbeck College, Institute of Structural and Molecular Biology, Malet Street, London WC1E 7HX, UK*

[2]*Department of Biomolecular Chemistry 271, Radboud University, Nijmegen, P.O. Box 9101, NL-6500 HB Nijmegen, The Netherlands*

*Received 22 May 2009; received in revised form 22 July 2009; accepted 23 July 2009*
Available online 30 July 2009

Small heat shock proteins (sHsps) are a family of large and dynamic oligomers highly expressed in long-lived cells of muscle, lens and brain. Several family members are upregulated during stress, and some are strongly cytoprotective. Their polydispersity has hindered high-resolution structure analyses, particularly for vertebrate sHsps. Here, crystal structures of excised α-crystallin domain from rat Hsp20 and that from human αB-crystallin show that they form homodimers with a shared groove at the interface by extending a β sheet. However, the two dimers differ in the register of their interfaces. The dimers have empty pockets that in large assemblies will likely be filled by hydrophobic sequence motifs from partner chains. In the Hsp20 dimer, the shared groove is partially filled by peptide in polyproline II conformation. Structural homology with other sHsp crystal structures indicates that in full-length chains the groove is likely filled by an N-terminal extension. Inside the groove is a symmetry-related functionally important arginine that is mutated, or its equivalent, in family members in a range of neuromuscular diseases and cataract. Analyses of residues within the groove of the αB-crystallin interface show that it has a high density of positive charges. The disease mutant R120G α-crystallin domain dimer was found to be more stable at acidic pH, suggesting that the mutation affects the normal dynamics of sHsp assembly. The structures provide a starting point for modelling higher assembly by defining the spatial locations of grooves and pockets in a basic dimeric assembly unit. The structures provide a high-resolution view of a candidate functional state of an sHsp that could bind non-native client proteins or specific components from cytoprotective pathways. The empty pockets and groove provide a starting model for designing drugs to inhibit those sHsps that have a negative effect on cancer treatment.



***Edited by R. Huber***

*Keywords:* αB-crystallin; Hsp20; Hsp27; Hsp22; chaperone

## Introduction

The heat shock proteins (Hsps) arose to keep unicellular organisms alive during harsh conditions through the heat shock response.[1] In vertebrates, they have expanded into large families involved in macromolecular quality control[2] and are integrated in networks that control the balance of cell death or cell survival.[3] Whereas structural biology, underpinned by yeast genetics, has unraveled many biological roles for the conserved ATP-dependent Hsps,[4] the cellular functions of the more diverse small Hsps (sHsps) are less well defined. Human sHsp family members are associated with processes related to human health, such as inflammation, muscle relaxation, longevity, cell division, cell survival and cell death.[5–10] Unsurprisingly, they are therapeutic targets and biomarkers for a whole range of diseases.[11]

Several members of the human sHsp family, Hsp27 (B1), αA-crystallin (B4), αB-crystallin (B5),

*Corresponding author. E-mail address: C.Slingsby@mail.cryst.bbk.ac.uk.

† C.B., O.A.B. and C.E.N. contributed equally to this work.

Abbreviations used: Hsp, heat shock protein; sHsp, small heat shock protein; ACD, α-crystallin domain.

Hsp20 (B6) and Hsp22 (B8), are highly expressed in long-lived cells of muscle, lens and brain.[12,13] Hsp27, αB-crystallin and Hsp22 are stress inducible.[14] (The nomenclature in parentheses follows HUGO guidelines and is mapped to the original names as outlined in Refs. 12 and 14.) The most pervasive function appears to be that of a molecular chaperone,[15] binding non-native "client" proteins during stress when energy levels are low before passing them on to a network of ATP-driven chaperones for either refolding or degradation.[16,17] The strongly cytoprotective function of αB-crystallin, Hsp27 and Hsp22 may also involve binding to specific components of apoptosis[11,18,19] and autophagy[20] pathways.

sHsps form large polydisperse assemblies[21] of variable oligomeric substructure[22] in heat-stimulated rapid equilibrium with smaller-assembly species.[23,24] Mammalian sHsps respond to stress/kinase transduction pathways by N-terminal phosphorylation, which can lead to sHsp disassembly.[25] It is possible that disassembly is part of the mechanism whereby functional binding sites are presented to their target substrates. Alternatively, or in addition, there may be more rapid conformational changes leading to an activated form of the parent oligomer, as has been suggested for yeast Hsp26.[26]

Sequence studies show that sHsps contain an α-crystallin domain (ACD) sequence preceded by a variable N-terminal region[27] (Fig. 1). Although the polydispersity of sHsps has hindered their structural characterization, a few ordered assemblies have been solved at atomic resolution. Crystal structures of monodisperse sHsp oligomers from *Methanococcus jannaschii* Hsp16.5,[28] wheat Hsp16.9[29] and tapeworm Tsp36[30] show that the ACD is a β sandwich: in common are strands β2–β9, with β2/β7 and β4/β8 strands forming sticky edges that are patched by flanking sequence extensions. Non-metazoan sHsps are assembled from dimers using an ACD β6-strand loop reciprocal exchange mechanism that contributes to patching the β2/β7-sandwich edge. In wheat Hsp16.9, this edge is also patched by intramolecular interactions with the hydrophobic N-terminal extension. Higher-order assembly is driven in part by interactions between a conserved sequence motif (I–X–I/V, with I and I/V highlighted in brown) in the C-terminal extension that patches the β4/β8 edge of the ACD from another dimer. In tapeworm Tsp36, where the sequence encompasses two ACDs and there is no conserved C-terminal sequence motif, the β4/β8 edge of one domain is patched instead by an N-terminal motif (I–X–V, with I and V highlighted in

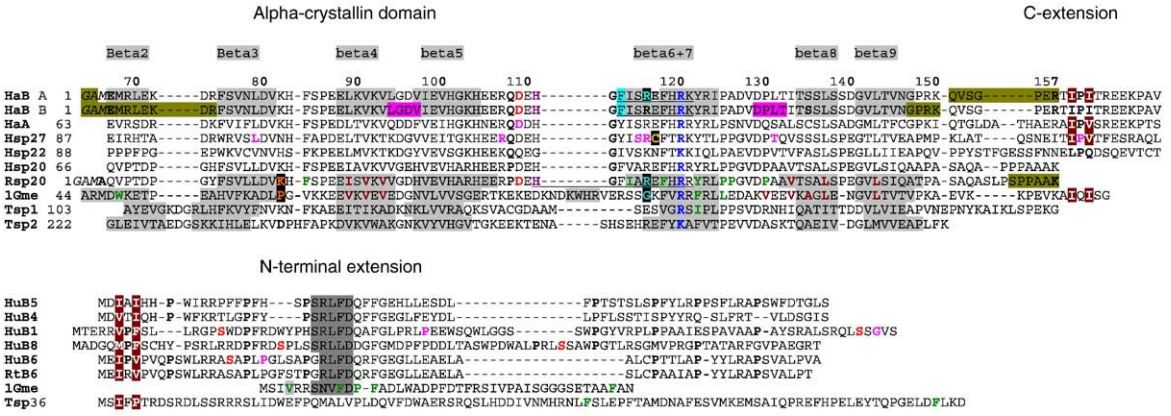


**Fig. 1.** sHsp sequence alignment. The sequences of hHaB (P02511 HSPB5 CRYAB human αB-crystallin), hHaA (P02489 HSPB4 human CRYAA αA-crystallin), hHsp27 (P04792 HSPB1 human Hsp27), hHsp22 (Q9UJY1 HSPB8 human Hsp22), hHsp20 (O14558 HSPB6 human Hsp20), rHsp20 (P97541 HSPB6 rat Hsp20), 1gme (Q41560 wheat *Triticum aestivum* Hsp16.9B) and Tsp36 (Q7YZT0 *Taenia saginata* Tsp36) are aligned structurally in the region of the ACD and the C-terminal extension. The numbers adjacent to the labels hHaB and rHsp20 are the numbers of the tagged constructs, whereas for the other sequences, the numbers refer to the UniProt scheme. The numbers along the first row correspond to the UniProt sequence of human αB-crystallin. Residues in italics are from the construct's N-terminal tag. In human αB-crystallin, the two chains of ACD dimer are labeled hHaB A and hHaB B, with regions of conformational difference highlighted in pink. Residues highlighted in dull green are not seen in the electron density. β strands B2–B9 are highlighted in light gray. The hydrophobic C-terminal motifs are highlighted in brown. The last residues of the two ACD constructs are shown in boldface: note that the αB ACD construct does not include the C-terminal hydrophobic motif, whereas the rat Hsp20 ACD construct includes the C-terminus, but that the wild-type sequence does not contain the C-terminal hydrophobic motif. Residues underlined in hHaB and rHsp20 are at the dimer interface in the crystal structure. In the text, UniProt numbers are quoted first, with the ACD construct number following in parentheses. Residues colored brown in rHsp20 and 1gme form the β4/β8 pockets. Non-polar residues colored green in rHsp20 line the shared groove, and residues colored green in 1gme interact with its N-terminal extension. Specific residues mentioned in the text are highlighted and color coordinated with the figures. The more variable N-terminal extensions are aligned by sequence below. The hydrophobic sequence motif at the N-terminus is highlighted in brown. In 1gme and Tsp36, residues colored green make intrachain interactions with residues colored green in their respective ACDs in the vicinity of conserved arginine colored in blue boldface, whereas V4 in 1gme makes an equivalent interaction with another chain in the dodecameric assembly. Additional disease mutations are in magenta. Serine phosphorylation sites in the mammalian sHsps are colored red.

brown) from a partner domain (Fig. 1). Structural information on patching the β2/β7-sandwich edge in metazoans is incomplete as the strand-exchange loop is too short for dimerization (Fig. 1).

Further molecular insights into the structure of the mammalian sHsps were gained through the creation of monodisperse species by removing sequence extensions. It has been shown that the αB ACD can fold independently into a dimeric species.[31] Recently, a solid-/solution-state NMR study of the ACD and full-length human αB-crystallin showed an elongated β6+β7 strand that was inferred to be at the dimer interface.[32] These smaller species were shown to function as chaperones *in vitro* using model substrates. Together, these data provide a model whereby sequence extensions play a role in higher assembly but when removed might unmask binding sites for client proteins. However, there is evidence that the N-terminal extension is required for *in vivo* function.[33]

Clues to the structural roles of human sHsps may also be gained from the phenotypes caused by point mutations. Extensive pathology ensues when mutations in stress-induced and/or highly expressed mammalian sHsps result in non-native or destabilized proteins that may sequester normal cellular components.[34] Earlier studies showed that a point mutation to a conserved arginine within the ACD of αA-crystallin and αB-crystallin causes cataract and skeletal muscle desmin-related myopathy, respectively (disease arginine colored in blue boldface in alignment),[35,36] and that αB-crystallin R120G formed enormous cytoplasmic inclusions in mammalian cells.[37] It is now known that the equivalent site is mutated in Hsp27 (HSPB1) and Hsp22 (HSPB8) and at several other sites in these proteins

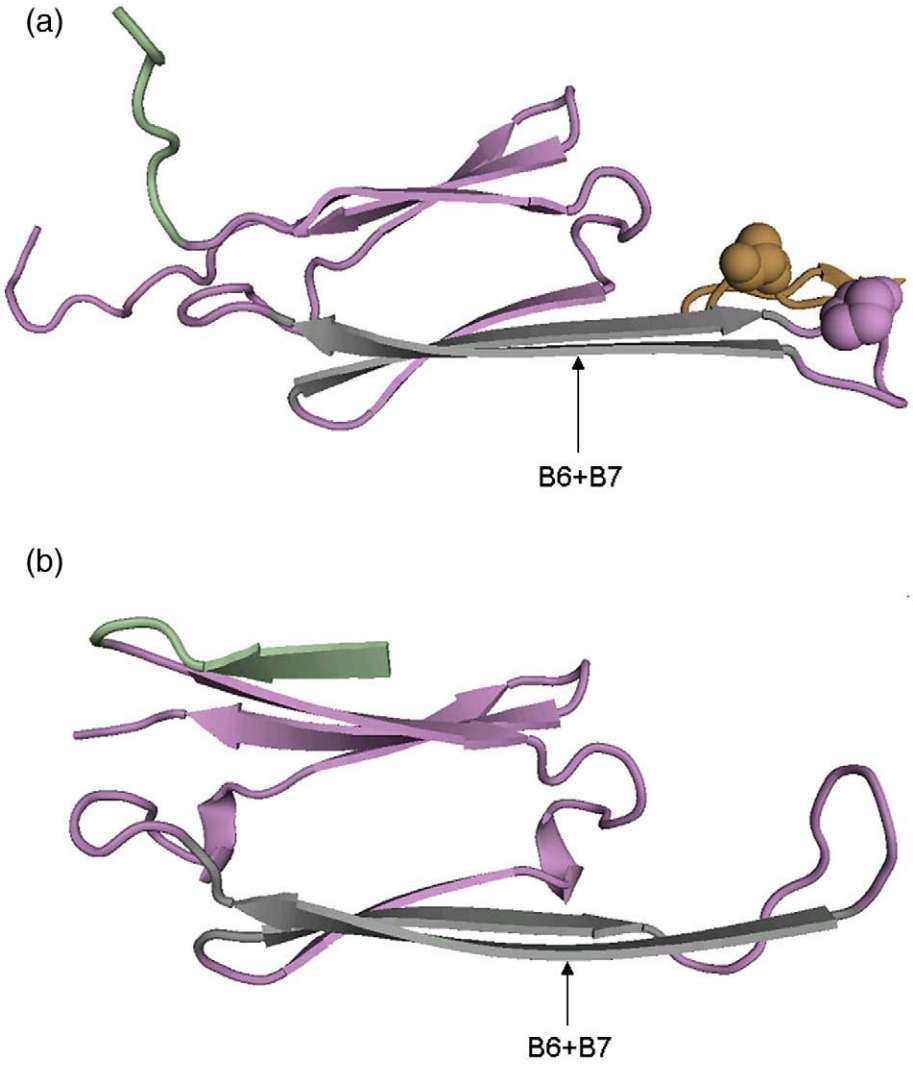


**Fig. 2.** Differences in the ACDs. (a) The ACD of rat hsp20 (plum ribbon) with the sequence equivalent of β2 strand in green, the β5 and β6+β7 strands in gray and the intervening residues (construct residues 45–50) in plum. A region of the hairpin loop (construct residues 43–51) from a symmetry-related molecule in the lattice is shown in khaki. Appended is Pro46 (in space fill) from each chain, showing it is close to the dyad axis. (b) The ACD of human αB (plum ribbon) with the sequence equivalent of β2 strand in green, the β5 and β6+β7 strands in gray and the intervening residues (construct residues 42–49) in plum.

(colored in magenta boldface in Fig. 1) in the most common inherited neuromuscular disorder, Charcot–Marie–Tooth disease.[38] Three family members (Hsp27, Hsp22 and αB) are often co-expressed in the same cell and can co-assemble into hetero-oligomers,[39] thus spreading the effect of any mutation. A mouse model for heart disease shows that cardiomycete-restricted overexpression of R120G αB-crystallin places the heart under reductive stress, causing increased expression of the redox chaperone, Hsp25/Hsp27.[40]

Here, the crystal structures of excised ACDs from rat Hsp20 and human αB-crystallin show that they form dimers in the crystal lattice. Each dimer structure has a shared groove containing the common disease arginine as a symmetry-related pair. The groove is the likely binding site for the N-terminal extension.

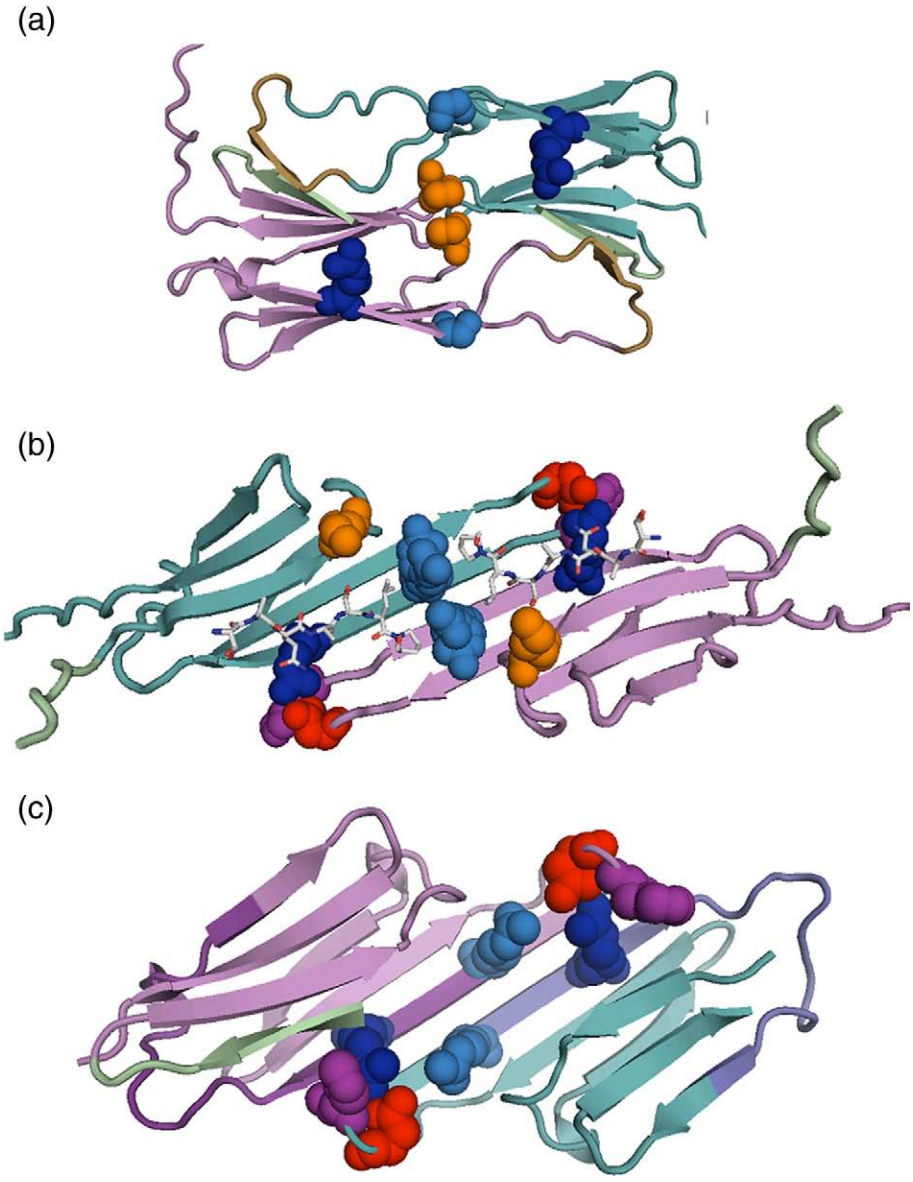


**Fig. 3.** The different dimers. (a) In wheat Hsp16.9 (PDB ID 1gme), the dimer (pink and blue ribbons) is formed by strand exchange by β6 strands (tan). Residues P62 (orange) and G104 (light blue) and R108 (blue) are appended in space fill β2 strand (green) is part of the ACD. (b) In rat hsp20 ACD dimer, there is a reorganization of the paired domains that results in a shared groove: note the relative positions to the dimer 2-folds of K81 (K20; orange) and R115 (R54; light blue), with sequences equivalent to those of Hsp16.9 P62 (orange) and G104 (light blue). The groove is partially filled by C-terminal extensions (ball-and-stick representation) from other chains in the lattice. The region equivalent to β2 strand (green) also engages in lattice interactions. The conserved R119 (R58; blue) from one chain interacts across the interface with D108 (D47; red) close to H110 (H49; plum) of the partner chain. (c) In αB ACD dimer, there is a change in register at the 2-fold interface that places R116 (R53; light blue), the sequence equivalent of Hsp20 R54 (light blue), closer to the side chain of partner R120 (R57; blue) yet still allows R120 (R57; blue) to ion pair across the interface with D109 (D46; red), close to H111 (H48; plum). Antibodies to the region of the αB ACD dimer interface and β8-loop region (dark pink ribbon and dark blue ribbon) have been found in the cerebrospinal fluid of multiple sclerosis patients.[41]

## Results

### Structure of ACDs of rat Hsp20 and human αB-crystallin

The high-resolution (1.3 Å) structure of the ACD of rat Hsp20 (P97541), residues 65–162, was solved by molecular replacement using core ACD wheat Hsp16.9 coordinates [Protein Data Bank (PDB) ID 1gme] in a C2 cell with one chain in the asymmetric unit.

The domain comprises two β sheets made from strands β3–β9, indicated in the sequence alignment (Fig. 1). The region of sequence equivalent to β2 strand in wheat Hsp16.9 is not part of the domain and is making interactions with other dimers in the lattice. A striking feature of the domain is a long hairpin comprising the β5 strand, the elongated (β6+β7) edge strand and the intervening turn (Fig. 2a).

The 2.9-Å X-ray structure of the ACD of human αB-crystallin (P02511), residues 67–157, was solved by molecular replacement using the rat Hsp20 ACD in a C2 cell with five domains in the asymmetric unit. In αB-crystallin ACD, the long hairpin is more curved compared with that of Hsp20 (Fig. 2b). The differing conformations of the hairpins could be related to the sequence change at **Q**108 (Q45 in the construct) in αB, which in Hsp20 is **P**107 (P46 in the construct), with the sequence-conserved glycine four residues downstream taking the strain (Fig. 1). The conformation of the Hsp20 hairpin may also be modulated by its interaction with a symmetry-related molecule close to a lattice dyad (Fig. 2a).

### Comparison of the dimer interfaces

In the unit cell of Hsp20 ACD, a crystallographic 2-fold axis is positioned between β6+β7 strands such that they form an extended antiparallel β sheet linking two domains (Fig. 3), similar to the proposed solution dimer of αB-crystallin based on NMR spectroscopy.[32] The X-ray structure of Hsp20 ACD shows that in this dimer the β2/β7 edge of the sandwich is opened out compared with that in wheat Hsp16.9 ACD (Fig. 3a). This reorganization of a pair of ACDs results in a shared groove (Fig. 3b) that has a superficial resemblance to the peptide binding grooves of major histocompatibility complex molecules or the conserved core of the luminal domain of the yeast ER sensor Ire1 that triggers the unfolded protein response.[42]

The αB-crystallin ACD forms a similar dimer as Hsp20 ACD (Fig. 3c). The dimers form a spiral in the crystal lattice using both non-crystallographic and crystallographic axes. Four domains in the asymmetric unit of the αB unit cell form two dimers in which the two chains have differences in conformation (Fig. 1). The fifth domain in the asymmetric unit interacts across a crystallographic dyad and is not well resolved. The two chains of the dimers within the asymmetric unit also show differences in that one chain has an ordered β2 strand as part of the ACD (Fig. 2b), similar to wheat Hsp16.9 (Fig. 3a).

Although these two mammalian sHsps have a similar dimer arrangement, the different hairpin conformation is associated with a shift in register at the dimer interface (Fig. 3b and c), with Hsp20 having a more extended interface than αB (Fig. 1). These differences, along with the additional β2 strand, result in the αB ACD dimer being more compact and the groove being less accessible than in Hsp20 (Fig. 3c). It is unclear from these data as to what extent the structural differences are due to sequence changes between the two sHsps or their different crystal lattice arrangements. The difference is of particular interest for human Hsp27, which is more closely related in sequence to αB-crystallin than Hsp20 at the interface: similar to αB, it has a glutamine rather than a proline in the hairpin loop and a serine rather than an alanine close to the interface (Figs. 1 and 3).

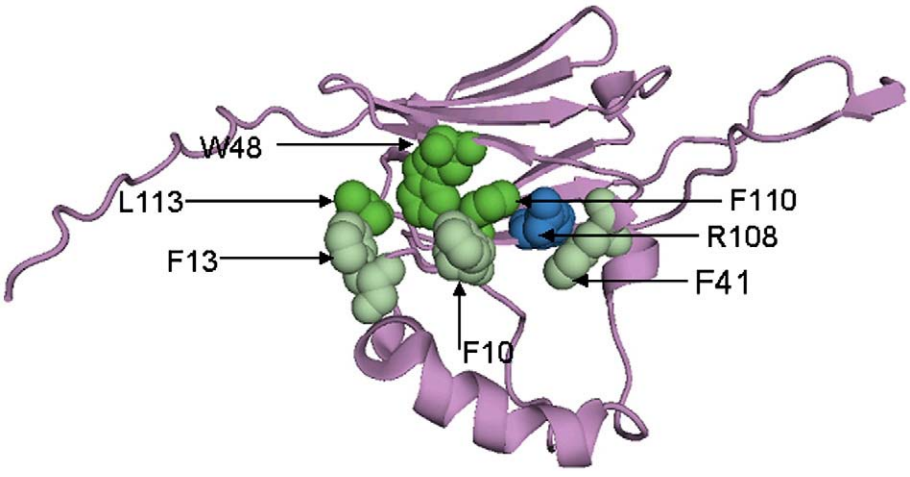


**Fig. 4.** Patching an edge in wheat Hsp16.9. Chain A (plum ribbon) extracted from wheat Hsp16.9 dodecamer showing that hydrophobic residues (dark green) in the β2/β7 edge are patched by hydrophobic residues (light green) from the intramolecular N-terminal extension in the vicinity of the conserved arginine (blue).

Human Hsp27 has a cysteine in an equivalent position to E117 (E54) in αB. This residue is closest to the interface dyad axis in αB and projects underneath the β sheet, whereas the equivalent residue in Hsp20 ACD is displaced somewhat from the dyad. If the detail of these two observed dimer interfaces is subject to modulation by the environment, then this would impact on the distance between the redox cysteines at the dimer interface of Hsp27, the redox chaperone.[43]

The solution-/solid-state NMR study showed that the region equivalent to strand β2 was not in a single

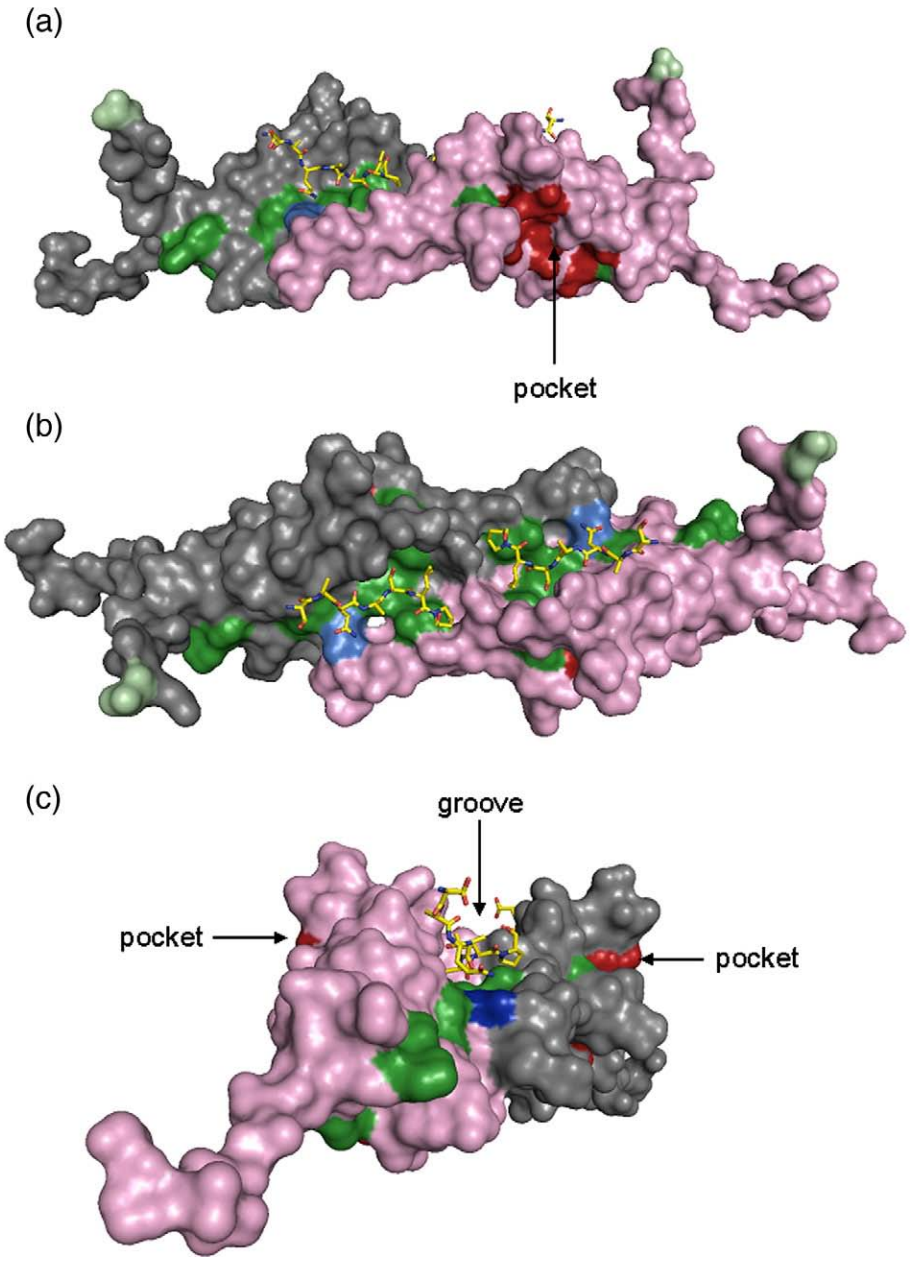


**Fig. 5.** Spatial arrangement of groove and pockets in Hsp20 dimer. (a) In the Hsp20 lattice, C-terminal sequence extensions (yellow ball-and-stick representation) bind into the central part of the shared groove of the dimer (pink and gray space fill). Conserved arginine (blue) and non-polar residues shown in green (and colored green in the Hsp20 sequence in Fig. 1) line the shared groove and lead to the β4/β8 pockets colored brown (and colored brown in Fig. 1 as well) in each domain. One possibility is that in full-length Hsp20, hydrophobic residues from one N-terminal chain interact along the entire groove with the N-terminal motif I–X–V (with I and V highlighted in brown in alignment) placed into the pockets of one domain. (b) Looking down at the shared groove of rat Hsp20 dimer. The two ACD chains are colored in gray and pink space fill, with non-polar residues lining the groove in green and the disease arginines in blue. These correspond to the residues colored green and blue in the Hsp20 ACD sequence in Fig. 1. The N-terminus of the ACD is shown in pale green space fill. The C-terminal extensions from symmetry-related molecules are shown lining the middle part of the groove in yellow ball-and-stick representation. This view has been rotated 90° along the horizontal axis compared with (a). (c) The groove shown when the upper image is rotated 90° perpendicular to the viewer.

conformation in the full-length αB-crystallin assembly.[32] The crystal structures show that the two chains of αB ACD in the dimer differ in this region, with one chain having a β2 strand in place in the β sheet, similar to wheat Hsp16.9, while the other is not visible in the lattice. In the Hsp20 ACD, both chains of the dimer are identical, with the region equivalent to β2 strand visible but involved in interactions with other dimers in the lattice. Together, these data indicate how the sequence equivalent to β2 strand might contribute toward flexibility in assembly of sHsps.

### The location of higher-assembly sites in the dimer

In the Hsp20 ACD dimer structure, the shared groove is partly occupied by two symmetry-related C-terminal extensions, in polyproline type II conformation, from other dimers in the lattice (Fig. 3b). Each peptide contributes to patching-exposed hydrophobics at the β2/β7-sandwich edge. In wheat Hsp16.9, hydrophobic residues from within the central region of the N-terminal extension interact with similar hydrophobics from the β2/β7 edge from the same chain ACD in the vicinity of the conserved arginine (Figs. 1 and 4), and a similar arrangement occurs in one of the tapeworm Tsp36 ACDs (see Fig. 6 in Ref. 30). In wheat Hsp16.9, the N-terminal extension both patches an intrachain ACD and contributes to higher assembly by linking dimers together.

On the other side of each β-sandwich domain of the dimer is a hydrophobic pocket formed between edge strands β4 and β8 (Figs. 1 and 5). sHsp higher assembly is also driven by the filling of these hydrophobic pockets with hydrophobic sequence motifs (I–X–I/V, with I and I/V highlighted in brown) from flanking C-terminal extensions from other dimers in the assembly[28,29] (Fig. 1). However, mammalian Hsp20 and Tsp36 are different from other sHsps in that they do not have the C-terminal hydrophobic motif (Fig. 1). In Tsp36, this patching function is performed by the I–X–P (with I and P highlighted in brown) N-terminal motif, which fills the β4/β8 pocket from another chain.[30] Similar hydrophobic N-terminal motifs are present in the sequences of mammalian sHsps (Fig. 1).

The crystal structures thus show the spatial arrangement of a groove and pockets in the ACD dimer assembly unit of mammalian sHsps.

### The disease arginine

The sequences of the two solved structures are aligned with the sequences of wheat Hsp16.9 and the two ACDs of tapeworm Tsp36, as well as several other human sHsps (Fig. 1). The site of the common disease arginine mutation is colored in blue boldface. In αB-crystallin, this is R120G (R57), and the equivalent arginine in Hsp20 is R119 (R58). This common disease mutation site is situated at the dimer interface on either side of the inside of the shared groove (Figs. 3 and 5). The αB-crystallin R120

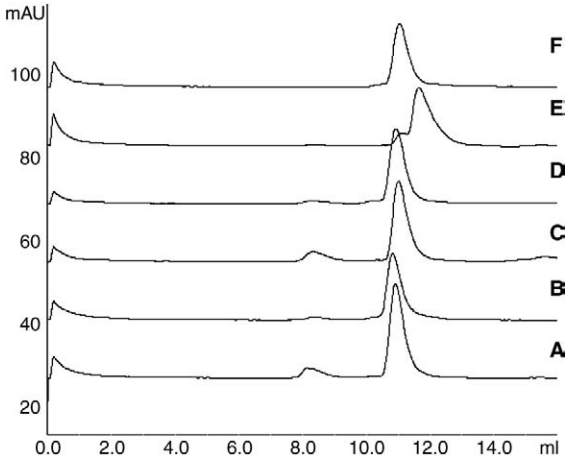


**Fig. 6.** The R120G αB ACD is more stable than the wild type. Gel-filtration profiles of αB ACD compared with αB R120G ACD over the pH range 6.0–8.5 show that the mutant ACD behaves more like a dimer than the wild-type sequence at pH 6.0.

(R57) forms two symmetry-related bidentate ion pairs with D109 (D46) across the dimer interface of the αB ACDs. Despite the altered secondary structure of the hairpin loop, these interface ion pairs are conserved in Hsp20 ACD. Comparison of the interfaces of the two solved dimers shows that there is a difference in the interactions around the "disease arginine" that is associated with the change in register involving the symmetry-related arginine four residues N-terminal in the sequences (colored in light blue boldface in Figs. 1 and 3). In the pH 6.5 Hsp20 ACD lattice, the equivalent to the disease arginine R119 (R58) is close to partner D108 (D47) and partner H110 (H49), whereas in the pH 9.0 αB ACD lattice, the disease arginine R120 (R57) is close to partner R116 (R54).

Solution NMR studies have shown that the wild-type αB ACD dimer undergoes changes consistent with monomer formation at acidic pH.[32] Insight into the electrostatic forces involved in this interface interaction was gained when the R120G αB ACD was purified and subject to size estimation. The R120G αB ACD is more stable as a dimer at acidic pH than the wild-type ACD sequence (Fig. 6). The R120G mutation thus appears to stabilize the dimer at acidic pH by removal of positive charge from the interface (Fig. 3b and c).

## Discussion

The Hsp20 ACD and αB ACD form similar dimers in their respective crystal lattices, although there are

significant differences in the detail of their dimer interfaces. There is evidence from solid-state NMR spectroscopy that this basic dimer assembly unit corresponds to an interface found in large assemblies of αB-crystallin built from the full-length sequence.[32] The presence of a shorter loop region in vertebrate sHsps compared with other sHsp sequences (Fig. 1) has resulted in a new interface whereby an extended β sheet forms the platform of a groove that contains hydrophobic residues from the β2/β7-sandwich edges of two ACDs, as well as their conserved arginines. Structural homology with other sHsp three-dimensional assemblies defined by X-ray crystallography leads to the consideration that this groove is likely the binding site for hydrophobic residues from the central region of an N-terminal extension. It is known from these high-resolution sHsp structures that the β4/β8-edge pockets are filled using hydrophobic sequence motifs from other chains in the assembly, found either at the extreme N-terminus or within the C-terminal extension. These dimer structures therefore show the relative spatial locations of assembly sites and provide a starting point for modelling higher assembly forms of the α-crystallins, human Hsp20 and Hsp27. However, it is unclear where the binding motifs are in the C-terminal extension sequence of Hsp22 (B8).

The simplest sHsp to model in broad outline is full-length human Hsp20, as it is a dimer,[44] with each chain having an N-terminal I–X–V (with I and V highlighted in brown) motif but no C-terminal hydrophobic motif. Taking into consideration the disposition of the pockets and the shared groove, one possible model is for one N-terminal extension to bind along the entire shared groove (as in peptide-bound major histocompatibility complex heterodimer) and then to wrap across the sandwich, placing its N-terminal I–X–V (with I and V highlighted in brown) motif into the ACD pocket from a partner domain (Figs. 1 and 5. This arrangement has similarity to wheat Hsp16.9 dodecamer where only 6 of 12 N-terminal extensions were resolved.[29] It is known that Hsp20 N-terminal S16 is phosphorylated and displaces cofilin from the adaptor protein 14-3-3γ, leading to disruption of actin stress fibres and smooth muscle relaxation.[45] It is plausible that phosphorylation of both Hsp20 N-terminal extensions prevents binding in the interface groove, making them available for binding to the dimeric adaptor protein, as demonstrated *in vitro*.[44]

Full-length α-crystallins and Hsp27 form larger and more disperse assemblies. Each chain has two candidate hydrophobic sequence motifs, one at the extreme N-terminus and another within the C-terminal extension. It is known from previous sHsp structures[28–30] that the hydrophobic sequence motif can bind across one edge of the domain in either direction. Furthermore, the sequence of the C-terminal hydrophobic motif of human αB-crystallin is a palindrome (Fig. 1). In addition, the structure of wheat Hsp16.9 dodecamer shows how different orientations of the linker preceding the C-terminal hydrophobic motif are used to build the higher assembly.[29] It may be that sHsp polydispersity stems from the multiple ways in which grooves and pockets in the ACDs accept sequence extensions.

It had previously been noted that within the N-terminal sequence there is a conserved sequence motif (**S**XX**FD**, highlighted in dark gray in Fig. 1) in a subset of mammalian sHsps that correlated with the presence of a hydrophobic residue of the ACD[29] now known to be a groove hydrophobic. It has been shown that removal of the sequence **S**RL**FD**QFFG from the N-terminal region of the α-crystallins resulted in smaller-assembly species with increased chaperone activity.[46] These observations are consistent with dimers retaining some chaperone activity and support a model whereby disassembly unmasks peptide binding sites on the ACD. These sites would be the pockets and grooves and would be candidate binding sites for the specific partners identified in pathways associated with cytoprotection.[9,18–20]

The disease arginine is located in two symmetry-related positions in the shared groove at the dimer interface. The observation that the αB-crystallin ACD dimer is stabilized at moderately acidic pH by the R120G mutation suggests that altered dynamics of the sHsp assembly contribute to the large cytoplasmic inclusions characteristic of the pathology caused by this mutation. This indicates that in normal sHsp function, the dimer has to dissociate to some extent, possibly to enable it to release bound peptides in the acidic environment of the aggregate clearance pathway.[47] It is worth noting that antibodies to the region of the αB ACD dimer interface and β8-loop region have been found in the cerebrospinal fluid of multiple sclerosis patients (Fig. 3b).[41] It may be that the monomer is the active binding species, with the long tracts of exposed regular backbone from the highly structured hairpin loops providing hydrogen-bond partners for backbone atoms of β strands awaiting folding or destruction.

Although sHsps are considered to be largely beneficial for health, the strongly cytoprotective effects of Hsp27 and αB-crystallin are harmful when present in certain kinds of cancer cells, especially when upregulated following chemotherapy.[11,48] The grooves and pockets defined in these two mammalian ACDs provide a starting point for designing small molecules that could cause sHsp malfunction by stabilizing interfaces and blocking access to client proteins.

## Methods

### Cloning

The human αB-crystallin domain (amino acids 67–157) was cloned into pPROEX™ HT(b) (GibcoBRL) to give an N-terminal His tag cleavable using tobacco etch virus protease. After cleavage, an additional GAM is present at

the N-terminus of the αB-crystallin domain 67–157. First, a PCR using the forward primer TTTTCCATGGAGATGCGCCTGGAGAAG containing an NcoI restriction site (underlined) and the reverse primer TTTTCTCGAGCTAGCGCTCAGGGCC containing an XhoI site (underlined) was generated using pET16(b) αB-crystallin as a template. The PCR product was gel purified and cleaved with NcoI and XhoI enzymes before being ligated into a previously digested pPROEX™ HT(b) using the same restriction enzymes.

Similarly, Hsp20 *Rattus norvegicus* ACD (amino acids 65–162) was cloned into pPROEX™ HT(b) by using the forward primer TTTTCCATGGCCCAGGTGCCCACGG containing an NcoI site (underlined) and the reverse primer TTTTAAGCTTTCACTTGGCAGCAGGTG containing a HindIII site (underlined) and pET16(b)Hsp20 as a template.

### Expression and purification

pPROEX™ HT(b) αB-crystallin 67–157 (human αB ACD) and pPROEX™ HT(b) Hsp20 *R. norvegicus* ACD 65–162 (rat Hsp20 ACD) were each transformed into BL21 (DE3) (Novagen) and then expressed and purified using the same protocol. Cells were grown in 2YT medium [1.6% (w/v) bacto-tryptone (Invitrogen), 1% bacto-yeast extract (Invitrogen) and 0.5% NaCl (Sigma), pH 7.2)] containing ampicillin (100 μg $ml^{-1}$) at 37 °C to an optical density at 600 nm of 1 before being induced with 1 mM isopropyl 1-thio-β-D-galactopyranoside at 22 °C overnight. After the cells were harvested, the pellet was resuspended in 25 mM Tris, pH 8.5, and 200 mM NaCl (buffer A) containing DNase I (final concentration of 10 μg/ml) and ethylenediaminetetraacetic-acid-free protease inhibitor cocktail (Roche) before sonication on ice. Cell debris and unbroken cells were removed by centrifugation (46,000*g* for 1 h at 4 °C), and the supernatant was filtered through a 0.45-μm filter. Cleared cell extract was applied to a 5-ml HisTrap™ FF column (GE Healthcare) preequilibrated with buffer A. The column was connected to an AKTA FPLC machine (GE Healthcare), and unbound proteins were washed with 10 column volumes of buffer A followed by 6 column volumes of buffer A containing 30 mM imidazole. Bound proteins were eluted with buffer A containing 250 mM imidazole. The tag was removed by dialysing overnight in 25 mM Tris, 200 mM NaCl, pH 8.5, 1 mM DTT and 0.5 mM ethylenediaminetetraacetic acid with the tobacco etch virus protease. The mixture was applied to a buffer-A-preequilibrated HisTrap™ FF column, and the cleaved protein eluted in the flow-through and with buffer A containing 15 mM imidazole. The protein was further purified by gel-filtration chromatography on a Superdex75 HR 10/60 or Superose6 HR 10/30 column preequilibrated with buffer A. Fractions were analyzed on 15% SDS-PAGE gels, and those containing pure protein were concentrated using 5K cutoff spin cartridges (VivaScience) to a protein concentration between 10 and 30 mg/ml and kept at −80 °C for further study.

### Crystallization

With the use of the hanging-drop vapor-diffusion technique, single crystals of human αB ACD grew in 4 days at 16 °C in 100 mM Bicine [*N*,*N*-bis(2-hydroxyethyl)glycine], pH 9.0, and 55% 2-methyl-2,4-pentandiol at a protein concentration of 20 mg/ml using 1 μl of protein solution plus 2 μl of reservoir solution. The crystals were flash frozen in liquid nitrogen with no further cryoprotection.

Full-length rat Hsp20 grew crystals, but only after limited proteolysis. The domain boundaries were estimated from mass spectrometry, and the ACD was cloned and expressed as above. Single crystals of rat Hsp20 ACD grew at 16 °C in 100 mM Mes (4-morpholineethanesulfonic acid), pH 6.5, between 45% and 52% v/v PEG (polyethylene glycol) 200, at a protein concentration of around 10 mg/ml using 1 μl of protein solution plus 2 μl of reservoir solution. The crystals were flash frozen in liquid nitrogen with no further cryoprotection.

### Crystallography

The cell dimensions for Hsp20 ACD crystals are as follows: $a = 28.1$ Å, $b = 37.96$ Å, $c = 56.04$ Å and $\beta = 91.47°$. Data were collected at a wavelength of 0.9757 Å; the $R_{merge}$ for all data was 7%, $R_{merge}$(1.38–1.3) was 21.5%, $I/\sigma I$ was 14.6 and $I/\sigma I$(1.38–1.3) was 4.2. Multiplicity was 3.8 with completeness of 94% up to 1.3 Å. The cell dimensions for αB ACD crystals are as follows: $a = 49.46$ Å, $b = 66.07$ Å, $c = 190.97$ Å and $\beta = 92.68°$. Diffraction from the crystals was anisotropic and collected at 1.0066 Å. The $R_{merge}$ was 17.1, $R_{merge}$(3.1 2.9) was 60%, $I/\sigma I$ was 9.6 and $I/\sigma I$(3.1–2.9) was 5.3. Multiplicity was 3.4 with completeness of 99%.

All data were integrated with iMosflm[49] and scaled with Scala[50] from the CCP4 program suite.[51] The rat Hsp20 ACD structure was solved using MrBump[52] and Phaser.[53] The successful model was derived from the coordinates for PDB structure 1gme. Following molecular replacement, the structure had an *R*-factor of 49.4% and an *R*-free of 50.3. Thirty-five cycles of automated building in ARP/wARP[54] resulted in a single polypeptide chain with 89 residues correctly placed. Further rounds of building, water placement and refinement with Phenix[55] reduced *R*-factor and *R*-free to 14.7% and 17.6%, respectively. The model has good geometry as defined by MolProbity[56] stereochemical checks. The human αB ACD structure was solved with Phaser[53] using a partially refined Hsp20 structure as a model. Initially, four chains were identified in the asymmetric unit, resulting in an *R*-factor and an *R*-free of 43.6% and 49.5%, respectively. After several rounds of refinement, it was possible to identify a disordered fifth molecule in electron density maps. Final refined *R*-factor and *R*-free values are 25.6% and 32%, respectively, with good geometry as defined by MolProbity checks.

### Solution dimer experiments

The molecular weight of human αB ACD was determined by sedimentation velocity at 20 °C in a Beckman Optima XL-I analytical ultracentrifuge at a rotor speed of 42,000 r.p.m. (An50Ti rotor) and double-sector cells (12-mm Epon charcoal-filled centre pieces) loaded with 395 μl of sample in the sample chamber and 405 μl of buffer in the reference chamber. Boundary detection was carried out using absorption optics set at 280 nm. The protein sample in 25 mM Tris, pH 8.5, and 200 mM NaCl with an absorption at 280 nm of 0.15 in a 1-cm path-length cell (1 mg/ml) was estimated to be 20.5 kDa for 75% of the signal and therefore predominantly a dimer ($s_{20,w} = 1.86$, $f/f_o = 1.46$). Size estimation of human αB ACD or human αB R120G ACD by gel-filtration chromatography over a range of pH levels was carried out using separate 50-μl samples at an absorption at 280 nm of 4 (28 mg/ml), loaded on a Superdex75 10/300 GL (GE Healthcare) preequilibrated with 25 mM Tris, pH 8.5, and 200 mM NaCl,

20 mM Kphosphate, pH 7.2, and 200 mM NaCl or 20 mM Mes, pH 6.0, and 200 mM NaCl (Fig. 6).

### Accession numbers

Coordinates and structure factors have been deposited in the PDB. For rat Hsp20 ACD, the PDB ID code is 2wj5 for the coordinate entry, and the code is r2wj5sf for the structure factors. For human αB ACD, the PDB ID code is 2wj7 for the coordinate entry, and the code is r2wj7sf for the structure factors.

## Acknowledgements

We gratefully acknowledge the financial support of the Medical Research Council of UK. We are grateful to the synchrotron beamline staff of ID14 at the European Synchrotron Radiation Facility (Grenoble, France) and I04 at Diamond Light Source (Oxfordshire, UK). We thank Dr. Andrew Purkiss for discussion and Drs. Bob Sarra and Tina Daviter (Birkbeck) for ultracentrifugation experiments.

## References

1. Lindquist, S. (1986). The heat-shock response. *Annu. Rev. Biochem.* **55**, 1151–1191.
2. Wickner, S., Maurizi, M. R. & Gottesman, S. (1999). Posttranslational quality control: folding, refolding, and degrading proteins. *Science*, **286**, 1888–1893.
3. Mizushima, N., Levine, B., Cuervo, A. M. & Klionsky, D. J. (2008). Autophagy fights disease through cellular self-digestion. *Nature*, **451**, 1069–1075.
4. Bukau, B., Weissman, J. & Horwich, A. (2006). Molecular chaperones and protein quality control. *Cell*, **125**, 443–451.
5. van Noort, J. M. (2008). Stress proteins in CNS inflammation. *J. Pathol.* **214**, 267–275.
6. Komalavilas, P., Penn, R. B., Flynn, C. R., Thresher, J., Lopes, L. B., Furnish, E. J. *et al.* (2008). The small heat shock-related protein, HSP20, is a cAMP-dependent protein kinase substrate that is involved in airway smooth muscle relaxation. *Am. J. Physiol.: Lung Cell. Mol. Physiol.* **294**, L69–L78.
7. Hsu, A. L., Murphy, C. T. & Kenyon, C. (2003). Regulation of aging and age-related disease by DAF-16 and heat-shock factor. *Science*, **300**, 1142–1145.
8. Fan, G. C., Chu, G. X. & Kranias, E. G. (2005). Hsp20 and its cardioprotection. *Trends Cardiovasc. Med.* **15**, 138–141.
9. Kamradt, M. C., Lu, M. L., Werner, M. E., Kwan, T., Chen, F., Strohecker, A. *et al.* (2005). The small heat shock protein alpha B-crystallin is a novel inhibitor of TRAIL-induced apoptosis that suppresses the activation of caspase-3. *J. Biol. Chem.* **280**, 11059–11066.
10. Parcellier, A., Schmitt, E., Brunet, M., Hammann, A., Solary, E. & Garrido, C. (2005). Small heat shock proteins HSP27 and alpha B-crystallin: cytoprotective and oncogenic functions. *Antioxid. Redox Signal.* **7**, 404–413.
11. Arrigo, A. P., Simon, S., Gibert, B., Kretz-Remy, C., Nivon, M., Czekalla, A. *et al.* (2007). Hsp27 (HspB1) and alpha B-crystallin (HspB5) as therapeutic targets. *FEBS Lett.* **581**, 3665–3674; Special Issue SI.
12. Kappe, G., Franck, E., Verschuure, P., Boelens, W. C., Leunissen, J. A. M. & de Jong, W. W. (2003). The human genome encodes 10 alpha-crystallin-related small heat shock proteins: HspB1–10. *Cell Stress Chaperones*, **8**, 53–61.
13. Quraishe, S., Asuni, A., Boelens, W. C., O'Connoro, V. & Wyttenbach, A. (2008). Expression of the small heat shock protein family in the mouse CNS: differential anatomical and biochemical compartmentalization. *Neuroscience*, **153**, 483–491.
14. Taylor, R. P. & Benjamin, I. J. (2005). Small heat shock proteins: a new classification scheme in mammals. *J. Mol. Cell. Cardiol.* **38**, 433–444.
15. Horwitz, J. (1992). alpha-Crystallin can function as a molecular chaperone. *Proc. Natl Acad. Sci. USA*, **89**, 10449–10453.
16. van Montfort, R., Slingsby, C. & Vierling, E. (2001). Structure and function of the small heat shock protein/α-crystallin family of molecular chaperones. In *Protein Folding in the Cell* (Horwich, A., ed). *Advances in Protein Structure Series*, vol. 59, pp. 150–156. Academic Press, San Diego, CA.
17. Haslbeck, M., Franzmann, T., Weinfurtner, D. & Buchner, J. (2005). Some like it hot: the structure and function of small heat-shock proteins. *Nat. Struct. Mol. Biol.* **12**, 842–846.
18. Mao, Y. W., Liu, J. P., Xiang, H. & Li, D. W. C. (2004). Human alpha A- and alpha B-crystallins bind to Bax and Bcl-X-S to sequester their translocation during staurosporine-induced apoptosis. *Cell Death Differ.* **11**, 512–526.
19. Stegh, A. H., Kesari, S., Mahoney, J. E., Jenq, H. T., Forloney, K. L., Protopopov, A. *et al.* (2008). Bcl2L12-mediated inhibition of effector caspase-3 and caspase-7 via distinct mechanisms in glioblastoma. *Proc. Natl Acad. Sci. USA*, **105**, 10703–10708.
20. Carra, S., Seguin, S. J., Lambert, H. & Landry, J. (2008). HspB8 chaperone activity toward poly(Q)-containing proteins depends on its association with Bag3, a stimulator of macroautophagy. *J. Biol. Chem.* **283**, 1437–1444.
21. Horwitz, J. (2009). Alpha crystallin: the quest for a homogeneous quaternary structure. *Exp. Eye Res.* **88**, 190–194.
22. Benesch, J. L. P., Ayoub, M., Robinson, C. V. & Aquilina, J. A. (2008). Small heat shock protein activity is regulated by variable oligomeric substructure. *J. Biol. Chem.* **283**, 28513–28517.
23. Bova, M. P., Ding, L. L., Horwitz, J. & Fung, B. K. K. (1997). Subunit exchange of alpha A-crystallin. *J. Biol. Chem.* **272**, 29511–29517.
24. Shashidharamurthy, R., Koteiche, H. A., Dong, J. H. & Mchaourab, H. S. (2005). Mechanism of chaperone function in small heat shock proteins—dissociation of the Hsp27 oligomer is required for recognition and binding of destabilized T4 lysozyme. *J. Biol. Chem.* **280**, 5281–5289.
25. Rogalla, T., Ehrnsperger, M., Preville, X., Kotlyarov, A., Lutsch, G., Ducasse, C. *et al.* (1999). Regulation of Hsp27 oligomerization, chaperone function, and protective activity against oxidative stress tumor necrosis factor alpha by phosphorylation. *J. Biol. Chem.* **274**, 18947–18956.
26. Franzmann, T. M., Menhorn, P., Walter, S. & Buchner, J. (2008). Activation of the chaperone Hsp26 is controlled by the rearrangement of its thermosensor domain. *Mol. Cell*, **29**, 207–216.

27. de Jong, W. W., Caspers, G. J. & Leunissen, J. A. M. (1998). Genealogy of the alpha-crystallin—small heat-shock protein superfamily. *Int. J. Biol. Macromol.* **22**, 151–162.
28. Kim, K. K., Kim, R. & Kim, S. H. (1998). Crystal structure of a small heat-shock protein. *Nature*, **394**, 595–599.
29. van Montfort, R. L. M., Basha, E., Friedrich, K. L., Slingsby, C. & Vierling, E. (2001). Crystal structure and assembly of a eukaryotic small heat shock protein. *Nat. Struct. Biol.* **8**, 1025–1030.
30. Stamler, R., Kappe, G., Boelens, W. & Slingsby, C. (2005). Wrapping the alpha-crystallin domain fold in a chaperone assembly. *J. Mol. Biol.* **353**, 68–79.
31. Feil, I. K., Malfois, M., Hendle, J., van der Zandt, H. & Svergun, D. I. (2001). A novel quaternary structure of the dimeric alpha-crystallin domain with chaperone-like activity. *J. Biol. Chem.* **276**, 12024–12029.
32. Jehle, S., van Rossum, B., Stout, J. R., Noguchi, S. M., Falber, K., Rehbein, K. *et al.* (2009). αB-Crystallin: a hybrid solid-state/solution-state NMR investigation reveals structural aspects of the heterogeneous oligomer. *J. Mol. Biol.* **385**, 1481–1497.
33. Giese, K. C., Basha, E., Catague, B. Y. & Vierling, E. (2005). Evidence for an essential function of the N terminus of a small heat shock protein *in vivo*, independent of *in vitro* chaperone activity. *Proc. Natl Acad. Sci. USA*, **102**, 18896–18901.
34. Ackerley, S., James, P. A., Kalli, A., French, S., Davies, K. E. & Talbot, K. (2006). A mutation in the small heat-shock protein HSPB1 leading to distal hereditary motor neuronopathy disrupts neurofilament assembly and the axonal transport of specific cellular cargoes. *Hum. Mol. Genet.* **15**, 347–354.
35. Litt, M., Kramer, P., la Morticella, D. M., Murphey, W., Lovrien, E. W. & Weleber, R. G. (1998). Autosomal dominant congenital cataract associated with a missense mutation in the human alpha crystallin gene CRYAA. *Hum. Mol. Genet.* **7**, 471–474.
36. Vicart, P., Caron, A., Guicheney, P., Li, Z., Prevost, M. C., Faure, A. *et al.* (1998). A missense mutation in the alphaB-crystallin chaperone gene causes a desmin-related myopathy. *Nat. Genet.* **20**, 92–95.
37. Simon, S., Michiel, M., Skouri-Panet, F., Lechaire, J. P., Vicart, P. & Tardieu, A. (2007). Residue R120 is essential for the quaternary structure and functional integrity of human alpha B-crystallin. *Biochemistry*, **46**, 9605–9614.
38. Houlden, H., Laura, M., Wavrant-De Vrieze, F., Blake, J., Wood, N. & Reilly, M. M. (2008). Mutations in the HSP27 (HSPB1) gene cause dominant, recessive, and sporadic distal HMN/CMT type 2. *Neurology*, **71**, 1660–1668.
39. Simon, S., Fontaine, J. M., Martin, J. L., Sun, X., Hoppe, A. D., Welsh, M. J. *et al.* (2007). Myopathy-associated alpha B-crystallin mutants—abnormal phosphorylation, intracellular location, and interactions with other small heat shock proteins. *J. Biol. Chem.* **282**, 34276–34287.
40. Rajasekaran, N. S., Connell, P., Christians, E. S., Yan, L. J., Taylor, R. P., Orosz, A. *et al.* (2007). Human alpha B-crystallin mutation causes oxido-reductive stress and protein aggregation cardiomyopathy in mice. *Cell*, **130**, 427–439.
41. Ousman, S. S., Tomooka, B. H., van Noort, J. M., Wawrousek, E. F., O'Connor, K. C., Hafler, D. A. *et al.* (2007). Protective and therapeutic role for alpha B-crystallin in autoimmune demyelination. *Nature*, **448**, 474–479.
42. Credle, J. J., Finer-Moore, J. S., Papa, F. R., Stroud, R. M. & Walter, P. (2005). On the mechanism of sensing unfolded protein in the endoplasmic reticulum. *Proc. Natl Acad. Sci. USA*, **102**, 18773–18784.
43. Preville, X., Salvemini, F., Giraud, S., Chaufour, S., Paul, C., Stepien, G. *et al.* (1999). Mammalian small stress proteins protect against oxidative stress through their ability to increase glucose-6-phosphate dehydrogenase activity and by maintaining optimal cellular detoxifying machinery. *Exp. Cell Res.* **247**, 61–78.
44. Chernik, I. S., Seit-Nebi, A. S., Marston, S. B. & Gusev, N. B. (2007). Small heat shock protein Hsp20 (HspB6) as a partner of 14-3-3gamma. *Mol. Cell. Biochem.* **295**, 9–17.
45. Dreiza, C. M., Brophy, C. M., Komalavilas, P., Furnish, E. J., Joshi, L., Pallero, M. A. *et al.* (2005). Transducible heat shock protein 20 (HSP20) phosphopeptide alters cytoskeletal dynamics. *FASEB J.* **19**, 261–263.
46. Pasta, S. Y., Raman, B., Ramakrishna, T. & Rao, C. M. (2003). Role of the conserved SRLFDQFFG region of alpha crystallin, a small heat shock protein: effect on oligomeric size, subunit exchange and chaperone-like activity. *J. Biol. Chem.* **278**, 51159–51166.
47. Rubinsztein, D. C. (2006). The roles of intracellular protein-degradation pathways in neurodegeneration. *Nature*, **443**, 780–786.
48. Sitterding, S. M., Wiseman, W. R., Schiller, C. L., Luan, C., Chen, F., Moyano, J. V. *et al.* (2008). αB-Crystallin: a novel marker of invasive basal-like and metaplastic breast carcinomas. *Ann. Diagn. Pathol.* **12**, 33–40.
49. Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Jt. CCP4/ESF-EAMCB Newsl. Protein Crystallogr.* **26**.
50. Evans, P. R. (2005). Scaling and assessment of data quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **62**, 72–82.
51. Collaborative Computational Project, No. 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50**, 760–763.
52. Keegan, R. M. & Winn, M. D. (2007). Automated search-model discovery and preparation for structure solution by molecular replacement. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **63**, 447–457.
53. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.
54. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols*, **3**, 1171–1179.
55. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W. *et al.* (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **58**, 1948–1954.
56. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X. *et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383.